Open access • Journal Article • DOI:10.1137/S0036144596300773

Computing an Eigenvector with Inverse Iteration — Source link

Ilse C. F. Ipsen

Published on: 01 Jun 1997 - Siam Review (Society for Industrial and Applied Mathematics)

Topics: Inverse iteration, Power iteration, Rayleigh quotient iteration, Fixed-point iteration and Arnoldi iteration

COMPUTING AN EIGENVECTOR WITH INVERSE ITERATION

ILSE C. F. IPSEN*

Abstract. The purpose of this paper is two-fold: to analyse the behaviour of inverse iteration for computing a single eigenvector of a complex, square matrix; and to review Jim Wilkinson's contributions to the development of the method. In the process we derive several new results regarding the convergence of inverse iteration in exact arithmetic.

In the case of normal matrices we show that residual norms decrease strictly monotonically. For eighty percent of the starting vectors a single iteration is enough.

In the case of non-normal matrices, we show that the iterates converge asymptotically to an invariant subspace. However the residual norms may not converge. The growth in residual norms from one iteration to the next can exceed the departure of the matrix from normality. We present an example where the residual growth is exponential in the departure of the matrix from normality. We also explain the often significant regress of the residuals after the first iteration: it occurs when the non-normal part of the matrix is large compared to the eigenvalues of smallest magnitude. In this case computing an eigenvector with inverse iteration is exponentially ill-conditioned (in exact arithmetic).

We conclude that the behaviour of the residuals in inverse iteration is governed by the departure of the matrix from normality, rather than by the conditioning of a Jordan basis or the defectiveness of eigenvalues.

Key words. eigenvector, invariant subspace, inverse iteration, departure from normality, ill-conditioned linear system

AMS subject classification. 15A06, 15A18, 15A42, 65F15

1. Introduction. Inverse Iteration was introduced by Helmut Wielandt in 1944 [56] as a method for computing eigenfunctions of linear operators. Jim Wilkinson turned it into a viable numerical method for computing eigenvectors of matrices. At present it is the method of choice for computing eigenvectors of matrices when approximations to one or several eigenvalues are available. It is frequently used in structural mechanics, for instance, to determine extreme eigenvalues and corresponding eigenvectors of Hermitian positive-(semi)definite matrices [2, 3, 20, 21, 28, 48].

Suppose we are given a real or complex square matrix A and an approximation $\hat{\lambda}$ to an eigenvalue of A. Inverse iteration generates a sequence of vectors x_k from a given starting vector x_0 by solving the systems of linear equations

$$(A - \hat{\lambda}I)x_k = s_k x_{k-1}, \qquad k \geq 1.$$

Here I is the identity matrix, and s_k is a positive number responsible for normalising x_k. If everything goes well, the sequence of iterates x_k converges to an eigenvector associated with an eigenvalue closest to $\hat{\lambda}$. In exact arithmetic, inverse iteration amounts to applying the power method to $(A - \hat{\lambda}I)^{-1}$.

The importance of inverse iteration is illustrated by three quotes from Wilkinson[1]:

> 'In our experience, inverse iteration has proved to be the most successful of methods we have used for computing eigenvectors of a tri-diagonal matrix from accurate eigenvalues.'

* Center for Research in Scientific Computation, Department of Mathematics, North Carolina State University, P. O. Box 8205, Raleigh, NC 27695-8205, USA (ipsen@math.ncsu.edu). This research was supported in part by NSF grants CCR-9102853 and CCR-9400921.

[1] [60, §III.54], [63, p 173], [45, §1]

> 'Inverse iteration is one of the most powerful tools in numerical analysis.'
> 'Inverse iteration is now the most widely used method for computing eigenvectors corresponding to selected eigenvalues which have already been computed more or less accurately.'

A look at software in the public domain shows that this is still true today [1, 44, 47].

The purpose of this paper is two-fold: to analyse the behaviour of inverse iteration; and to review Jim Wilkinson's contributions to the development of inverse iteration. Although inverse iteration looks like a deceptively simple process, its behaviour is subtle and counter-intuitive, especially for non-normal (e.g. non-symmetric) matrices. It is important to understand the behaviour of inverse iteration in exact arithmetic, for otherwise we cannot develop reliable numerical software. Fortunately, as Wilkinson recognised already [45, p 355], the idiosyncrasies of inverse iteration originate from the mathematical process rather than from finite precision arithmetic. This means a numerical implementation of inverse iteration in finite precision arithmetic does not behave very differently from the exact arithmetic version. Therefore we can learn a lot about a numerical implementation by studying the theory of inverse iteration. That's what we'll do in this paper.

We make two assumptions in our discussion of inverse iteration. First, the shift $\hat{\lambda}$ remains fixed for each eigenvector during the course of the iterations; this excludes variants of inverse iteration such as Rayleigh quotient iteration[2], and the interpretation of inverse iteration as a Newton method[3]. Second, only a single eigenvector is computed as opposed to a basis for an invariant subspace[4].

To illustrate the additional difficulties in the computation of several vectors, consider a real symmetric matrix. When the eigenvalues under consideration are well-separated, inverse iteration computes numerically orthogonal eigenvectors. But when the eigenvalues are poorly separated, it may not be clear with which eigenvalue a computed eigenvector is to be affiliated. Hence one perturbs the eigenvalues to enforce a clear separation. Even then the computed eigenvectors can be far from orthogonal. Hence one orthogonalises the iterates against previously computed iterates to enforce orthogonality. But explicit orthogonalisation is expensive and one faces a trade-off between orthogonality and efficiency. Therefore one has to decide which eigenvalues to perturb and by how much; and which eigenvectors to orthogonalise and against how many previous ones and how often. One also has to take into account that the tests involved in the decision process can be expensive. A detailed discussion of these issues can be found for instance in [44, 8].

1.1. Motivation. This paper grew out of commentaries about Wielandt's work on inverse iteration [26] and the subsequent development of the method [27]. Several reasons motivated us to take a closer look at Wilkinson's work. Among all contributions to inverse iteration, Wilkinson's are by far the most extensive and the most important. They are contained in five papers[5] and in chapters of his two books [60, 61]. Since four of the five papers were published after the books, there is no one place where all of his results are gathered. Overlap among the papers and books, and the gradual development of ideas over several papers makes it difficult to realise what Wilkinson

[2] [11, §5.4], [41, §§4.6-9], [40], [46, §IV.1.3] [62, §3]
[3] [10, §3], [11, §5.9], [38, §2, §3], [45, §4]
[4] [10, 29, 43, 8]
[5] [44, 45, 57, 62, 63]

has accomplished. Therefore we decided to compile and order his main results and to set out his ideas.

Wilkinson's numerical intuition provided him with many insights and empirical observations for which he did not provide rigorous proofs. To put his ideas on a theoretical footing, we extend his observations and supply simple proofs from first principles. To this end it is necessary to clearly distinguish the mathematical properties of inverse iteration from finite precision issues. The importance of this distinction was first realised in Shiv Chandrasekaran's thesis [8] where a thorough analysis of inverse iteration is presented for the computation of a complete set of eigenvectors of a real, symmetric matrix.

1.2. Caution. Our primary means for analysing the convergence of inverse iteration is the backward error rather than the forward error. The backward error for iterate x_k is $\|r_k\|$, where $r_k = (A - \hat{\lambda}I)x_k$ is the residual. When $\|r_k\|$ is small then x_k and $\hat{\lambda}$ are an eigenpair of a matrix close to A (§2.3). In contrast, the forward error measures how close x_k is to an eigenvector of A. We concentrate on the backward error because

$$\|r_k\| = \|(A - \hat{\lambda}I)x_k\| = \|s_k x_{k-1}\| = s_k$$

is the only readily available computational means for monitoring the progress of inverse iteration.

Unfortunately, however, a small backward error does not imply a small forward error. In the case of normal matrices, for instance, a measure of the forward error is the acute angle θ_k between x_k and the eigenspace associated with all eigenvalues closest to $\hat{\lambda}$. The resulting $\sin\theta$ theorem [12, §2], [14, §4] bounds the forward error in terms of the backward error and an amplification factor γ, which is the separation between $\hat{\lambda}$ and all eigenvalues farther away:

$$\sin\theta_k \leq \|r_k\|/\gamma.$$

This means, even though the backward error may be small, x_k can be far away from the eigenspace if the eigenvalues closest to $\hat{\lambda}$ are poorly separated from those remaining.

Nevertheless, in the absence of information about the eigenvalue distribution of A the only meaningful computational pursuit is a small backward error. If the backward error is small then we can be certain, at least, that we have solved a nearby problem. Therefore we concentrate our efforts on analysing the behaviour of successive residuals, and on finding out under what conditions a residual is small.

1.3. New Results. To our knowledge, the following observations and results are new.

Normal Matrices. Residual norms decrease strictly monotonically (§3.3). For eighty percent of the starting vectors a single iteration is enough, because the residual is as small as the accuracy of the computed eigenvalue (§3.1).

Diagonalisable Matrices. Inverse iteration distinguishes between eigenvectors, and vectors belonging to an invariant subspace that are not eigenvectors (§4.4). The square root of the residual growth is a lower bound for an eigenvector condition number (§4.3).

Non-Normal Matrices[6]**.** For every matrix one can find iterates where the residual growth from one iteration to the next is at least as large as the departure of the

[6] Non-normal matrices include diagonalisable as well as defective matrices.

matrix from normality; and one can also find iterates where the residual growth is at least as large as the departure of the inverse matrix from normality (§5.3).

We introduce a measure for the *relative* departure of a matrix from normality by comparing the size of the non-normal part to the eigenvalues of smallest magnitude (§5.2). There are matrices whose residual growth can be exponential in the relative departure from normality (§5.4). This explains the often significant regress of inverse iteration after the first iteration.

Increasing the accuracy of the approximate eigenvalue $\hat{\lambda}$ increases the relative departure of $A - \hat{\lambda}I$ from normality. When the relative departure from normality exceeds one, computing an eigenvector of A with inverse iteration is exponentially ill-conditioned (§5.2).

We conclude that the residual growth in inverse iteration is governed by the departure of the matrix from normality, rather than by the conditioning of a Jordan basis or the defectiveness of eigenvalues (§5.2).

1.4. Overview. In §2 we discuss the basic aspects of inverse iteration: the underlying idea (§2.1); how to solve the linear system (§2.2); the purpose of the residual (§2.3, §2.4); and the choice of starting vectors (§2.5, §2.6).

In §3 we exhibit the good behaviour of inverse iteration in the presence of a normal matrix: one iteration usually suffices (§3.1); and the residuals decrease strictly monotonically (§3.2, §3.3).

In §4 we show what happens when inverse iteration is applied to a diagonalisable matrix: residuals can grow with the ill-conditioning of the eigenvectors (§4.1, §4.3); the accuracy of the approximate eigenvalue can exceed the size of the residual (§4.2); and residuals do not always reveal when the iterates have arrived at their destination (§4.4).

In §5 we describe the behaviour of inverse iteration in terms of the departure from normality: upper and lower bounds on the residual growth (§5.2, §5.3); an example of exponential residual growth (§5.4); and the relation of the residual to conditioning of a Jordan basis and defectiveness of eigenvalues (§5.1).

In §6 we illustrate that inverse iteration in finite precision arithmetic behaves very much like in exact arithmetic. We examine the effects of finite precision arithmetic on the residual size (§6.1), the performance of starting vectors (§6.2), and the solution of linear systems (§6.3). A short description of a numerical software implementation (§6.4) concludes the chapter.

In §7 we prove the convergence of inverse iteration in exact arithmetic.

In §8 (Appendix 1) we supply facts about Jordan matrices required in §5; and in §9 (Appendix 2) we present relations between different measures of departure from normality.

1.5. Notation. Our protagonist is a real or complex $n \times n$ matrix A with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. When the matrix is normal or, in particular Hermitian, we call it H.

Our goal is to compute an eigenvector of A. But since we have access only to an approximate eigenvalue $\hat{\lambda}$, the best we can get is an approximate eigenvector $\hat{x}$. We use a hat to represent approximate or computed quantities, such as $\hat{\lambda}$ and $\hat{x}$. We assume most of the time that $\hat{\lambda}$ is not an eigenvalue of A, hence $A - \hat{\lambda}I$ is nonsingular, where I (or I_n) is the identity matrix of order n. If, on the contrary, $\hat{\lambda}$ is an eigenvalue of A then the problem is easy to solve because we only have to compute a null vector of $A - \hat{\lambda}I$.

The norm $\|\cdot\|$ is the two-norm, i.e. $\|x\| = \sqrt{x^*x}$, where the superscript * denotes the conjugate transpose. The ith column of the identity matrix I is called the *canonical vector* e_i, $i \geq 1$.

2. The Method. In this section we describe the basics of inverse iteration: the idea behind it, solution of the linear systems, rôle of the residual, and choice of starting vectors.

2.1. Wilkinson's Idea. Wilkinson had been working on and talking about inverse iteration as early as 1957 [39]. In those days it was believed that inverse iteration was doomed to failure because of its main feature: the solution of an ill-conditioned system of linear equations [38, p 22] (see also §6.3). In a radical departure from conventional wisdom, Wilkinson made inverse iteration work[7] [38, §1].

Although Wilkinson gives the credit for inverse iteration to Wielandt [45, 60, 62], he himself presented it in 1958 [57] as a result of trying to improve Givens's method for computing a single eigenvector of a symmetric tridiagonal matrix (a different improvement of Givens's method is discussed in [42]).

We modify Wilkinson's idea slightly and present it for a general complex matrix A rather than for a real symmetric tridiagonal matrix. His idea is the following: If $\hat{\lambda}$ is an eigenvalue of the $n \times n$ matrix A, then $A - \hat{\lambda}I$ is singular, and $n-1$ equations from $(A - \hat{\lambda}I)\hat{x} = 0$ determine, up to a scalar multiple, an eigenvector associated with $\hat{\lambda}$. However, if $\hat{\lambda}$ is not an eigenvalue of A, $A - \hat{\lambda}I$ is non-singular and the only solution to $(A - \hat{\lambda}I)\hat{x} = 0$ is zero. To get a non-zero approximation to an eigenvector, select a non-zero vector y that solves $n-1$ equations from $(A - \hat{\lambda}I)\hat{x} = 0$.

Why should y be a good approximation to an eigenvector? First consider the case where y solves the leading $n-1$ equations. Partition A so as to distinguish its leading principal submatrix A_1 of order $n-1$, and partition y conformally,

$$A \equiv \begin{pmatrix} A_1 & a_1 \\ a_2 & \alpha \end{pmatrix}, \qquad y \equiv \begin{pmatrix} y_1 \\ \upsilon \end{pmatrix}.$$

One can always find a number υ so that the smaller system

$$(A_1 - \hat{\lambda}I)y_1 = -\upsilon a_1$$

has a non-zero solution y_1: If $\hat{\lambda}$ is an eigenvalue of A_1, set $\upsilon = 0$. Since $A_1 - \hat{\lambda}I$ is singular, $(A_1 - \hat{\lambda}I)y_1 = 0$ has a non-zero solution. If $\hat{\lambda}$ is not an eigenvalue of A_1, set υ to some non-zero value. Since $A_1 - \hat{\lambda}I$ is non-singular, the system has a non-zero solution. Therefore there exists a non-zero vector y that solves the leading $n-1$ equations of the larger system $(A - \hat{\lambda}I)\hat{x} = 0$. Including the last equation and setting

$$\beta \equiv \begin{pmatrix} a_2 & \alpha - \hat{\lambda} \end{pmatrix} y$$

implies that y is a non-zero solution to

$$(A - \hat{\lambda}I)y = \beta e_n,$$

where the canonical vector e_n is the nth column of the identity matrix.

Suppose λ is an eigenvalue of A and x is an associated eigenvector. If βe_n contains a contribution of x, then this contribution is multiplied in y by $1/(\lambda - \hat{\lambda})$. Moreover, if λ is a simple eigenvalue that is closer to $\hat{\lambda}$ than any other eigenvalue λ_j then

$$\frac{1}{|\lambda - \hat{\lambda}|} > \frac{1}{|\lambda_j - \hat{\lambda}|}.$$

[7] I am grateful to Michael Osborne for providing this historical context.

For a diagonalisable matrix A this means the contribution of the eigenvector x in y is amplified by a larger amount than the contributions of all other eigenvectors. In this case y is closer to x than is βe_n.

Instead of solving the leading $n-1$ equations one can solve any set of $n-1$ equations. Omitting the ith equation leads to a multiple of e_i as the right-hand side (for real, symmetric tridiagonal matrices, Parlett and Dhillon [42] select the doomed equation according to the pivots from a pair of 'complementary' triangular factorisations). In general, the right-hand side can be any vector x_0, as long as it contains a contribution of the desired eigenvector x so that the solution x_1 of $(A-\hat{\lambda}I)x_1 = x_0$ is closer to x than is x_0.

Therefore, if there is only a single, simple eigenvalue λ closest to $\hat{\lambda}$ and if A is diagonalisable then the iterates x_k of inverse iteration converge to (multiples of) an eigenvector x associated with λ, provided x_0 contains a contribution of x. A more detailed convergence proof is given in §7.

2.2. Solution of the Linear System. In practice, one first solves the linear system before normalising the iterate. Given a scalar $\hat{\lambda}$ so that $A-\hat{\lambda}I$ is non-singular, and a vector x_0 with $\|x_0\| = 1$, perform the following iterations for $k \geq 1$:

$$\begin{aligned} (A-\hat{\lambda}I)z_k &= x_{k-1} \\ x_k &= z_k/\|z_k\| \end{aligned}$$

Here z_k is the unnormalised iterate. The corresponding normalised iterate satisfies $\|x_k\| = 1$. Hence the normalisation constant is $s_k = 1/\|z_k\|$.

Already in his very first paper on inverse iteration [57, pp 92-3] Wilkinson advocated that the linear system $(A-\hat{\lambda}I)z_k = x_{k-1}$ be solved by Gaussian elimination with partial pivoting,

$$P(A-\hat{\lambda}I) = LU, \qquad Lc_k = Px_{k-1}, \qquad Uz_k = c_k,$$

where P is a permutation matrix, L is unit lower triangular and U is upper triangular. Since the matrix $A-\hat{\lambda}I$ does not change during the course of the iteration, the factorisation is performed only once and the factors L and U are used in all iterations.

For reasons to be explained in §2.6, Wilkinson chooses c_1 equal to e, a multiple of the vector of all ones[8]. This amounts to the implicit choice $x_0 \equiv P^T Le$ for the starting vector. It saves the lower triangular system solution in the very first inverse iteration. The resulting algorithm is

		Factor $P(A-\hat{\lambda}I) = LU$
Iteration	1	Set $c_1 \equiv e$
		Solve $Uz_1 = c_1$
		Set $x_1 \equiv z_1/\|z_1\|$
Iteration	$k \geq 2$	Solve $Lc_k = Px_{k-1}$
		Solve $Uz_k = c_k$
		Set $x_k \equiv z_k/\|z_k\|$

To reduce the operation count for the linear system solution, the matrix is often reduced to Hessenberg form or, in the real symmetric case, to tridiagonal form before the start of inverse iteration [1]. While Gaussian elimination with partial pivoting requires $O(n^3)$ arithmetic operations for a general $n \times n$ matrix, it requires merely $O(n^2)$

[8] [44, p 435], [57, pp 93-94], [60, §III.54], [61, §5.54, §9.54], [62, p 373]

operations for a Hessenberg matrix [61, 9.54], and $O(n)$ operations for an Hermitian tridiagonal matrix [57], [60, §III.51].

2.3. An Iterate and Its Residual. Wilkinson showed[9] that the residual

$$r_k \equiv (A - \hat{\lambda} I) x_k$$

is a measure for the accuracy of the shift $\hat{\lambda}$ and of the iterate x_k. Here we present Wilkinson's argument for exact arithmetic; and in §6.1 we discuss it in the context of finite precision arithmetic.

From

$$r_k = (A - \hat{\lambda} I) x_k = \frac{1}{\|z_k\|} (A - \hat{\lambda} I) z_k = \frac{1}{\|z_k\|} x_{k-1}$$

and $\|x_{k-1}\| = 1$ follows

$$\|r_k\| = 1/\|z_k\|.$$

Therefore the residual is inversely proportional to the norm of the unnormalised iterate[10]. Since $\|z_k\|$ is required for the computation of x_k, the size of the residual comes for free in inverse iteration. It is used as a criterion for terminating the iterations. Once the residual is small enough, inverse iteration stops because then $\hat{\lambda}$ and x_k are an eigenpair of a nearby matrix:

THEOREM 2.1 (§15 IN [18]). *Let A be a complex, square matrix, and let $r_k = (A - \hat{\lambda} I) x_k$ be the residual for some number $\hat{\lambda}$ and vector x_k with $\|x_k\| = 1$.*

Then there is a matrix E_k with $(A + E_k - \hat{\lambda} I) x_k = 0$ and $\|E_k\| \leq \mu$ if and only if $\|r_k\| \leq \mu$.

Proof. Suppose $(A + E_k - \hat{\lambda} I) x_k = 0$ and $\|E_k\| \leq \mu$. Then

$$r_k = (A - \hat{\lambda} I) x_k = -E_k x_k$$

implies $\|r_k\| \leq \|E_k\| \leq \mu$.

Now suppose $\|r_k\| \leq \mu$. Then

$$(A - \hat{\lambda} I) x_k = r_k = r_k x_k^* x_k,$$

implies

$$(A + E_k - \hat{\lambda} I) x_k = 0, \qquad \text{where} \qquad E_k = -r_k x_k^*.$$

Since E_k has rank one,

$$\|E_k\| = \|r_k\| \leq \mu.$$

□

Thus, a small residual implies that $\hat{\lambda}$ and x_k are accurate in the backward sense.

2.4. The Smallest Possible Residual. Since a small residual r_k indicates that the iterate x_k is an eigenvector of a matrix close to A, we would like to know how small a residual can possibly be.

From the definition of the residual $r_k = (A - \hat{\lambda} I) x_k$ and the fact that

$$\|Mx\| \geq \|x\| / \|M^{-1}\|$$

for any non-singular matrix M, it follows that

$$\|r_k\| \geq 1/\|(A - \hat{\lambda} I)^{-1}\|.$$

[9] [18, §15], [44, pp 419-420], [45, p 356], [60, §49], [61, §3.53], [62, p 371]

[10] [44, p 420], [45, p 352], [60, §III.52], [61, §5.55], [62, p 372], [63, p 176]

Thus, when $A - \hat{\lambda}I$ is almost singular, $\|(A - \hat{\lambda}I)^{-1}\|$ is large and the residual can be very small.

The lower bound for $\|r_k\|$ is attained when x_k is a right singular vector associated with the smallest singular value of $A - \hat{\lambda}I$ [45, p 358]. Denoting by u the corresponding left singular vector gives

$$(A - \hat{\lambda}I)x_k = \frac{u}{\|(A - \hat{\lambda}I)^{-1}\|}$$

and the residual has minimal norm $\|r_k\| = 1/\|(A - \hat{\lambda}I)^{-1}\|$. The starting vector x_0 that gives the smallest possible residual after one iteration is therefore a left singular vector associated with the smallest singular value of $A - \hat{\lambda}I$.

Different Interpretations. The bound $\|r_k\| \geq 1/\|(A - \hat{\lambda}I)^{-1}\|$ has several different meanings.

One interpretation is based on the *separation*[11] *between two matrices* A and B,

$$\mathrm{sep}(A, B) \equiv \min_{\|X\|=1} \|AX - XB\|.$$

In the special case when $B \equiv \hat{\lambda}$ is a scalar, the separation between A and $\hat{\lambda}$ is [55, §2]

$$\mathrm{sep}(A, \hat{\lambda}) = \min_{\|x\|=1} \|Ax - x\hat{\lambda}\| = \min_{\|x\|=1} \|(A - \hat{\lambda}I)x\| = 1/\|(A - \hat{\lambda}I)^{-1}\|.$$

Thus, the residual is a lower bound for the separation between the matrix A and the shift $\hat{\lambda}$,

$$\|r_k\| \geq \mathrm{sep}(A, \hat{\lambda}).$$

A second interpretation is based on the *resolvent* of A at the point $\hat{\lambda}$ [33, §I.5.2],

$$\mathcal{R}(A, \hat{\lambda}) \equiv (A - \hat{\lambda}I)^{-1}.$$

Since $\|r_k\| = 1/\|z_k\|$, the iterate norm represents a lower bound on the resolvent norm at $\hat{\lambda}$,

$$\|z_k\| \leq \|\mathcal{R}(A, \hat{\lambda})\|.$$

Yet a third interpretation is based on the ϵ-*pseudospectrum*[12] of A,

$$\mathcal{S}_\epsilon(A) \equiv \left\{\lambda : \|(A - \lambda I)^{-1}\| \geq \epsilon^{-1}\right\}.$$

Thus, when $\|r_k\| \leq \epsilon$ then $\hat{\lambda}$ is contained in the ϵ-pseudospectrum of A,

$$\hat{\lambda} \in \mathcal{S}_\epsilon(A).$$

Regrettably, we found notions like separation between matrices, pseudospectrum, and resolvent of limited help in analysing inverse iteration, because they focus on the operator $(A - \hat{\lambda}I)^{-1}$ rather than on its application to a particular argument.

[11] [50, §2], [50, §4.3], [55, §1]
[12] [33, §VIII.5.1], [52], [55, §3]

2.5. Good Starting Vectors. Wilkinson's goal was to find starting vectors x_0 'so that one iteration produces a *very good eigenvector*' [62, p 372]. Little or no work should be involved in determining these starting vectors. Here we assume exact arithmetic. The finite precision case is discussed in §6.2.

Varah [54] and Wilkinson[13] showed that there exists at least one canonical vector that, when used as a starting vector x_0, gives an iterate z_1 of almost maximal norm and hence an almost minimal residual r_1. Varah's results constitute the basis for Wilkinson's argument [62, p 374].

Suppose column l, $1 \leq l \leq n$, has the largest norm among all columns of $(A - \hat{\lambda}I)^{-1}$,

$$\|(A - \hat{\lambda}I)^{-1}e_l\| \geq \frac{1}{\sqrt{n}}\|(A - \hat{\lambda}I)^{-1}\|.$$

Setting $x_0 = e_l$ gives a residual r_1 that deviates from its minimal value by a factor of at most $\sqrt{n}$,

$$\frac{1}{\|(A - \hat{\lambda}I)^{-1}\|} \leq \|r_1\| \leq \sqrt{n}\,\frac{1}{\|(A - \hat{\lambda}I)^{-1}\|},$$

and the first iterate exhibits almost maximal growth,

$$\frac{1}{\sqrt{n}}\|(A - \hat{\lambda}I)^{-1}\| \leq \|z_1\| \leq \|(A - \hat{\lambda}I)^{-1}\|.$$

Varah [54] showed that the above argument is true for any orthonormal basis, not just the canonical basis. Choose a unitary matrix W, i.e. $W^*W = WW^* = I$. If column l has the largest norm among all columns of $(A - \hat{\lambda}I)^{-1}W$ then

$$\|(A - \hat{\lambda}I)^{-1}We_l\| \geq \frac{1}{\sqrt{n}}\|(A - \hat{\lambda}I)^{-1}W\| = \|(A - \hat{\lambda}I)^{-1}\|.$$

Thus, $\|r_1\|$ is almost minimal and z_1 exhibits almost maximal growth when $x_0 = We_l$. More generally, when x_0 is a column of largest norm of $(A - \hat{\lambda}I)^{-1}W$, where W is any non-singular matrix, the upper bound on $\|r_1\|$ contains the condition number of W [54, Theorem 2].

Wilkinson's Argument. In spite of their apparent dissimilarity, Wilkinson's argument [62, pp 372-3], [45, p 353] and our argument above are basically the same.

Wilkinson argues as follows. Selecting, in turn, each canonical vector e_i as a starting vector amounts to solving the linear system

$$(A - \hat{\lambda}I)Z = I,$$

where Ze_i is the unnormalised first iterate for inverse iteration with starting vector $x_0 = e_i$. Since $\hat{\lambda}$ is an exact eigenvalue of $A + F$ for some F, $A - \hat{\lambda} + F$ is singular and $\|Z\| \geq 1/\|F\|$. If $\hat{\lambda}$ is a good approximation to an eigenvalue of A then $\|F\|$ must be small, at least one column of Z must be large, and at least one of the canonical vectors must give rise to a large first iterate.

Since $Z = (A - \hat{\lambda}I)^{-1}$, Wilkinson basically argues that $\|(A - \hat{\lambda}I)^{-1}\|$ is large if $\hat{\lambda}$ is close to an eigenvalue of A. The quantity $\|F\|$ is an indication of the distance of $A - \hat{\lambda}I$ to singularity. Thus, the norm of at least one column of Z is inversely proportional to the distance of $A - \hat{\lambda}I$ from singularity, cf. [44, p 420].

In comparison, our argument shows that an appropriate canonical vector leads to an iterate z_1 with $\|z_1\| \geq \frac{1}{\sqrt{n}}\,\|(A - \hat{\lambda}I)^{-1}\|$. As $1/\|(A - \hat{\lambda}I)^{-1}\|$ is the distance of

[13] [44, p 420], [45, p 353], [62, pp 372-3]

$A - \hat{\lambda}I$ to singularity [13, §3], this means, as in Wilkinson's case, that the magnitude of z_1 is inversely proportional to the distance of $A - \hat{\lambda}I$ from singularity.

Hence, the basis for both, Wilkinson's and our arguments is that the ideal starting vector x_0 should result in a z_1 whose size reflects the distance of $A - \hat{\lambda}I$ from singularity.

2.6. Wilkinson's Choice of Starting Vector. Wilkinson's motivation for putting so much care into the choice of starting vector x_0 is to reduce the residual r_1 of the first iterate as much as possible [62, p 374]. This is especially important for non-normal matrices because the residuals can grow significantly in subsequent iterations [44, p 420]. Although he cannot expect to find x_0 that produces a minimal r_1, he tries to select x_0 that is likely to produce a small r_1 [44, p 420].

There are certain choices of x_0 that Wilkinson rules out immediately[14]. In case of real symmetric tridiagonal matrices T for instance it is 'quite common' that a canonical vector contains only a very small contribution of an eigenvector [57, p 91]. When T contains a pair of zero off-diagonal elements, it splits into two disjoint submatrices and the eigenvectors associated with one submatrix have zero elements in the positions corresponding to the other submatrix. Hence canonical vectors with ones in the position corresponding to zero eigenvector elements are likely to produce large r_1. Wilkinson remarks [57, p 91]:

> It is clear that none of the e_i is likely to produce consistently accurate eigenvectors and, in particular, that e_1 and e_n will, in general, be the least satisfactory.

Since no one canonical vector works satisfactorily all the time, Wilkinson plays it safe and chooses a vector that represents all of them: the vector consisting of all ones. Wilkinson's choice, already alluded to in §2.2, can be justified as follows.

If A is a complex, square matrix and $\hat{\lambda}$ is not an eigenvalue of A then $A - \hat{\lambda}I$ and the upper triangular matrix U from Gaussian elimination $P(A - \hat{\lambda}I) = LU$ are non-singular. Let

$$U \equiv \begin{pmatrix} U_{n-1} & u_{n-1} \\ & u_{nn} \end{pmatrix}, \qquad U^{-1} = \begin{pmatrix} U_{n-1}^{-1} & -U_{n-1}^{-1} u_{n-1}/u_{nn} \\ & 1/u_{nn} \end{pmatrix},$$

where U_{n-1} is the leading principal submatrix of order $n-1$ and u_{nn} is the trailing diagonal element of U. If $\hat{\lambda}$ is a good approximation to an eigenvalue of A, so that $A - \hat{\lambda}I$ is almost singular, and if $|u_{nn}|$ is small then choosing $c_1 = e_n$ yields

$$z_1 = U^{-1} e_n = \frac{1}{u_{nn}} \begin{pmatrix} -U_{n-1}^{-1} u_{n-1} \\ 1 \end{pmatrix}.$$

Hence $\|z_1\| \geq 1/|u_{nn}|$ is large. This means, if the trailing diagonal element of U is small then choosing $c_1 = e_n$ results in a small residual r_1.

More generally, let U_l and U_{l-1} be the leading principal submatrices of orders l and $l-1$ of U, so

$$U_l \equiv \begin{pmatrix} U_{l-1} & u_{l-1} \\ & u_{ll} \end{pmatrix}.$$

If the lth diagonal element u_{ll} of U is small then choosing $c_1 = e_l$ gives

$$z_1 = U^{-1} e_l = \frac{1}{u_{ll}} \begin{pmatrix} -U_{l-1}^{-1} u_{l-1} \\ 1 \end{pmatrix}$$

and $\|z_1\| \geq 1/|u_{ll}|$ is large.

[14] [57, p 91], [60, III.54], [62, p 373]

However, it can happen that $A - \hat{\lambda}I$ is almost singular while no diagonal element of U is small[15]. According to §2.5, U^{-1} contains a column l whose norm is almost as large as $\|U^{-1}\|$. But l is unknown, so Wilkinson proposes[16] what he considers to be a 'very safe' choice, a multiple of the vector of all ones, i.e. $c_1 \equiv e$ [57, p 94]. In the special case of symmetric tridiagonal matrices Wilkinson remarks [61, p 322]:

> No solution to this problem has been found which can be demonstrated rigorously to be satisfactory in all cases, but [the choice $c_1 = e$] has proved to be extremely effective in practice and there are good reasons for believing that it cannot fail.

Back to general, complex matrices: Although the choice $c_1 = e$ often results in $z_1 = U^{-1}e$ with large norm, this is not guaranteed [54, p 788]. It is thus necessary to have contingency plans. Varah [54, p 788] and Wilkinson [44, §5.2] present orthogonal matrices whose first column is e, a multiple of the vector of all ones, and whose remaining columns stand by to serve as alternative starting vectors. If $c_1 = e$ does not produce a sufficiently large iterate z_1, one column after another of such a matrix takes its turn as c_1 until an iterate z_1 with sufficiently large norm appears. Wilkinson recommends this strategy 'in practice' [62, p 373] and remarks [62, pp 373-74]:

> In the majority of cases the first of these [i.e. $c_1 = e$] immediately gives an acceptable z and no case has been encountered in practice where it was necessary to proceed to the third vector [i.e. the third column of the orthogonal matrix].

In spite of all the effort devoted to choosing a starting vector x_0 that is likely to produce a small residual r_1, Wilkinson states [45, p 353], cf. also [45, p 358]:

> In fact if we take a random x_0 the probability that one iteration will produce an x_1 having a pathologically small residual is very high indeed; this would only fail to be true if we happen to choose an x_0 which is almost completely deficient in all those e_s [for which $\|(A - \hat{\lambda}I)^{-1}e_s\|$ is large].

This was confirmed more recently for real symmetric matrices by a rigorous statistical analysis [29]. In case of general matrices Wilkinson remarks [45, p 360]:

> The ordinary process of inverse iteration will almost always succeed in one iteration; if it does not do so one only has to re-start with an initial vector orthogonal to the first. This process can be continued until one reaches an initial vector which gives success in one iteration. It is rare for the first vector to fail and the average number of iterations is unlikely to be as high as 1.2.

3. Normal Matrices. In this section we examine the size of the residual and its change during the course of the iteration in the special case when the matrix is normal.

We use the following notation. Let H be a normal matrix with eigendecomposition $H = Q\Lambda Q^*$ where Q is unitary; and let

$$\epsilon \equiv \min_i |\lambda_i - \hat{\lambda}|$$

be the accuracy of the shift $\hat{\lambda}$. We assume for the most part that $\hat{\lambda}$ is not an eigenvalue of H, because otherwise a single linear system solution would suffice to produce an eigenvector, see the discussion in §2.1 and [42, §2].

[15] [54, p 788], [57, p 93], [61, §5.56, §9.54]

[16] [44, p 435], [57, p 93-94], [60, §III.54], [61, §5.54, §9.54], [62, p 373]

Wilkinson showed that for Hermitian matrices the size of the residual is constrained solely by the accuracy of the shift[17]. The Bauer-Fike theorem [4, Theorem IIIa] implies that this is true for the larger class of normal matrices as well:

THEOREM 3.1. *Let H be a normal matrix, and let $r_k = (A-\hat{\lambda}I)x_k$ be the residual for some number $\hat{\lambda}$ and vector x_k with $\|x_k\| = 1$.*

Then

$$\|r_k\| \geq \epsilon.$$

Proof. If $\hat{\lambda}$ is an eigenvalue of H then $\epsilon = 0$ and the statement is obviously true. If $\hat{\lambda}$ is not an eigenvalue of H then $H - \hat{\lambda}I$ is non-singular. Because H is normal, $\|(H - \hat{\lambda}I)^{-1}\| = 1/\epsilon$, and

$$\|r_k\| = \|(H - \hat{\lambda}I)x_k\| \geq \frac{1}{\|(H - \hat{\lambda}I)^{-1}\|} = \epsilon.$$

□

3.1. Once Is Enough. We show that in most cases the residual of a normal matrix after the first iteration is as good as the accuracy of the shift $\hat{\lambda}$.

Partition the eigendecomposition as

$$\Lambda = \begin{pmatrix} \Lambda_1 & \\ & \Lambda_2 \end{pmatrix}, \qquad Q = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix},$$

where Λ_1 contains all eigenvalues λ_i at the same, minimal distance to $\hat{\lambda}$,

$$\epsilon = \|\Lambda_1 - \hat{\lambda}I\|,$$

while Λ_2 contains the remaining eigenvalues that are further away from $\hat{\lambda}$,

$$\epsilon < \min_j |(\Lambda_2)_{jj} - \hat{\lambda}| = 1/\|(\Lambda_2 - \hat{\lambda}I)^{-1}\|.$$

This partition covers two common special cases: $\hat{\lambda}$ approximates an eigenvalue λ of multiplicity $l \geq 1$, i.e. $\Lambda_1 = \lambda I$; and $\hat{\lambda}$ approximates two distinct real eigenvalues at the same minimal distance to its left and right, i.e. $\hat{\lambda} - \lambda_1 = \lambda_2 - \hat{\lambda}$. The columns of Q_1 span the invariant subspace associated with the eigenvalues in Λ_1.

The result below states that the residual decreases with the angle between the starting vector and the desired eigenspace. It applies in all cases but the worst, when the starting vector is orthogonal to the desired invariant subspace.

THEOREM 3.2. *Let inverse iteration be applied to a normal, non-singular matrix $H-\hat{\lambda}I$; let $r_1 = (H-\hat{\lambda}I)x_1$ be the residual for the first iterate x_1; and let $0 \leq \theta \leq \pi/2$ be the angle between starting vector x_0 and* range(Q_1).

If $\theta < \pi/2$, then

$$\|r_1\| \leq \epsilon / \cos\theta.$$

Proof. Decompose x_0 into its contribution c in range(Q_1) and its contribution s outside,

$$x_0 = Q \begin{pmatrix} c \\ s \end{pmatrix}, \qquad \text{where} \qquad c \equiv Q_1^* x_0, \qquad s \equiv Q_2^* x_0.$$

Since x_0 has unit norm, $\|c\|^2 + \|s\|^2 = 1$, and $\|c\| = \cos\theta$ [17, §12.4.3].

[17] [9, §2.3], [45, §8], [59, §2], [60, §III.49]

To bound $\|r_1\|$ from above, bound $\|z_1\|$ from below, as $1/\|r_1\| = \|z_1\|$. Then

$$z_1 = Q(\Lambda - \hat{\lambda}I)^{-1} \begin{pmatrix} c \\ s \end{pmatrix}$$

implies

$$\|z_1\| \geq \|Q_1^* z_1\| = \|(\Lambda_1 - \hat{\lambda}I)^{-1} c\| \geq \frac{\|c\|}{\|\Lambda_1 - \hat{\lambda}I\|} = \frac{\cos\theta}{\epsilon}.$$

Since $\theta < \pi/2$, $\cos\theta \neq 0$ and

$$\|r_1\| \leq \epsilon / \cos\theta.$$

□

The previous observations imply that the deviation of $\|r_1\|$ from its minimal value depends only on the angle between the starting vector and the desired invariant subspace:

COROLLARY 3.3. *Under the assumptions of Theorem 3.2*

$$\epsilon \leq \|r_1\| \leq \epsilon / \cos\theta.$$

Already in 1958 Wilkinson surmised with regard to the number of inverse iterations for real, symmetric tridiagonal matrices 'this will seldom mean producing anything further than x_3' [57, p 93]. In case of Hermitian matrices, Wilkinson proposes several years later that two iterations are enough to produce a vector that is as accurate as can be expected [60, §§III.53-54]. The following result promises (in exact arithmetic) something even better, for the larger class of normal matrices: For over eighty percent of the starting vectors, one iteration suffices to drive the residual down to its minimal value.

COROLLARY 3.4. *If, in addition to the assumptions of Theorem 3.2, θ does not exceed $75°$ then*

$$\|r_1\| \leq 4\epsilon.$$

Proof. Apply the bound in Theorem 3.2 and use the fact that $\theta \leq 75°$ implies

$$\cos\theta \geq \frac{\sqrt{2}}{4}(\sqrt{3} - 1) \geq \frac{1}{4}.$$

□

This result is confirmed in practice. With regard to his particular choice of starting vector $c_1 = e$ Wilkinson remarked in the context of symmetric tridiagonal matrices [61, p 323]:

> In practice we have usually found this choice of [starting vector] to be so effective that after the first iteration, $[x_1]$ is already a *good* approximation to the required eigenvector [...]

3.2. Residual Norms Never Increase. Although he did not give a proof, Wilkinson probably knew deep down that this was true, as the following quotation shows [62, p 376]:

> For a well-conditioned eigenvalue and therefore for all eigenvalues of a normal matrix subsequent iterations are perfectly safe and indeed give valuable protection against an unfortunate choice of initial vector.

We formalise Wilkinson's observation and show that residuals of a normal matrix are monotonically non-increasing (in exact arithmetic). This observation is crucial for establishing that inverse iteration always converges for normal matrices.

THEOREM 3.5. *Let inverse iteration be applied to a non-singular normal matrix* $H - \hat{\lambda}I$, *and let* $r_k = (H - \hat{\lambda}I)x_k$ *be the residual for the* k*th iterate* x_k.

Then

$$\frac{\|r_k\|}{\|r_{k-1}\|} \leq 1, \qquad k \geq 1.$$

Proof. As $\|r_k\| = 1/\|z_k\|$ (§2.3), the ratio of two successive residuals is

$$\frac{\|r_k\|}{\|r_{k-1}\|} = \frac{1}{\|z_k\|}\frac{1}{\|(H-\hat{\lambda}I)x_{k-1}\|} = \frac{1}{\|(H-\hat{\lambda}I)^{-1}x_{k-1}\|\,\|(H-\hat{\lambda}I)x_{k-1}\|}.$$

Now use the fact that $\|Hx\| = \|H^*x\|$ for any vector x when H is normal [15, Theorem 1], [24, Problem 1, p 108], and then apply the Cauchy-Schwartz inequality,

$$\begin{aligned}\|(H-\hat{\lambda}I)^{-1}x_{k-1}\|\,\|(H-\hat{\lambda}I)x_{k-1}\| &= \|(H-\hat{\lambda}I)^{-*}x_{k-1}\|\,\|(H-\hat{\lambda}I)x_{k-1}\| \\ &\geq |x_{k-1}^*(H-\hat{\lambda}I)^{-1}(H-\hat{\lambda}I)x_{k-1}| = 1.\end{aligned}$$

□

Theorem 3.5 does *not* imply that the residual norms go to zero. That's because they are bounded below by ϵ (Theorem 3.1). The residual norms are not even assured to converge to ϵ. If an iterate lies in an invariant subspace of $H - \hat{\lambda}I$ all of whose eigenvalues have magnitude $\gamma > \epsilon$ then the residual norms cannot be smaller than γ.

3.3. Monotonic Convergence of Residual Norms. In the previous section (cf. Theorem 3.5) we established that residual norms are monotonically non-increasing. Now we show that the residual size remains constant whenever inverse iteration has found an iterate that lies in an invariant subspace all of whose eigenvalues have the same magnitude (these eigenvalues, though, are not necessarily the ones closest to $\hat{\lambda}$).

THEOREM 3.6. *Let inverse iteration be applied to a non-singular normal matrix* $H - \hat{\lambda}I$, *and let* $r_k = (H - \hat{\lambda}I)x_k$ *be the residual for the* k*th iterate* x_k.

Then $\|r_{k-1}\| = \|r_k\|$ *if and only if iterate* x_{k-1} *lies in an invariant subspace of* $H - \hat{\lambda}I$ *all of whose eigenvalues have the same magnitude.*

Proof. Suppose x_{k-1} belongs to an invariant subspace of eigenvalues whose magnitude is γ, where $\gamma > 0$. Partition the eigendecomposition as

$$\Lambda - \hat{\lambda}I = \begin{pmatrix} \tilde{\Lambda}_1 & \\ & \tilde{\Lambda}_2 \end{pmatrix}, \qquad Q = \begin{pmatrix} \tilde{Q}_1 & \tilde{Q}_2 \end{pmatrix},$$

where $|(\tilde{\Lambda}_1)_{ii}| = \gamma$. Then $x_{k-1} = \tilde{Q}_1 y$ for some y with $\|y\| = 1$. The residual for x_{k-1} is

$$r_{k-1} = (H - \hat{\lambda}I)x_{k-1} = \tilde{Q}_1\tilde{\Lambda}_1 y$$

and satisfies $\|r_{k-1}\| = \gamma$. The next iterate

$$z_k = (H - \hat{\lambda}I)^{-1}x_{k-1} = \tilde{Q}_1\tilde{\Lambda}_1^{-1}y$$

satisfies $\|z_k\| = 1/\gamma$. Hence

$$\|r_{k-1}\| = \gamma = 1/\|z_k\| = \|r_k\|.$$

Conversely, if $\|r_{k-1}\| = \|r_k\|$ then, as in the proof of Theorem 3.5, $\|r_k\| = 1/\|z_k\|$ implies

$$1 = \frac{\|r_k\|}{\|r_{k-1}\|} = \frac{1}{\|(H-\hat{\lambda}I)^{-1}x_{k-1}\|\,\|(H-\hat{\lambda}I)x_{k-1}\|}$$

and

$$1 = \|(H-\hat{\lambda}I)^{-1}x_{k-1}\|\,\|(H-\hat{\lambda}I)x_{k-1}\| \geq |x_{k-1}^*(H-\hat{\lambda}I)^{-1}(H-\hat{\lambda}I)x_{k-1}| = 1.$$

Since the Cauchy Schwartz inequality holds with equality, the two vectors involved must be parallel, i.e. $(H-\hat{\lambda}I)x_{k-1} = \mu(H-\hat{\lambda}I)^{-*}x_{k-1}$ for some scalar $\mu \neq 0$. Consequently

$$(H-\hat{\lambda}I)^*(H-\hat{\lambda}I)x_{k-1} = \mu x_{k-1},$$

where $\mu > 0$ since $(H-\hat{\lambda}I)^*(H-\hat{\lambda}I)$ is Hermitian positive-definite. Let m be the multiplicity of μ. Partition the eigendecomposition of

$$(H-\hat{\lambda}I)^*(H-\hat{\lambda}I) = Q(\Lambda-\hat{\lambda}I)^*(\Lambda-\hat{\lambda}I)Q^*$$

as

$$(\Lambda-\hat{\lambda}I)^*(\Lambda-\hat{\lambda}I) = \begin{pmatrix} \mu I_m & \\ & M_2 \end{pmatrix}, \qquad Q = (\tilde{Q}_1 \quad \tilde{Q}_2),$$

where $\tilde{Q}_1$ has m columns. Since x_{k-1} is an eigenvector for μ, we must have $x_{k-1} = \tilde{Q}_1 y$ for some y with $\|y\| = 1$. Thus x_{k-1} belongs to the invariant subspace of all eigenvalues λ_i of H with $|\lambda_i - \hat{\lambda}|^2 = \mu$. □

Thus, inverse iteration converges strictly monotonically in the following sense.

COROLLARY 3.7. *Let inverse iteration be applied to a non-singular normal matrix. Then the norms of the residuals decrease strictly monotonically until some iterate belongs to an invariant subspace all of whose eigenvalues have the same magnitude.*

The above result insures that the iterates approach an invariant subspace. A stronger conclusion, convergence of the iterates to an eigenspace, holds when there is a single (possibly multiple) eigenvalue closest to $\hat{\lambda}$.

COROLLARY 3.8. *Let inverse iteration be applied to a non-singular normal matrix $H-\hat{\lambda}I$. If, for the partitioning in §3.1, $\Lambda_1 = \lambda_1 I$ for some λ_1 and $Q_1^* x_0 \neq 0$, then the norms of the residuals decrease strictly monotonically until some iterate belongs to an eigenspace associated with λ_1.*

In §7 we present more details about the space targeted by the iterates.

4. Diagonalisable Matrices. In contrast to a normal matrix, the eigenvectors of a diagonalisable matrix can, in general, not be chosen to be orthonormal; in fact, they can be arbitrarily ill-conditioned. In this section we examine how ill-conditioning of eigenvectors can affect the size of a residual and the increase in the norms of successive residuals.

We use the following notation. Let A be a diagonalisable matrix with eigendecomposition $A = V\Lambda V^{-1}$; and let

$$\epsilon \equiv \min_i |\lambda_i - \hat{\lambda}|$$

be the accuracy of the shift $\hat{\lambda}$. The condition number of the eigenvector matrix V is $\kappa(V) \equiv \|V\|\,\|V^{-1}\|$.

4.1. Diagonalisable Matrices Are Different. We showed in §2.5, §2.6 and §3.1 that, with an appropriate starting vector, the residual after one iteration is almost negligible. Consequently one should expect iterations beyond the first one to bring more improvement. Unfortunately this is not true. Wilkinson[18] and Varah [54, p 786] point out that inverse iteration applied to a non-normal matrix can produce an almost negligible first residual and much larger subsequent residuals. Wilkinson remarks [62, p 374]:

> It might be felt that the preoccupation with achieving what is required in one iteration is a little pedantic, and that it would be more reasonable to concede the occasional necessity of performing a second iteration and even, rarely, a third. In fact, early inverse iterations procedures were usually designed to perform two steps of inverse iteration in all cases on the grounds that it cost little and would give a useful safety factor. The results proved to be disappointing and it was frequently found that the residuals were vastly above noise level.

Wilkinson points out that this residual growth is neither due to round-off errors [63, p 176] nor is it related to the process of inverse iteration as such [63, p 174]. He puts the blame on ill-conditioned eigenvectors when the matrix is diagonalisable[19]. The following example illustrates such residual growth caused by ill-conditioned eigenvectors.

EXAMPLE 1 (S.C. EISENSTAT). *The residual of a diagonalisable matrix can be much smaller than the accuracy of the shift; and the residual growth can be proportional to the eigenvector condition number.*

Consider the 2×2 *matrix*

$$A - \hat{\lambda} I = \begin{pmatrix} \epsilon & \nu \\ & 1 \end{pmatrix},$$

where $0 < \epsilon < 1$ *and* $\nu > 1$. *Then* $A - \hat{\lambda} I = V \Lambda V^{-1}$ *with*

$$\Lambda = \begin{pmatrix} \epsilon & \\ & 1 \end{pmatrix}, \qquad V = \begin{pmatrix} 1 & \frac{\nu}{1-\epsilon} \\ & 1 \end{pmatrix},$$

and

$$\kappa(V) \geq \frac{\nu^2}{(1-\epsilon)^2}.$$

Fix ϵ *and let* ν *grow, so the ill-conditioning of the eigenvectors grows with* ν,

$$\kappa(V) \to \infty \qquad \text{as} \quad \nu \to \infty.$$

The starting vector $x_0 = e_2$ *produces the unnormalised first iterate*

$$z_1 = (A - \hat{\lambda} I)^{-1} x_0 = \frac{1}{\epsilon} \begin{pmatrix} -\nu \\ \epsilon \end{pmatrix}$$

and the residual is bounded by

$$\|r_1\| = 1/\|z_1\| = \frac{\epsilon}{\sqrt{\nu^2 + \epsilon^2}} \leq \frac{\epsilon}{\nu}.$$

The first residual can be much smaller than the shift ϵ *and tends to zero as* ν *becomes large.*

---

[18] [45, §§6-8], [63], [62, §4]

[19] [45, §7], [62, pp 374-6], [63, pp 174-5]

The normalised first iterate

$$x_1 = \frac{1}{\sqrt{\nu^2+\epsilon^2}} \begin{pmatrix} -\nu \\ \epsilon \end{pmatrix}$$

produces the unnormalised second iterate

$$z_2 = (A - \hat{\lambda} I)^{-1} x_1 = \frac{1}{\epsilon\sqrt{\nu^2+\epsilon^2}} \begin{pmatrix} -\nu(1+\epsilon) \\ \epsilon^2 \end{pmatrix}$$

and the residual

$$\|r_2\| = 1/\|z_2\| = \frac{\epsilon\sqrt{\nu^2+\epsilon^2}}{\sqrt{\nu^2(1+\epsilon)^2+\epsilon^4}}.$$

Since

$$\|r_2\| \geq \frac{\epsilon\sqrt{\nu^2+\epsilon^2}}{\sqrt{(\nu^2+\epsilon^2)(1+\epsilon)^2}} = \frac{\epsilon}{1+\epsilon} \geq \frac{1}{2}\epsilon,$$

the second residual is limited below by the accuracy of the shift, regardless of ν.

As a consequence, the residual growth is bounded below by

$$\frac{\|r_2\|}{\|r_1\|} \geq \frac{1}{2}\,\frac{\epsilon\nu}{\epsilon} = \frac{1}{2}\,\nu,$$

independently of the shift. This means[20]

$$\frac{\|r_2\|}{\|r_1\|} = O\left(\sqrt{\kappa(V)}\right) \to \infty \qquad \text{as} \qquad \nu \to \infty,$$

and the residual growth increases with the ill-conditioning of the eigenvectors.

In the following sections we analyse the behaviour of inverse iteration for diagonalisable matrices: Ill-conditioned eigenvectors can push the residual norm far below the accuracy of the shift, and they can cause significant residual growth from one iteration to the next.

4.2. Lower Bound on the Residual. The following lower bound on the residual decreases with ill-conditioning of the eigenvectors. Hence the residual can be much smaller than the accuracy of the shift if the eigenvectors are ill-conditioned.

THEOREM 4.1 (THEOREM 1 IN [5]). *Let* A *be a diagonalisable matrix with eigenvector matrix* V*, and let* $r_k = (A - \hat{\lambda} I)x_k$ *be the residual for some number* $\hat{\lambda}$ *and vector* x_k *with* $\|x_k\| = 1$.

Then

$$\|r_k\| \geq \epsilon/\kappa(V).$$

Proof. If $\hat{\lambda}$ is an eigenvalue of A then $\epsilon = 0$ and the statement is obviously true. If $\hat{\lambda}$ is not an eigenvalue of A then $A - \hat{\lambda} I$ is non-singular. Hence

$$\|(A - \hat{\lambda} I)^{-1}\| \leq \kappa(V)\, \|(\Lambda - \hat{\lambda} I)^{-1}\| = \kappa(V)/\epsilon$$

and

$$\|r_k\| = \|(A - \hat{\lambda} I)x_k\| \geq \frac{1}{\|(A - \hat{\lambda} I)^{-1}\|} \geq \frac{\epsilon}{\kappa(V)}.$$

□

[20] A number α satisfies $\alpha = O(\beta^m)$ if there exists a positive constant γ such that $|\alpha| \leq \gamma\beta^m$ for sufficiently large β [19, §4.1.1].

This bound can also be interpreted as an inclusion domain for eigenvalues of the matrix polynomial $A - \hat{\lambda}I$,

$$\kappa(V)\,\|r_k\| \geq \epsilon = \min_i |\lambda_i - \hat{\lambda}|.$$

4.3. Upper Bound on Residual Increase. We derive an upper bound on the residual growth that grows with the ill-conditioning of the eigenvectors. It also provides a means for estimating the eigenvector ill-conditioning: The square-root of the residual growth is a lower bound for the eigenvector condition number.

THEOREM 4.2. *Let inverse iteration be applied to a non-singular diagonalisable matrix $A - \hat{\lambda}I$ with eigenvector matrix V, and let $r_k = (A - \hat{\lambda}I)x_k$ be the residual for the kth iterate x_k.*

Then

$$\frac{\|r_k\|}{\|r_{k-1}\|} \leq \kappa(V)^2.$$

Proof. This proof is similar to the one for Theorem 3.5. Since $\|r_k\| = 1/\|z_k\| = 1/\|(A - \hat{\lambda}I)^{-1}x_{k-1}\|$ we get

$$\begin{aligned}\frac{\|r_k\|}{\|r_{k-1}\|} &= \frac{1}{\|(A - \hat{\lambda}I)^{-1}x_{k-1}\|\,\|(A - \hat{\lambda}I)x_{k-1}\|} = \frac{1}{\|V(\Lambda - \hat{\lambda}I)^{-1}y\|\,\|V(\Lambda - \hat{\lambda}I)y\|} \\ &\leq \frac{\|V^{-1}\|^2}{\|(\Lambda - \hat{\lambda}I)^{-1}y\|\,\|(\Lambda - \hat{\lambda}I)y\|},\end{aligned}$$

where $y \equiv V^{-1}x_{k-1}$. The normality of Λ implies

$$\|(\Lambda - \hat{\lambda}I)^{-1}y\|\,\|(\Lambda - \hat{\lambda}I)y\| \geq |y^*y| \geq 1/\|V\|^2.$$

Hence

$$\frac{\|r_k\|}{\|r_{k-1}\|} \leq \|V^{-1}\|^2\,\|V\|^2 = \kappa(V)^2.$$

□

4.4. Residuals May Not Be Very Useful. In contrast to normal matrices, the residual norms of diagonalisable matrices do not necessarily converge even if the iterates lie in an invariant subspace associated with eigenvalues of the same magnitude.

When the matrix is not normal, inverse iteration distinguishes eigenvectors from vectors that belong to an invariant subspace but are not eigenvectors. If x_k is an eigenvector then clearly the residuals of all succeeding iterates have norm $\|r_k\|$. But this may not be true when x_k merely belongs to an invariant subspace but is not an eigenvector. The following example illustrates this.

EXAMPLE 2. *The residual norms of a diagonalisable matrix can change even if the iterates lie in an invariant subspace all of whose eigenvalues have the same magnitude.*

Consider the 2×2 matrix

$$A - \hat{\lambda}I = \begin{pmatrix} \epsilon & \nu \\ & -\epsilon \end{pmatrix}, \qquad \epsilon > 0.$$

The invariant subspace of $A - \hat{\lambda}I$ associated with eigenvalues of magnitude ϵ consists of all 2×1 vectors. Since

$$(A - \hat{\lambda}I)^{-1} = \begin{pmatrix} 1/\epsilon & \nu/\epsilon^2 \\ & -1/\epsilon \end{pmatrix} = \frac{1}{\epsilon^2}(A - \hat{\lambda}I)$$

we get

$$z_k = (A - \hat{\lambda}I)^{-1}x_{k-1} = \frac{1}{\epsilon^2}(A - \hat{\lambda}I)x_{k-1} = \frac{1}{\epsilon^2}r_{k-1}$$

and

$$\|r_k\| = \frac{1}{\|z_k\|} = \frac{\epsilon^2}{\|r_{k-1}\|}.$$

Thus, successive residuals of non-normal matrices can differ in norm although all iterates belong to the same invariant subspace,

$$\frac{\|r_k\|}{\|r_{k-1}\|} = \frac{\epsilon^2}{\|r_{k-1}\|^2}.$$

Fortunately, for this particular matrix there happens to be a fix: $(A - \hat{\lambda}I)^2 = \epsilon^2 I$ *is a normal matrix. Hence the results for normal matrices apply to every other iterate of* $A - \hat{\lambda}I$. *Since* $(A - \hat{\lambda}I)^2$ *is also a scalar matrix, all* 2×1 *vectors are eigenvectors of* $(A - \hat{\lambda}I)^2$, *and* $x_{k-1} = x_{k+1}$. *Thus, inverse iteration converges for all even-numbered iterates, and for all odd-numbered iterates.*

This example illustrates that residual norms of diagonalisable matrices do not necessarily reveal when the iterates have arrived in an invariant subspace.

5. Non-Normal Matrices. In this section we use the Jordan and Schur decompositions to analyse the residual for the class of all complex, square matrices: diagonalisable as well as defective. In particular, we show that the residuals of non-normal matrices can be much smaller than the accuracy of the shift when the Jordan basis is highly ill-conditioned or when the matrix has a highly defective eigenvalue. We also derive tight bounds on the residual growth in terms of the departure of the matrix from normality.

5.1. Lower Bounds for the Residual. We derive lower bounds on the residual in terms of the accuracy of the shift, as well as the conditioning of a Jordan basis and the defectiveness of the eigenvalues.

Let $A = VJV^{-1}$ be a Jordan decomposition where the Jordan matrix J is a block-diagonal matrix with Jordan blocks λ_i or

$$\begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$$

on the diagonal; and let

$$\epsilon \equiv \min_i |\lambda_i - \hat{\lambda}|$$

be the accuracy of the shift $\hat{\lambda}$. The condition number of the Jordan basis V is $\kappa(V) \equiv \|V\| \, \|V^{-1}\|$.

THEOREM 5.1. *Let* $A - \hat{\lambda}I$ *be non-singular with Jordan decomposition* $A = VJV^{-1}$, *and let* $r_k = (A - \hat{\lambda}I)x_k$ *be the residual for some vector* x_k *with* $\|x_k\| = 1$.
Then

$$\|r_k\| \geq \frac{1}{\kappa(V)} \, \frac{1}{\|(J - \hat{\lambda}I)^{-1}\|}.$$

Proof. The inequality follows from the lower bound (§2.4),

$$\|r_k\| \geq 1/\|(A - \hat{\lambda}I)^{-1}\|,$$

and the submultiplicative inequality

$$\|(A-\hat{\lambda}I)^{-1}\| \leq \kappa(V)\,\|(J-\hat{\lambda}I)^{-1}\|.$$

□

Below we present several bounds for $\|(J-\hat{\lambda}I)^{-1}\|$. The first upper bound was derived in [11, Proposition 1.12.4]; and a first-order version of the lower bound was derived in [18, §3].

THEOREM 5.2. *Let $J-\hat{\lambda}I$ be a non-singular Jordan matrix J; let m be the order of a largest Jordan block of J; and let l be the order of any Jordan block of $J-\hat{\lambda}I$ whose diagonal element has absolute value ϵ.*

Then

$$\|(J-\hat{\lambda}I)^{-1}\| \leq \frac{(1+\epsilon)^{m-1}}{\epsilon^m},$$

and

$$\frac{1}{\epsilon^l}\sqrt{\frac{1-\epsilon^{2l}}{1-\epsilon^2}} \leq \|(J-\hat{\lambda}I)^{-1}\| \leq \frac{\sqrt{m}}{\epsilon^m}\sqrt{\frac{1-\epsilon^{2m}}{1-\epsilon^2}}.$$

Proof. The first bound appears as Theorem 8.2 in Appendix 1 (§8), while the remaining two are part of Theorem 8.4. □

The lower bound on $\|(J-\hat{\lambda}I)^{-1}\|$ is proportional to $1/\epsilon^l$, which means that $\|(J-\hat{\lambda}I)^{-1}\|$ increases exponentially with the defectiveness of an eigenvalue closest to $\hat{\lambda}$. We conclude that the residual can be much smaller than the accuracy of $\hat{\lambda}$ when the Jordan basis is ill-conditioned, or when A has a highly defective eigenvalue:

COROLLARY 5.3. *Let $A-\hat{\lambda}I$ be non-singular with Jordan decomposition $A = VJV^{-1}$; let m the order of the largest Jordan block of J; and let $r_k = (A-\hat{\lambda}I)x_k$ be teh residual for some vector x_k with $\|x_k\| = 1$.*

Then

$$\|r_k\| \geq c\,\epsilon^m/\kappa(V),$$

where

$$c \equiv \max\left\{\left(\frac{1}{1+\epsilon}\right)^{m-1}, \frac{1}{\sqrt{m}}\sqrt{\frac{1-\epsilon^2}{1-\epsilon^{2m}}}\right\}.$$

Proof. The inequality follows from Theorems 5.1 and 5.2. □

In the case of diagonalisable matrices, $m = 1$ and we recover the bound from Theorem 4.1.

5.2. Upper Bounds on Residual Increase. To bound the residual growth for general matrices we use a Schur decomposition

$$A = Q(\Lambda - N)Q^*,$$

where Q is unitary, Λ is diagonal and N is strictly upper triangular. When $N = 0$, A is normal. When A is non-normal, one can measure the *departure of A from normality*[21]. Henrici proposes as one such measure, among others, the Frobenius norm of N, $\|N\|_F$ [22, §1.2]. Although N is not unique, $\|N\|_F$ is.

Here we are interested in the two-norm, $\|N\|$, which is not unique. But since Frobenius and two norms are related by [17, §2.2]

$$\frac{1}{\sqrt{n}}\|N\|_F \leq \|N\| \leq \|N\|_F,$$

[21] e.g. [15, 22, 35, 36]

we are content to know that $\|N\|$ is at most $\sqrt{n}$ away from a unique bound. We measure the departure of $A - \hat{\lambda}I$ from normality by

$$\eta \equiv \|N\| \, \|(\Lambda - \hat{\lambda}I)^{-1}\| = \|N\|/\epsilon.$$

This is a relative measure because it compares the size of the non-normal part of $A - \hat{\lambda}I$ to the eigenvalues of smallest magnitude. We use it to bound the residual growth.

THEOREM 5.4. *Let inverse iteration be applied to a non-singular matrix $A - \hat{\lambda}I$ whose Schur decomposition has nilpotent part N; let $m-1 \equiv \mathrm{rank}(N)$, where $m-1 < n$; and let $r_k = (A - \hat{\lambda}I)x_k$ be the residual for the kth iterate x_k.*

Then

$$\frac{\|r_k\|}{\|r_{k-1}\|} \leq (1+\eta) \, \frac{1-\eta^m}{1-\eta}.$$

Proof. From

$$\frac{\|r_k\|}{\|r_{k-1}\|} = \frac{1}{\|(A - \hat{\lambda}I)^{-1}x_{k-1}\| \, \|(A - \hat{\lambda}I)x_{k-1}\|}$$

and the Schur decomposition of $A - \hat{\lambda}I$ with $\hat{\Lambda} \equiv \Lambda - \hat{\lambda}I$ and $y \equiv Q^* x_{k-1}$ follows

$$\|(A - \hat{\lambda}I)x_{k-1}\| = \|\hat{\Lambda}(I - \hat{\Lambda}^{-1}N)y\|,$$

and

$$\|(A - \hat{\lambda}I)^{-1}x_{k-1}\| = \|\hat{\Lambda}^{-1}(I - N\hat{\Lambda}^{-1})^{-1}y\| = \|\hat{\Lambda}^{-*}(I - N\hat{\Lambda}^{-1})^{-1}y\|.$$

The last equality holds because $\hat{\Lambda}$ is normal [15, Theorem 1], [24, Problem 1, p 108]. Application of the Cauchy-Schwartz inequality gives

$$\begin{aligned}
\|(A - \hat{\lambda}I)^{-1}x_{k-1}\| \, \|(A - \hat{\lambda}I)x_{k-1}\| &= \|\hat{\Lambda}^{-*}(I - N\hat{\Lambda}^{-1})^{-1}y\| \, \|\hat{\Lambda}(I - \hat{\Lambda}^{-1}N)y\| \\
&\geq |y^*(I - N\hat{\Lambda}^{-1})^{-*}(I - \hat{\Lambda}^{-1}N)y| \\
&\geq \frac{1}{\|(I - \hat{\Lambda}^{-1}N)^{-1}(I - \hat{\Lambda}^{-*}N^*)\|} \\
&\geq \frac{1}{\|(I - \hat{\Lambda}^{-1}N)^{-1}\| \, (1+\eta)}.
\end{aligned}$$

Since N is nilpotent of rank $m-1$,

$$\|(I - \hat{\Lambda}^{-1}N)^{-1}\| = \|\sum_{j=0}^{m-1} (\hat{\Lambda}^{-1}N)^j\| \leq \sum_{j=0}^{m-1} \eta^j = \frac{1-\eta^m}{1-\eta}.$$

Applying the bounds with η to the expression for $\|r_k\|/\|r_{k-1}\|$ gives the desired result. □

When A is normal, $\eta = 0$ and the upper bound equals one, which is consistent with the bound for normal matrices in Theorem 3.5.

When $\eta > 1$, the upper bound grows exponentially with η, since

$$\frac{\|r_k\|}{\|r_{k-1}\|} \leq \frac{\eta+1}{\eta-1} \, \eta^m = O(\eta^m).$$

Thus, the problem of computing an eigenvector with inverse iteration is exponentially ill-conditioned when $A - \hat{\lambda}I$ has a large departure from normality. The closer $\hat{\lambda}$ is to an eigenvalue of A, the smaller is ϵ, and the larger is η. Thus, increasing the accuracy

of a computed eigenvalue $\hat{\lambda}$ increases the departure of $A - \hat{\lambda}I$ from normality, which in turn makes it harder for inverse iteration to compute an eigenvector.

When $\eta < 1$, the upper bound can be simplified to

$$\frac{\|r_k\|}{\|r_{k-1}\|} \leq \frac{1+\eta}{1-\eta}.$$

It equals one when $\eta = 0$ and is strictly increasing for $0 \leq \eta < 1$.

The next example presents a 'weakly non-normal' matrix with bounded residual growth but unbounded eigenvector condition number.

EXAMPLE 3. *The residual growth of a 2×2 matrix with $\eta < 1$ never exceeds four. Yet, when the matrix is diagonalisable, its eigenvector condition number can be arbitrarily large.*

Consider the 2×2 matrix

$$A - \hat{\lambda}I = \begin{pmatrix} \epsilon & \nu \\ & \lambda \end{pmatrix}, \qquad 0 < \epsilon \leq |\lambda|.$$

Since $m = 2$, the upper bound in Theorem 5.4 is

$$\frac{\|r_k\|}{\|r_{k-1}\|} \leq (1+\eta)^2.$$

For $\eta < 1$ this implies

$$\frac{\|r_k\|}{\|r_{k-1}\|} \leq 4.$$

Thus the residual growth remains small – whether the matrix has two distinct eigenvalues, or a double defective eigenvalue.

When the eigenvalues are distinct, as in Example 1, then $\epsilon < |\lambda|$ and A is diagonalisable. The upper bound in Theorem 4.2,

$$\frac{\|r_k\|}{\|r_{k-1}\|} \leq \kappa(V)^2,$$

can be made arbitrarily large by shrinking the eigenvalue separation, since

$$\kappa(V) \geq \frac{|\nu|^2}{|\lambda - \epsilon|^2}.$$

In general, ill-conditioned eigenvectors are not necessarily responsible for a large departure of $A - \hat{\lambda}I$ from normality [7, Example 9.1]. Hence the bound on the residual growth for diagonalisable matrices in Theorem 4.2 may be totally unrealistic. We show in the next section that the residual growth is primarily affected by the departure of $A - \hat{\lambda}I$ from normality, rather than by eigenvector conditioning. Example 3 seems to suggest that any harm done by defectiveness is limited, as long as the departure from normality remains small.

5.3. The Bounds Make Sense. In this section we demonstrate that the bound in Theorem 5.4 is realistic in the following sense. Based on a different measure for departure from normality, we show that for any matrix there always exist iterates whose residual growth is at least as large as the departure of $A - \hat{\lambda}I$ from normality.

An alternative measure for departure from normality, which is invariant under scalar multiplication, is the *Henrici number* [6, Definition 1.1],[7, Definition 9.1]

$$\mathrm{He(A)} \equiv \frac{\|\mathrm{A^*A - AA^*}\|}{\|\mathrm{A^2}\|} \leq 2\frac{\|\mathrm{A}\|^2}{\|\mathrm{A^2}\|}.$$

When A is normal matrix then $\mathrm{He(A)} = 0$. Relations of the Henrici number to other measures of non-normality are described in Appendix 2 (§9).

One of Wilkinson's explanations for residual growth[22] is based on the singular value decomposition [18, §15], [45, §8]: The starting vector x_0 that produces the smallest possible residual r_1 is a left singular vector u_n associated with the smallest singular value of $A - \hat{\lambda}I$ (cf. §2.4). The resulting z_1 lies in the direction of the corresponding right singular vector v_n. Representing v_n in terms of the left singular vector basis may result in a very small coefficient for u_n if an eigenvalue closest to $\hat{\lambda}$ is ill-conditioned. In this case r_2 is likely to be much larger than r_1.

In the same spirit, we consider two possibilities for making $\|r_2\|/\|r_1\|$ large: either make $\|r_1\|$ minimal, as Wilkinson suggested; or else make $\|r_2\|$ maximal. As a consequence the residual growth is no less than the departure from normality of $(A-\hat{\lambda}I)^{-1}$ in the first case, and of $A - \hat{\lambda}I$ in the second case.

THEOREM 5.5. *Let inverse iteration be applied to a non-singular matrix $A - \hat{\lambda}I$, and let $r_k = (A - \hat{\lambda}I)x_k$ be the residual for the kth iterate x_k.*

There exists a starting vector x_0 so that $\|r_1\|$ is minimal and

$$\frac{\|r_2\|}{\|r_1\|} \geq \frac{\|(A-\hat{\lambda}I)^{-1}\|^2}{\|(A-\hat{\lambda}I)^{-2}\|} \geq \frac{1}{2}\mathrm{He}((\mathrm{A}-\hat{\lambda}\mathrm{I})^{-1}).$$

There also exists a starting vector x_0 so that $\|r_2\|$ is maximal and

$$\frac{\|r_2\|}{\|r_1\|} \geq \frac{\|A-\hat{\lambda}I\|^2}{\|(A-\hat{\lambda}I)^2\|} \geq \frac{1}{2}\mathrm{He}(\mathrm{A}-\hat{\lambda}\mathrm{I}).$$

Proof. In the first case, let u_n and v_n be respective left and right singular vectors associated with the smallest singular value of $A - \hat{\lambda}I$, i.e.

$$(A-\hat{\lambda}I)v_n = \frac{u_n}{\|(A-\hat{\lambda}I)^{-1}\|}.$$

If $x_0 = u_n$ then $x_1 = v_n$ and $\|r_1\| = 1/\|(A-\hat{\lambda}I)^{-1}\|$ is minimal (§2.4). Therefore

$$\begin{aligned}\frac{\|r_2\|}{\|r_1\|} &= \frac{1}{\|z_2\|}\frac{1}{\|(A-\hat{\lambda}I)x_1\|} = \frac{1}{\|(A-\hat{\lambda}I)^{-1}x_1\|\,\|(A-\hat{\lambda}I)x_1\|} \\ &= \frac{\|(A-\hat{\lambda}I)^{-1}\|^2}{\|(A-\hat{\lambda}I)^{-2}x_0\|} \geq \frac{\|(A-\hat{\lambda}I)^{-1}\|^2}{\|(A-\hat{\lambda}I)^{-2}\|} \geq \frac{1}{2}\mathrm{He}((\mathrm{A}-\hat{\lambda}\mathrm{I})^{-1}).\end{aligned}$$

In the second case, let u_1 and v_1 be respective left and right singular vectors associated with the largest singular value of $A - \hat{\lambda}I$, i.e.

$$(A-\hat{\lambda}I)v_1 = \|A-\hat{\lambda}I\|\,u_1.$$

If $x_0 = (A-\hat{\lambda}I)u_1/\|(A-\hat{\lambda}I)u_1\|$ then $x_1 = u_1$ and $x_2 = v_1$. Then $\|r_2\| = \|A-\hat{\lambda}I\|$ is maximal, and

$$\begin{aligned}\frac{\|r_2\|}{\|r_1\|} &= \frac{1}{\|(A-\hat{\lambda}I)^{-1}x_1\|\,\|(A-\hat{\lambda}I)x_1\|} = \frac{\|A-\hat{\lambda}I\|^2}{\|(A-\hat{\lambda}I)^2x_2\|} \\ &\geq \frac{\|A-\hat{\lambda}I\|^2}{\|(A-\hat{\lambda}I)^2\|} \geq \frac{1}{2}\mathrm{He}(\mathrm{A}-\hat{\lambda}\mathrm{I}).\end{aligned}$$

[22] For simplicity we consider residual growth from the first to the second iteration, but the argument of course applies to any pair of successive iterations.

□

Thus the best possible starting vector, i.e. a right singular vector associated with the smallest singular value of $A-\hat{\lambda}I$, can lead to significant residual growth. Wilkinson was very well aware of this and he recommended [44, p 420]:

> For non-normal matrices there is much to be said for choosing the initial vector in such a way that the full growth occurs in one iteration, thus ensuring a small residual. This is the only simple way we have of recognising a satisfactory performance. For well conditioned eigenvalues (and therefore for all eigenvalues of normal matrices) there is no loss in performing more than one iteration and subsequent iterations offset an unfortunate choice of the initial vector.

5.4. A Particular Example of Extreme Residual Increase. In this section we present a matrix for which the bound in Theorem 5.4 is tight and the residual growth is exponential in the departure from normality.

This example is designed to pin-point the apparent regress of inverse iteration after the first iteration, which Wilkinson documented extensively[23]: The first residual is tiny, because it is totally under the control of non-normality. But subsequent residuals are much larger: The influence of non-normality has disappeared and they behave more like residuals of normal matrices. In the example below the residuals satisfy

$$\|r_1\| \leq \frac{1}{\eta^{n-1}}, \qquad \|r_2\| \geq \frac{1}{n}, \qquad \|r_3\| \geq \frac{2}{n+1},$$

where $\eta > 1$. For instance, when $n = 1000$ and $\eta = 2$ then the first residual is completely negligible while subsequent residuals are significantly larger than single precision,

$$\|r_1\| \leq 2^{-999} \leq 2 \cdot 10^{-301}, \qquad \|r_2\| \geq 10^{-3}, \qquad \|r_3\| \geq 10^{-3}.$$

Let the matrix A have a Schur decomposition $A = Q(\Lambda - N)Q^*$ with

$$Q = \Lambda - \hat{\lambda}I = I, \qquad N = \eta Z, \qquad \eta > 1,$$

and

$$Z \equiv \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}.$$

Thus $\epsilon = 1$, $\|N\|/\epsilon = \eta > 1$, $m = n-1$, and

$$(A - \hat{\lambda}I)^{-1} = \begin{pmatrix} 1 & \eta & \eta^2 & \dots & \eta^{n-1} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \eta^2 \\ & & & \ddots & \eta \\ & & & & 1 \end{pmatrix}.$$

Let the starting vector be $x_0 = e_n$. Then $\|r_1\|$ is almost minimal because x_1 is a multiple of a column of $(A - \hat{\lambda}I)^{-1}$ with largest norm (§2.5).

[23] [18, p 615], [45, P 355], [62, p 375], [63]

The first iteration produces the unnormalised iterate $z_1 = (A - \hat{\lambda} I)^{-1} e_n$ with

$$\|z_1\|^2 = \sum_{i=0}^{n-1} (\eta^i)^2 = \frac{\eta^{2n} - 1}{\eta^2 - 1},$$

and the normalised iterate

$$x_1 = \frac{z_1}{\|z_1\|} = \left(\frac{\eta^{2n} - 1}{\eta^2 - 1} \right)^{-1/2} \begin{pmatrix} \eta^{n-1} \\ \vdots \\ \eta^2 \\ \eta \\ 1 \end{pmatrix}.$$

The residual norm can be bounded in terms of the $(1, n)$ element of $(A - \hat{\lambda} I)^{-1}$,

$$\|r_1\| = \frac{1}{\|z_1\|} = \frac{1}{\|(A - \hat{\lambda} I)^{-1} e_n\|} \leq \frac{1}{\|((A - \hat{\lambda} I)^{-1})_{1n}\|} = \frac{1}{\eta^{n-1}}.$$

The second iteration produces the unnormalised iterate

$$z_2 = (A - \hat{\lambda} I)^{-1} x_1 = \left(\frac{\eta^{2n} - 1}{\eta^2 - 1} \right)^{-1/2} \begin{pmatrix} n\eta^{n-1} \\ \vdots \\ 3\eta^2 \\ 2\eta \\ 1 \end{pmatrix}$$

with

$$\|z_2\|^2 = \left(\frac{\eta^{2n} - 1}{\eta^2 - 1} \right)^{-1} \sum_{i=0}^{n-1} \left((i+1)\eta^i \right)^2 \leq n^2,$$

and the normalised iterate

$$x_2 = \frac{z_2}{\|z_2\|} = \left(\sum_{i=0}^{n-1} \left((i+1)\eta^i \right)^2 \right)^{-1/2} \begin{pmatrix} n\eta^{n-1} \\ \vdots \\ 3\eta^2 \\ 2\eta \\ 1 \end{pmatrix}.$$

The residual norm is

$$\|r_2\| = \frac{1}{\|z_2\|} \geq \frac{1}{n}.$$

The third iteration produces the unnormalised iterate

$$z_3 = (A - \hat{\lambda} I)^{-1} x_2 = \left(\sum_{i=0}^{n-1} \left((i+1)\eta^i \right)^2 \right)^{-1/2} \begin{pmatrix} \frac{n(n+1)}{2}\eta^{n-1} \\ \vdots \\ 6\eta^2 \\ 3\eta \\ 1 \end{pmatrix}$$

with

$$\|z_3\|^2 = \left(\sum_{i=0}^{n-1} \left((i+1)\eta^i \right)^2 \right)^{-1} \sum_{i=0}^{n-1} \left(\frac{(i+1)(i+2)}{2} \eta^i \right)^2 \leq \left(\frac{n+1}{2} \right)^2.$$

The residual norm is

$$\|r_3\| = \frac{1}{\|z_3\|} \geq \frac{2}{n+1}.$$

In spite of the setback in residual size after the first iteration, the gradation of the elements in z_1, z_2, and z_3 indicates that the iterates eventually converge to the eigenvector e_1.

6. Finite Precision Arithmetic. In this section we illustrate that finite precision arithmetic has little effect on inverse iteration, in particular on residual norms, starting vectors and solutions to linear systems. Wilkinson agreed: 'the inclusion of rounding errors in the inverse iteration process makes surprisingly little difference' [45, p 355].

We use the following notation. Suppose $\hat{x}_{k-1}$ with $\|\hat{x}_{k-1}\| = 1$ is the current iterate computed in finite precision, and $\hat{z}_k$ is the new iterate computed in finite precision. That is, the process of linear system solution in §2.2 produces a matrix F_k depending on $A - \hat{\lambda}I$ and $\hat{x}_{k-1}$ so that $\hat{z}_k$ is the exact solution to the linear system [60, §III.25], [61, §9.48]

$$(A - \hat{\lambda}I - F_k)\hat{z}_k = \hat{x}_{k-1}.$$

We assume that the normalisation is error free, so the normalised computed iterate $\hat{x}_k \equiv \hat{z}_k/\|\hat{z}_k\|$ satisfies

$$(A - \hat{\lambda}I - F_k)\hat{x}_k = \hat{s}_k\, \hat{x}_{k-1},$$

where the normalisation constant is $\hat{s}_k \equiv 1/\|\hat{z}_k\|$.

In the case of floating point arithmetic, the backward error F_k is bounded by [60, §III.25]

$$\|F_k\| \leq p(n)\, \rho\, \|A - \hat{\lambda}I\|\, \epsilon_M,$$

where $p(n)$ is a low degree polynomial in the matrix size n. The machine epsilon ϵ_M determines the accuracy of the computer arithmetic; it is the *smallest* positive floating point number that when added to 1.0 results in a floating point number larger than 1.0 [32, §2.3].

The growth factor ρ is the ratio between the element of largest magnitude occurring during Gaussian elimination and the element of largest magnitude in the original matrix $A - \hat{\lambda}I$ [58, §2]. For Gaussian elimination with partial pivoting $\rho \leq 2^{n-1}$ [59, §§8, 29]. Although one can find practical examples with exponential growth factors [16], numerical experiments with random matrices suggest growth factors proportional to $n^{2/3}$ [53]. According to public opinion, ρ is small [17, §3.4.6]. Regarding the possibility of large elements in U, hence a large ρ, Wilkinson remarked [61, §9.50]:

> I take the view that this danger is negligible.

while Kahan believes [30, p 782],

> Intolerable pivot-growth is a phenomenon that happens only to numerical analysts who are looking for that phenomenon.

It seems therefore reasonable to assume that ρ is not too large, and that the backward error from Gaussian elimination with partial pivoting is small.

For real symmetric tridiagonal matrices T, the sharper bound

$$\|F_k\| \leq c\, \sqrt{n}\, \|T - \hat{\lambda}I\|\, \epsilon_M$$

holds, where c is a small constant [60, §3.52], [61, §5.55].

6.1. The Finite Precision Residual. The exact iterate x_k is an eigenvector of a matrix close to A if its unnormalised version z_k has sufficiently large norm (cf. §2.3). Wilkinson demonstrated that this is also true in finite precision arithmetic [60, §III.52], [61, §5.55]. The meaning of 'sufficiently large' is determined by the size of the backward error F_k from the linear system solution.

The residual of the computed iterate

$$\hat{r}_k \equiv (A - \hat{\lambda}I)\hat{x}_k = -F_k\hat{x}_k + \hat{s}_k\hat{x}_{k-1}$$

is bounded by

$$\frac{1}{\|\hat{z}_k\|} - \|F_k\| \leq \|\hat{r}_k\| \leq \frac{1}{\|\hat{z}_k\|} + \|F_k\|.$$

This means, if $\hat{z}_k$ is sufficiently large then the size of the residual is about as small as the backward error. For instance, if the iterate norm is inversely proportional to the backward error, i.e.

$$\|\hat{z}_k\| \geq \frac{1}{c_1\|F_k\|}$$

for some constant $c_1 > 0$, then the residual is at most a multiple of the backward error,

$$\|\hat{r}_k\| \leq (1 + c_1)\,\|F_k\|.$$

A lower bound on $\|\hat{z}_k\|$ can therefore serve as a criterion for terminating the inverse iteration process.

For a real symmetric tridiagonal matrix T Wilkinson suggested [61, p 324] terminating the iteration process one iteration after

$$\|\hat{z}_k\|_\infty \geq \frac{1}{100n\,\epsilon_M}$$

is satisfied, assuming T is normalised so $\|T - \hat{\lambda}I\|_\infty \approx 1$. Wilkinson did his computations on the ACE, a machine with a 46-bit mantissa where $\epsilon_M = 2^{-46}$. The factor $100n$ covers the term $c\sqrt{n}$ in the bound on the backward error for tridiagonal matrices in §6. According to Wilkinson [61, p 325]:

> In practice this has never involved doing more than three iterations and usually only two iterations are necessary. [...] The factor $1/100n$ has no deep significance and merely ensures that we seldom perform an unnecessary extra iteration.

Although Wilkinson's suggestion of performing one additional iteration beyond the stopping criterion does not work well for general matrices (due to possible residual increase, cf. §5 and [54, p 768]), it is effective for symmetric tridiagonal matrices [29, §4.2].

In the more difficult case when one wants to compute an entire eigenbasis of a real, symmetric matrix, the stopping criterion requires that the relative residual associated with a projected matrix be sufficiently small [8, §5.3].

6.2. Good Starting Vectors in Finite Precision. In exact arithmetic there are always starting vectors that lead to an almost minimal residual in a single iteration (cf. §2.5). Wilkinson proved that this is also true for finite precision arithmetic [62, p 373]. That is, the transition to finite precision arithmetic does not affect the size of the first residual significantly.

Suppose the lth column has the largest norm among all columns of $(A - \hat{\lambda}I)^{-1}$, and the computed first iterate $\hat{z}_1$ satisfies

$$(A - \hat{\lambda}I + F_1)\hat{z}_1 = e_l.$$

This implies

$$(I + (A - \hat{\lambda}I)^{-1}F_1)\hat{z}_1 = (A - \hat{\lambda}I)^{-1}e_l$$

and

$$\frac{1}{\sqrt{n}}\,\|(A - \hat{\lambda}I)^{-1}\| \le \|(A - \hat{\lambda}I)^{-1}e_l\| \le \|I + (A - \hat{\lambda}I)^{-1}F_1\|\,\|\hat{z}_1\|.$$

Therefore

$$\|\hat{z}_1\| \ge \frac{1}{\sqrt{n}}\,\frac{\|(A - \hat{\lambda}I)^{-1}\|}{1 + \|(A - \hat{\lambda}I)^{-1}\|\,\|F_1\|}.$$

If $\|(A - \hat{\lambda}I)^{-1}\|\,\|F_1\|$ is small then z_1 and $\hat{z}_1$ have about the same size.

Our bound for $\|\hat{z}_1\|$ appears to be more pessimistic than Wilkinson's [62, p 373] which says essentially that there is a canonical vector e_l such that

$$(A - \hat{\lambda}I + F_1)\hat{z}_1 = e_l$$

and

$$\|\hat{z}_1\| = \|(A - \hat{\lambda}I + F_1)^{-1}e_l\| \ge \frac{1}{\sqrt{n}}\,\|(A - \hat{\lambda}I + F_1)^{-1}\|.$$

But

$$\begin{aligned}\|(A - \hat{\lambda}I + F_1)^{-1}\| &= \|\left(I + (A - \hat{\lambda}I)^{-1}F_1\right)^{-1}(A - \hat{\lambda}I)^{-1}\| \\ &\ge \frac{\|(A - \hat{\lambda}I)^{-1}\|}{\|I + (A - \hat{\lambda}I)^{-1}F_1\|} \ge \frac{\|(A - \hat{\lambda}I)^{-1}\|}{1 + \|(A - \hat{\lambda}I)^{-1}\|\,\|F_1\|}\end{aligned}$$

implies that Wilkinson's result is an upper bound of our result, and it is not much more optimistic than ours.

Unless x_0 contains an extraordinarily small contribution of the desired eigenvector x, Wilkinson argued that the second iterate x_2 is as good as can be expected in finite precision arithmetic [60, §III.53]. Jessup and Ipsen [29] performed a statistical analysis to confirm the effectiveness of random starting vectors for real, symmetric tridiagonal matrices.

6.3. Solution of Ill-Conditioned Linear Systems. A major concern in the early days of inverse iteration was the ill-conditioning of the linear system involving $A - \hat{\lambda}I$ when $\hat{\lambda}$ is a good approximation to an eigenvalue of A. It was believed that the computed solution to $(A - \hat{\lambda}I)z = \hat{x}_{k-1}$ would be totally inaccurate [45, p 340]. Wilkinson went to great lengths to allay these concerns[24]. He reasoned that only the direction of a solution is of interest but not the exact multiple: A computed iterate with a large norm lies in 'the correct direction' and 'is wrong only by [a] scalar factor' [45, p 342].

We quantify Wilkinson's argument and compare the computed first iterate to the exact first iterate (as before, of course, the argument applies to any iterate). The respective exact and finite precision computations are

$$(A - \hat{\lambda}I)z_1 = x_0, \qquad (A - \hat{\lambda}I + F_1)\hat{z}_1 = x_0.$$

Below we make the standard assumption[25] that $\|(A - \hat{\lambda}I)^{-1}F_1\| < 1$, which means that $A - \hat{\lambda}I$ is sufficiently well-conditioned with respect to the backward error, so

[24] [44, §6], [45, §2], [60, §III.53], [61, §5.57], [61, §§9.48, 49], [62, §5]

[25] [17, Lemma 2.7.1], [51, Theorem 4.2], [58, §9.(S)], [60, §III.12]

nonsingularity is preserved despite the perturbation. The following result assures that the computed iterate is not much smaller than the exact iterate.

THEOREM 6.1. *Let $A - \hat{\lambda}I$ be non-singular; let $\|(A - \hat{\lambda}I)^{-1}F_1\| < 1$; and let*

$$(A - \hat{\lambda}I)z_1 = x_0, \qquad (A - \hat{\lambda}I + F_1)\hat{z}_1 = x_0.$$

Then

$$\|\hat{z}_1\| \geq \frac{1}{2}\|z_1\|.$$

Proof. Since $\hat{\lambda}$ is not an eigenvalue of A, $A - \hat{\lambda}I$ is non-singular and

$$\left(I + (A - \hat{\lambda}I)^{-1}F_1\right)\hat{z}_1 = z_1.$$

The assumption $\|(A - \hat{\lambda}I)^{-1}F_1\| < 1$ implies

$$\|z_1\| \leq \left(1 + \|(A - \hat{\lambda}I)^{-1}F_1\|\right)\ \|\hat{z}_1\| \leq 2\|\hat{z}_1\|.$$

□

Since computed and exact iterate are of comparable size, the ill-conditioning of the linear system does not damage the accuracy of an iterate. When $\hat{z}_1$ is a column of maximal norm of $(A - \hat{\lambda}I + F_1)^{-1}$, Theorem 6.1 implies the bound from §6.2.

6.4. An Example of Numerical Software. In this section we briefly describe a state-of-the-art implementation of inverse iteration from the numerical software library LAPACK [1].

Computing an eigenvector of a real symmetric or complex Hermitian matrix H with LAPACK requires three steps [1, 2.3.3]:

1. Reduce H to a real, symmetric tridiagonal matrix T by means of an orthogonal or unitary similarity transformation Q, $H = QTQ^*$.
2. Compute an eigenvector x of T by xSTEIN.
3. Backtransform x to an eigenvector Qx of H.

The reduction to tridiagonal form is the most expensive among the three steps. For a matrix H of order n, the first step requires $O(n^3)$ operations, the second $O(n)$ and the third $O(n^2)$ operations.

The particular choice of Q in the reduction to tridiagonal form depends on the sparsity structure of H. If H is full and dense or if it is sparse and stored in packed format, then Q should be chosen as a product of Householder reflections. If H is banded with bandwidth w then Q should be chosen as a product of Givens rotations, so the reduction requires only $O(w^2\, n)$ operations. Unless requested, Q is not determined explicitly but stored implicitly in factored form, as a sequence of Householder reflections or Givens rotations.

Given a computed eigenvalue $\hat{\lambda}$, the LAPACK subroutine xSTEIN[26] computes an eigenvector of a real symmetric tridiagonal matrix T as follows. We assume at first that all off-diagonal elements of T are non-zero.

Step 1: Compute the LU factors of $T - \hat{\lambda}I$ by Gaussian elimination with partial pivoting:

$$P(T - \hat{\lambda}I) = LU.$$

[26] The prefix 'x' stands for the data type: real single (S) or double (D) precision, or complex single (C) or double precision (Z).

Step 2: Select a random (unnormalised) starting vector z_0 with elements from a uniform $(-1, 1)$ distribution.

Step 3: Execute at most five of the following iterations:

Step i.1: Normalise the current iterate so its one-norm is on the order of machine epsilon: $x_{k-1} = s_{k-1} * z_{k-1}$ where

$$s_{k-1} = n \, \|T\|_1 \, \max\{\epsilon_M, \, |u_{nn}|\} \, / \, \|z_{k-1}\|_1,$$

and u_{nn} is the trailing diagonal element of U.

Step i.2: Solve the two triangular systems to compute the new unnormalised iterate:

$$Lc_k = Px_{k-1}, \qquad Uz_k = c_k.$$

Step i.3: Check whether the infinity-norm of the new iterate has grown sufficiently:

$$\text{Is} \quad \|z_k\|_\infty \geq 1/\sqrt{10n}\,?$$

1. If yes, then perform two additional iterations. Normalise the final iterate x_{k+2} so that $\|x_{k+2}\|_2 = 1$ and the largest element in magnitude is positive. Stop.
2. If no, and if the current number of iterations is less than five, start again at Step i.1.
3. Otherwise terminate unsuccessfully.

We comment on the different steps of xSTEIN:

Step 1. The matrix T is input to xSTEIN in the form of two arrays of length n, one array containing the diagonal elements of T and the other containing the off-diagonal elements. Gaussian elimination with pivoting on $T - \hat{\lambda} I$ results in a unit lower triangular matrix L with at most one non-zero subdiagonal and an upper triangular matrix U with at most two non-zero superdiagonals. xSTEIN uses an array of length $5n$ to store the starting vector z_0, the subdiagonal of L, and the diagonal and superdiagonals of U in a single array. In subsequent iterations, the location that initially held z_0 is overwritten with the current iterate.

Step i.1. Instead of normalising the current iterate x_{k-1} so it has unit-norm, xSTEIN normalises x_{k-1} to make its norm as small as possible. The purpose is to avoid overflow in the next iterate z_k.

Step i.2. In contrast to Wilkinson's practice of saving one triangular system solution in the first iteration (cf. §2.2), xSTEIN executes the first iteration like all others.

To avoid overflow in the elements of z_k, xStein gradually increases the magnitude of very small diagonal elements of U: If entry i of z_k would be larger than the reciprocal of the smallest normalised machine number, then u_{ii} (by which we divide to obtain this entry) has its magnitude increased by $2^p \, \epsilon_M \, \max_{i,j} |u_{i,j}|$, where p is the number of already perturbed diagonal elements.

Step i.3. The stopping criterion determines whether the norm of the new iterate z_k has grown in comparison to the norm of the current iterate x_{k-1} (in previous chapters the norm of x_{k-1} equals one). To see this, divide the stopping criterion by $\|x_{k-1}\|_1$ and use the fact that $\|x_{k-1}\|_1 = s_{k-1}$. This amounts to asking whether

$$\frac{\|z_k\|_\infty}{\|x_{k-1}\|_1} \geq \frac{1}{\sqrt{10\,n}\, n \, \|T\|_1 \, \epsilon_M}.$$

Thus the stopping criterion in xSTEIN is similar in spirit to Wilkinson's stopping criterion in §6.1 (Wilkinson's criterion does not contain the norm of T because he assumed that $T - \hat{\lambda}I$ is normalised so its norm is close to one).

When $|u_{nn}|$ is on the order of ϵ_M then $T - \hat{\lambda}I$ is numerically singular. The convergence criterion expects a lot more growth from an iterate when the matrix is close to singular than when it is far from singular.

In the preceding discussion we assumed that T has non-zero off-diagonal elements. When T does have zero off-diagonal elements, it splits into several disjoint submatrices T_i whose eigenvalues are the eigenvalues of T. xSTEIN requires as input the index i of the submatrix T_i to which $\hat{\lambda}$ belongs, and the boundaries of each submatrix. Then xSTEIN computes an eigenvector x of T_i and expands it to an eigenvector of T by filling zeros into the remaining entries above and below x.

7. Asymptotic Convergence In Exact Arithmetic. In this section we give a convergence proof for inverse iteration applied to a general, complex matrix. In contrast to a normal matrix (cf. §3.3), the residual norms of a non-normal matrix do not decrease strictly monotonically (cf. §5.3 and §5.4). In fact, the residual norms may even fail to converge (cf. Example 2). Thus we need to establish convergence of the iterates proper.

The absence of monotonic convergence is due to the transient dominance of the departure from normality over the eigenvalues. The situation is similar to the 'hump' phenomenon: If all eigenvalues of a matrix B are less than one in magnitude, then the powers of B converge to zero asymptotically, $\|B^k\| \to 0$ as $k \to \infty$ [24, §3.2.5, §5.6.12]. But before asymptotic convergence sets in, $\|B^k\|$ can become much larger than one temporarily before dying down. This transient growth is quantified by the hump $\max_k \|B^k\|/\|B\|$ [23, §2].

Wilkinson established convergence conditions for diagonalisable matrices[27] and for symmetric matrices [60, §III.50]. We extend Wilkinson's argument to general matrices. A simple proof demonstrates that unless the choice of starting vector is particularly unfortunate, the iterates approach the invariant subspace of all eigenvalues closest to $\hat{\lambda}$. Compared to the convergence analysis of multi-dimensional subspaces [43] our task is easier because it is less general: Each iterate represents a 'perturbed subspace' of dimension one; and the 'exact subspace' is defined conveniently, by grouping the eigenvalues as follows.

Let $A - \hat{\lambda}I = V(J - \hat{\lambda}I)V^{-1}$ be a Jordan decomposition and partition

$$J = \begin{pmatrix} J_1 & \\ & J_2 \end{pmatrix}, \qquad V = \begin{pmatrix} V_1 & V_2 \end{pmatrix}, \qquad V^{-1} = \begin{pmatrix} W_1^* \\ W_2^* \end{pmatrix},$$

where J_1 contains all eigenvalues λ_i of A closest to $\hat{\lambda}$, i.e. $|\lambda_i - \hat{\lambda}| = \epsilon$; while J_2 contains the remaining eigenvalues. Again we assume that $\hat{\lambda}$ is not an eigenvalue of A, so $\epsilon > 0$.

The following theorem ensures that inverse iteration gradually removes from the iterates their contribution in the undesirable invariant subspace, i.e the subspace associated with eigenvalues farther away from $\hat{\lambda}$. If the starting vector x_0 has a non-zero contribution in the complement of the undesirable subspace then the iterates approach the desirable subspace, i.e. the subspace associated with the eigenvalues closest to $\hat{\lambda}$. This result is similar to the well-known convergence results for the power method, e.g. [34, §10.3].

[27] [44, §1], [62, §1], [63, p 173]

THEOREM 7.1. *Let inverse iteration be applied to a non-singular matrix* $A - \hat{\lambda}I$. *Then iterate* x_k *is a multiple of* $y_{k1} + y_{k2}$, $k \geq 1$, *where* $y_{k1} \in \text{range}(V_1)$, $y_{k2} \in \text{range}(V_2)$, *and* $y_{k2} \to 0$ *as* $k \to \infty$.

If $W_1^* x_0 \neq 0$ *then* $y_{k1} \neq 0$ *for all* k.

Proof. Write

$$(A - \hat{\lambda}I)^{-1} = \epsilon^{-1} V (\epsilon (J - \hat{\lambda}I)^{-1}) V^{-1},$$

where

$$\epsilon (J - \hat{\lambda}I)^{-1} = \begin{pmatrix} \epsilon (J_1 - \hat{\lambda}I)^{-1} & \\ & \epsilon (J_2 - \hat{\lambda}I)^{-1} \end{pmatrix}.$$

The eigenvalues of $\epsilon (J_1 - \hat{\lambda}I)^{-1}$ have absolute value one; while the ones of $\epsilon (J_2 - \hat{\lambda}I)^{-1}$ have absolute value less than one. Then

$$(A - \hat{\lambda}I)^{-k} x_0 = \epsilon^{-k} \, (y_{k1} + y_{k2}),$$

where

$$y_{k1} \equiv V_1 (\epsilon (J_1 - \hat{\lambda}I)^{-1})^k W_1^* x_0, \qquad y_{k2} \equiv V_2 (\epsilon (J_2 - \hat{\lambda}I)^{-1})^k W_2^* x_0.$$

Since all eigenvalues of $\epsilon (J_2 - \hat{\lambda}I)^{-1}$ are less than one in magnitude, $(\epsilon (J_2 - \hat{\lambda}I)^{-1})^k \to 0$ as $k \to \infty$ [24, §3.2.5, §5.6.12]. Hence $y_{k2} \to 0$. Because

$$x_k = \frac{(A - \hat{\lambda}I)^{-k} x_0}{\| (A - \hat{\lambda}I)^{-k} x_0 \|},$$

x_k is a multiple of $\epsilon^k \, (A - \hat{\lambda}I)^{-k} x_0 = y_{k1} + y_{k2}$.

If $W_1^* x_0 \neq 0$, the non-singularity of $\epsilon (J_1 - \hat{\lambda}I)^{-1}$ implies $y_{k1} \neq 0$. □

Because the contributions y_{k1} of the iterates in the desired subspace depend on k, the above result only guarantees that the iterates approach the desired invariant subspace. It does not imply that they converge to an eigenvector.

Convergence to an eigenvector occurs, for instance, when there is a single, non-defective eigenvalue closest to $\hat{\lambda}$. When this is a multiple eigenvalue, the eigenvector targeted by the iterates depends on the starting vector. The iterates approach unit multiples of this target vector. Below we denote by $|x|$ the vector whose elements are the absolute values of the elements of x.

COROLLARY 7.2. *Let inverse iteration be applied to the non-singular matrix* $A - \hat{\lambda}I$. *If* $J_1 = \lambda_1 I$ *in the Jordan decomposition of* $A - \hat{\lambda}I$, *then iterate* x_k *is a multiple of* $\omega^k y_1 + y_{k2}$, *where* $|\omega| = 1$, $y_1 \equiv V_1 W_1^* x_0$ *is independent of* k, $y_{k2} \in \text{range}(V_2)$, *and* $y_{k2} \to 0$ *as* $k \to \infty$.

If $W_1^* x_0 \neq 0$ *then* $|x_k|$ *approaches a multiple of* $|y_1|$ *as* $k \to \infty$.

Proof. In the proof of Theorem 7.1 set $\omega \equiv \epsilon / (\lambda_1 - \hat{\lambda})$. Then $\epsilon (J_1 - \hat{\lambda}I)^{-1} = \omega I$ and $y_{k1} = \omega^k y_1$. □

8. Appendix 1: Facts about Jordan Blocks. In this section we give bounds on the norm of the inverse of a Jordan block. First we give an upper bound for a single Jordan block.

LEMMA 8.1 (PROPOSITION 1.12.4 IN [11]). *Let*

$$J = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

be of order m and $\lambda \neq 0$. Then

$$\|J^{-1}\| \leq \frac{(1+|\lambda|)^{m-1}}{|\lambda|^m}.$$

Now we bound the norm of a matrix consisting of several Jordan blocks.

THEOREM 8.2 (PROPOSITION 1.12.4 IN [11], THEOREM 8 IN [31]). *Let J be a Jordan matrix whose Jordan blocks have diagonal elements λ_i; let*

$$\epsilon \equiv \min_i |\lambda_i|;$$

and let m be the order of a Jordan block J_j for which $\|J^{-1}\| = \|J_j^{-1}\|$.

If $\epsilon > 0$ then

$$\|J^{-1}\| \leq \frac{(1+\epsilon)^{m-1}}{\epsilon^m}.$$

Proof. Let λ_j be the diagonal element of J_j. Lemma 8.1 implies

$$\|J^{-1}\| = \|J_j^{-1}\| \leq \frac{(1+|\lambda_j|)^{m-1}}{|\lambda_j|^m}.$$

Because of $\epsilon \leq |\lambda_j|$ we get (proof of Proposition 1.12.4 in [11])

$$\frac{(1+|\lambda_j|)^{m-1}}{|\lambda_j|^m} = \frac{1}{|\lambda_j|}\left(\frac{1+|\lambda_j|}{|\lambda_j|}\right)^{m-1} \leq \frac{1}{\epsilon}\left(\frac{1+\epsilon}{\epsilon}\right)^{m-1}.$$

□

Now we derive a lower bound and a second upper bound, which differ by a factor of at most $\sqrt{m}$. A first-order version of the lower bound appeared in [18, §3], and the one-norm upper bound was proved in [37, Lemma 2]. As before, we start with a single Jordan block.

LEMMA 8.3. *Let*

$$J = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

be of order m and $\lambda \neq 0$. Then

$$\|J^{-1}e_m\| \leq \|J^{-1}\| \leq \|J^{-1}e_m\|_1 \leq \sqrt{m}\,\|J^{-1}e_m\|,$$

where for $|\lambda| \neq 1$

$$\|J^{-1}e_m\| = \frac{1}{|\lambda|^m}\sqrt{\frac{1-|\lambda|^{2m}}{1-|\lambda|^2}}, \qquad \|J^{-1}e_m\|_1 = \frac{1}{|\lambda|^m}\sqrt{\frac{1-|\lambda|^m}{1-|\lambda|}}.$$

Proof. To prove the lower bound and the two-norm upper bound, note that the inverse of J is [25, §6.2.13]

$$J^{-1} = \begin{pmatrix} \frac{1}{\lambda} & \frac{-1}{\lambda^2} & \cdots & \frac{(-1)^{m+1}}{\lambda^m} \\ & \frac{1}{\lambda} & & \frac{(-1)^m}{\lambda^{m-1}} \\ & & \ddots & \vdots \\ & & & \frac{1}{\lambda} \end{pmatrix}.$$

The last column of J^{-1} has largest norm among all columns,

$$\|J^{-1}e_m\| \le \|J^{-1}\| \le \sqrt{m}\,\|J^{-1}e_m\|,$$

and the square of its two-norm is a geometric progression in $1/|\lambda|^2$,

$$\|J^{-1}e_m\|^2 = \frac{1}{|\lambda|^2}\sum_{i=0}^{m-1}\frac{1}{|\lambda|^{2i}} = \frac{1}{|\lambda|^2}\,\frac{1-\frac{1}{|\lambda|^{2m}}}{1-\frac{1}{|\lambda|^2}}.$$

Taking square-roots on both sides gives

$$\|J^{-1}e_m\| = \frac{1}{|\lambda|^m}\sqrt{\frac{1-|\lambda|^{2m}}{1-|\lambda|^2}}.$$

To prove the one-norm upper bound, apply the Neumann lemma [17, Lemma 2.3.3] to $J = \lambda I + N$,

$$\|J^{-1}\| = \frac{1}{|\lambda|}\left\|\left(I+\frac{1}{|\lambda|}N\right)^{-1}\right\| \le \frac{1}{|\lambda|}\sum_{i=0}^{m-1}\frac{1}{|\lambda|^i} = \|J^{-1}e_m\|_1,$$

where $\|J^{-1}e_m\|_1$ is a geometric progression in $1/|\lambda|$.

The relation between the two upper bounds follows from the fact that $\|x\|_1 \le \sqrt{m}\,\|x\|$ for any m-vector x [17, (2.2.5)]. □

We extend the above bounds to matrices consisting of several Jordan blocks.

THEOREM 8.4. *Let J be a Jordan matrix whose diagonal blocks have diagonal elements λ_i, and let*

$$\epsilon \equiv \min_i |\lambda_i| > 0.$$

Furthermore let m be the order of a Jordan block J_j for which $\|J^{-1}\| = \|J_j^{-1}\|$; and let l be the order of any Jordan block whose diagonal elements have absolute value ϵ. Then

$$\frac{1}{\epsilon^l}\sqrt{\frac{1-\epsilon^{2l}}{1-\epsilon^2}} \le \|J^{-1}\| \le \frac{1}{\epsilon^m}\,\frac{1-\epsilon^m}{1-\epsilon}.$$

Proof. To prove the upper bound, let λ_j be the diagonal element of J_j. Since $\epsilon \le |\lambda_j|$, we bound the one-norm upper bound from Lemma 8.3 in terms of ϵ,

$$\|J^{-1}\| = \|J_j^{-1}\| \le \|J_j^{-1}e_m\|_1 = \frac{1}{|\lambda_j|}\sum_{i=0}^{m-1}\frac{1}{|\lambda_j|^i} \le \frac{1}{\epsilon}\sum_{i=0}^{m-1}\frac{1}{\epsilon^i}.$$

To prove the lower bound, let J_i be any Jordan block whose diagonal element has absolute value ϵ. Then Lemma 8.3 implies

$$\|J^{-1}\| \ge \|J_i^{-1}\| \ge \frac{1}{\epsilon^l}\sqrt{\frac{1-\epsilon^{2l}}{1-\epsilon^2}}.$$

□

9. Appendix 2: Departure from Normality. In this section we present relations between different measures for departure from normality.

The following result relates $\|AA^* - A^*A\|_F$ and $\|N\|_F$ in the Frobenius norm.

THEOREM 9.1 (THEOREM 1 IN [22]). *Let A be $n \times n$ with Schur decomposition $A = Q(\Lambda - N)Q^*$. Then*

$$\|N\|_F^2 \leq \sqrt{\frac{n^3 - n}{12}} \|A^*A - AA^*\|_F.$$

Below is a corresponding relation in the two-norm.

THEOREM 9.2. *Let A be $n \times n$ with Schur decomposition $A = Q(\Lambda - N)Q^*$. Then*

$$\frac{\|N\|^2}{n^2} \leq \|A^*A - AA^*\| \leq 2\|N\|^2 + 4\|\Lambda\|\,\|N\|.$$

Proof. The upper bound follows from the triangle inequality and the submultiplicative property of the two-norm. Regarding the lower bound, there exists a column Ne_l, $1 \leq l \leq n$, such that (cf. §2.5)

$$\frac{1}{n}\|N\|^2 \leq \|Ne_l\|^2 = (N^*N)_{ll},$$

where $(N^*N)_{ll}$ is the lth diagonal element of N^*N. Because

$$\sum_{i=l}^{n} (N^*N - NN^*)_{ii} = (N^*N)_{ll} + \sum_{i=l+1}^{n} \sum_{j=1}^{l-1} N_{ji}^* N_{ji},$$

where N_{ji} is element (j,i) of N,

$$(N^*N)_{ll} \leq \sum_{i=l}^{n} (N^*N - NN^*)_{ii} \leq n \max_{1 \leq i \leq n} |(N^*N - NN^*)_{ii}| \leq n\,\|A^*A - AA^*\|.$$

The last inequality holds because $(N^*N - NN^*)_{ii}$ is the ith diagonal element of $Q^*(A^*A - AA^*)Q$ and because for any matrix M, $\|M\| \geq \max_i |M_{ii}|$. Putting all inequalities together gives the desired lower bound. □

The Henrici number in the Frobenius norm [6, Definition 1.1],[7, Definition 9.1],

$$\mathrm{He_F(A)} \equiv \frac{\|\mathrm{A^*A - AA^*}\|_\mathrm{F}}{\|\mathrm{A}^2\|_\mathrm{F}} \leq 2\,\frac{\|\mathrm{A}\|_\mathrm{F}^2}{\|\mathrm{A}^2\|_\mathrm{F}},$$

is a lower bound for the two-norm eigenvector condition number of a diagonalisable matrix:

THEOREM 9.3 (THEOREM 8 IN [49] ADAPTED TO THE TWO-NORM). *Let A be $n \times n$. If A is diagonalisable with eigenvector matrix V then*

$$\kappa(V)^4 \geq 1 + \frac{1}{2}\,\mathrm{He_F(A)}^2.$$

Acknowledgements. I thank Carl Meyer for encouraging me to write this paper; Stan Eisenstat and Rich Lehoucq for helpful discussions; Bob Funderlic and Gene Golub for pointers to the literature; and Inder Dhillon for carefully reading the paper and making many helpful suggestions.

[1] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Oustrouchov, and D. Sorensen, *LAPACK Users' Guide*, Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[2] K. Bathë and E. Wilson, *Discussion of a paper by K.K. Gupta*, Int. J. Num. Meth. Engng, 6 (1973), pp. 145–6.

[3] ———, *Numerical Methods in Finite Element Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1976.

[4] F. Bauer and C. Fike, *Norms and exclusion theorems*, Numer. Math., 2 (1960), pp. 137–41.

[5] F. Bauer and A. Householder, *Moments and characteristic roots*, Numer. Math., 2 (1960), pp. 42–53.

[6] F. Chaitin-Chatelin, *Is nonnormality a serious difficulty?*, Tech. Rep. TR/PA/94/18, CERFACS, Toulouse, France, 1994.

[7] F. Chaitin-Chatelin and V. Frayssé, *Lectures on Finite Precision Computations*, SIAM, Philadelphia, 1996.

[8] S. Chandrasekaran, *When is a Linear System Ill-Conditioned?*, PhD thesis, Department of Computer Science, Yale University, 1994.

[9] S. Chandrasekaran and I. Ipsen, *Backward errors for eigenvalue and singular value decompositions*, Numer. Math., 68 (1994), pp. 215–23.

[10] F. Chatelin, *Ill conditioned eigenproblems*, in Large Scale Eigenvalue Problems, North-Holland, Amsterdam, 1986, pp. 267–82.

[11] ———, *Valeurs Propres de Matrices*, Masson, Paris, 1988.

[12] C. Davis and W. Kahan, *The rotation of eigenvectors by a perturbation, III*, SIAM J. Numer. Anal., 7 (1970), pp. 1–46.

[13] J. Demmel, *On condition numbers and the distance to the nearest ill-posed problem*, Numer. Math., 51 (1987), pp. 251–89.

[14] S. Eisenstat and I. Ipsen, *Relative perturbation results for eigenvalues and eigenvectors of diagonalisable matrices*, Tech. Rep. CRSC-TR96-6, Center for Research in Scientific Computation, Department of Mathematics, North Carolina State University, 1996.

[15] L. Elsner and M. Paardekooper, *On measures of nonnormality of matrices*, Linear Algebra Appl., 92 (1987), pp. 107–23.

[16] L. Foster, *Gaussian elimination with partial pivoting can fail in practice*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 1354–62.

[17] G. Golub and C. van Loan, *Matrix Computations*, The Johns Hopkins Press, Baltimore, second ed., 1989.

[18] G. Golub and J. Wilkinson, *Ill-conditioned eigensystems and the computation of the Jordan canonical form*, SIAM Review, 18 (1976), pp. 578–619.

[19] D. Greene and D. Knuth, *Mathematics for the Analysis of Algorithms*, Birkhäuser, Boston, 1981.

[20] K. Gupta, *Eigenproblem solution by a combined Sturm sequence and inverse iteration technique*, Int. J. Num. Meth. Engng, 7 (1973), pp. 17–42.

[21] ———, *Development of a unified numerical procedure for free vibration analysis of structures*, Int. J. Num. Meth. Engng, 17 (1981), pp. 187–98.

[22] P. Henrici, *Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices*, Numer. Math., 4 (1962), pp. 24–40.

[23] N. Higham and P. Knight, *Matrix powers in finite precision arithmetic*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 343–58.

[24] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[25] ———, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[26] I. Ipsen, *Helmut Wielandt's contributions to the numerical solution of complex eigenvalue problems*, in Helmut Wielandt, Mathematische Werke, Mathematical Works, B. Huppert and H. Schneider, eds., vol. II: Matrix Theory and Analysis, Walter de Gruyter, Berlin.

[27] ———, *A history of inverse iteration*, in Helmut Wielandt, Mathematische Werke, Mathematical Works, B. Huppert and H. Schneider, eds., vol. II: Matrix Theory and Analysis, Walter de Gruyter, Berlin.

[28] P. Jensen, *The solution of large symmetric eigenvalue problems by sectioning*, SIAM J. Numer. Anal, 9 (1972), pp. 534–45.

[29] E. Jessup and I. Ipsen, *Improving the accuracy of inverse iteration*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 550–72.

[30] W. Kahan, *Numerical linear algebra*, Canadian Math. Bull., 9 (1966), pp. 757–801.

[31] W. Kahan, B. Parlett, and E. Jiang, *Residual bounds on approximate eigensystems of nonnormal matrices*, SIAM J. Numer. Anal., 19 (1982), pp. 470–84.
[32] D. Kahaner, C. Moler, and S. Nash, *Numerical Methods and Software*, Prentice Hall, Englewood Cliffs, NJ, 1989.
[33] T. Kato, *Perturbation Theory for Linear Operators*, Springer Verlag, Berlin, 1995.
[34] P. Lancaster and M. Tismenetsky, *The Theory of Matrices, Second Edition*, Academic Press, Orlando, 1985.
[35] S. Lee, *Bounds for the departure from normality and the Frobenius norm of matrix eigenvalues*, Tech. Rep. ORNL/TM-12853, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 1994.
[36] ———, *A practical upper bound for the departure from normality*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 462–68.
[37] R. Li, *On perturbation bounds for eigenvalues of a matrix*. Unpublished manuscript, November 1985.
[38] M. Osborne, *Inverse iteration, Newton's method, and non-linear eigenvalue problems*, in The Contributions of Dr. J.H. Wilkinson to Numerical Analysis, Symposium Proceedings Series No. 19, The Institute of Mathematics and its Applications, 1978.
[39] ———, June 1996. Private communication.
[40] A. Ostrowski, *On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. I-VI*, Arch. Rational Mech. Anal., 1, 2, 3, 3, 3, 4 (1958-59), pp. 233–41, 423–8, 325–40, 341–7, 472–81, 153–65.
[41] B. Parlett, *The Symmetric Eigenvalue Problem*, Prentice Hall, Englewood Cliffs, NJ, 1980.
[42] B. Parlett and I. Dhillon, *On Fernando's method to find the most redundant equation in a tridiagonal system*, Linear Algebra Appl., (to appear).
[43] B. Parlett and W. Poole, *A geometric theory for the QR, LU and power iterations*, SIAM J. Numer. Anal., 10 (1973), pp. 389–412.
[44] G. Peters and J. Wilkinson, *The calculation of specified eigenvectors by inverse iteration, contribution II/18*, in Linear Algebra, Handbook for Automatic Computation, Volume II, Springer Verlag, Berlin, 1971, pp. 418–39.
[45] ———, *Inverse iteration, ill-conditioned equations and Newton's method*, SIAM Review, 21 (1979), pp. 339–60.
[46] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, Manchester University Press, New York, 1992.
[47] B. Smith, J. Boyle, J. Dongarra, B. Garbow, Y. Ikebe, V. Klema, and C. Moler, *Matrix Eigensystem Routines: EISPACK Guide, 2nd Ed.*, Springer-Verlag, Berlin, 1976.
[48] M. Smith and S. Huttin, *Frequency modification using Newton's method and inverse iteration eigenvector updating*, AIAA Journal, 30 (1992), pp. 1886–91.
[49] R. Smith, *The condition numbers of the matrix eigenvalue problem*, Numer. Math., 10 (1967), pp. 232–40.
[50] G. Stewart, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems*, SIAM Review, 15 (1973), pp. 727–64.
[51] ———, *Introduction to Matrix Computations*, Academic Press, New York, 1973.
[52] L. Trefethen, *Pseudospectra of matrices*, in Numerical Analysis 1991, Longman, 1992, pp. 234–66.
[53] L. Trefethen and R. Schreiber, *Average-case stability of Gaussian elimination*, SIAM J. Matrix Anal. Appl., 11 (1990), pp. 335–60.
[54] J. Varah, *The calculation of the eigenvectors of a general complex matrix by inverse iteration*, Math. Comp., 22 (1968), pp. 785–91.
[55] ———, *On the separation of two matrices*, SIAM J. Numer. Anal., 16 (1979), pp. 216–22.
[56] H. Wielandt, *Beiträge zur mathematischen Behandlung komplexer Eigenwertprobleme, Teil V: Bestimmung höherer Eigenwerte durch gebrochene Iteration*, Bericht B 44/J/37, Aerodynamische Versuchsanstalt Göttingen, Germany, 1944. (11 pages).
[57] J. Wilkinson, *The calculation of the eigenvectors of codiagonal matrices*, Computer J., 1 (1958), pp. 90–6.
[58] ———, *Error analysis for direct methods of matrix inversion*, J. Assoc. Comp. Mach., 8 (1961), pp. 281–330.
[59] ———, *Rigorous error bounds for computed eigensystems*, Computer J., 4 (1961), pp. 230–41.
[60] ———, *Rounding Errors in Algebraic Processes*, Prentice Hall, Englewood Cliffs, NJ, 1963.
[61] ———, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.

[62] ———, *Inverse iteration in theory and practice*, in Symposia Matematica, vol. X, Institutio Nationale di Alta Matematica Monograf, Bologna, 1972, pp. 361–79.

[63] ———, *Note on inverse iteration and ill-conditioned eigensystems*, in Acta Universitatis Carolinae – Mathematica et Physica, No. 1-2, Universita Karlova, Fakulta Matematiky a Fyziky, 1974, pp. 173–7.