Ruprecht-Karls-Universität Heidelberg
Faculty of mathematics and computer science
Institute of applied mathematics

Using complex shifts in Rayleigh Quotient Iteration to compute close eigenvalues

Bachelor Thesis

Author: Nils Friess

Supervisor: Prof. Dr. Robert Scheichl

Submitted: May 28, 2020

Contents

Todo list

1 Introduction

This thesis introduces a novel shift-and-invert method to compute eigenvalues and eigenvectors of real symmetric matrices. The method is a modification of *Rayleigh Quotient Iteration (RQI)*. Numerical examples suggest that our method overcomes some of the drawbacks of RQI at the expense of a slower convergence rate.

The thesis is structured as follows. In this chapter, we introduce a model problem from physics and collect the necessary definitions and results from numerical linear algebra. We also review some of the most basic iterative methods for eigenvalue problems. The second chapter is devoted to the study of the Rayleigh Quotient and Rayleigh Quotient Iteration. Among others, we give a proof for the property that makes RQI advantageous over other methods, viz., the local cubic convergence. In the third chapter we introduce our novel method that we call *Complex Rayleigh Quotient Iteration*. After motivating the method, different numerical experiments are carried out to better understand its behaviour.

1.1 Model Problem

There is a plethora of examples which lead to eigenvalue problems in almost all of the natural sciences, engineering and many other disciplines. The example below leads to an eigenvalue problem with eigenvalues that are very closely spaced. As we will see later, traditional methods do not deal well with this problem.

Example 1.1 (Discretization of Schrödinger operator with band-gap spectrum)**.**

Write up example

1.2 Results from numerical linear algebra

In this section we introduce the definitions and results from linear algebra that will be needed later. We also introduce the most basic iterative eigenvalue algorithms.

1.2.1 The Symmetric Eigenvalue Problem

In Example 1.1 a problem of the form

"Find $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $\lambda \in \mathbb{R}$, such that $\boldsymbol{A}\mathbf{v} = \lambda\mathbf{v}$ "

arised, where $\boldsymbol{A}$ was a real symmetric matrix. This problem is referred to as the *symmetric eigenvalue problem*. A more general definition is given in the following.

Definition 1.2. Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$. A scalar $\lambda \in \mathbb{C}$ is called *eigenvalue* of $\boldsymbol{A}$ if there exists a nonzero vector $\mathbf{v} \in \mathbb{C}^n$ such that

$$\boldsymbol{A}\mathbf{v} = \lambda\mathbf{v}\,. \tag{1.1}$$

The vector $\mathbf{v}$ is called an *eigenvector* of $\boldsymbol{A}$ associated with λ. The tuple $(\lambda, \mathbf{v})$ is called an *eigenpair*. The set of all eigenvalues of $\boldsymbol{A}$ is referred to as the *spectrum* and is denoted by $\sigma(\boldsymbol{A})$. To indicate that eigenvalues belong to a particular matrix $\boldsymbol{M}$ we sometimes write $\lambda(\boldsymbol{M})$.

Computing eigenpairs is a non-trivial task. Rewriting (1.1) gives $\boldsymbol{A}\mathbf{v} - \lambda\mathbf{v} = \mathbf{0}$ or $(\boldsymbol{A} - \lambda\boldsymbol{I})\mathbf{v} = \mathbf{0}$, where $\boldsymbol{I}$ is the identity matrix. Since $\mathbf{v}$ cannot be the zero vector, this equation has a solution if and only if the matrix $\boldsymbol{A} - \lambda\boldsymbol{I}$ is singular. Thus, eigenvalues of $\boldsymbol{A}$ are exactly the roots of the *characteristic polynomial*

$$\chi_{\boldsymbol{A}}(t) \coloneqq \det(\boldsymbol{A} - t\boldsymbol{I})\,.$$

This fact, despite being of theoretical importance, cannot be used to calcu-

late eigenvalues numerically for two reasons. First, the computation of the coefficients of the polynomial is not stable [7, p. 37]. And even if it was, it is well-known that even small perturbations in the coefficients of $\chi_{\boldsymbol{A}}(t)$ can lead to devastating errors in the roots [cf. 28, p. 190]. Thus, other methods are necessary to solve (1.1) which gave rise to iterative algorithms. These methods date back to 1846 when Jacobi published a pioneering paper on a method to compute eigenvalues of symmetric matrices [11]. Below we present essential facts from linear algebra preparing us for discussing such iterative methods in Section 1.2.2.

Remark 1.3 (Generalisations of eigenvalue problems)**.** The problem stated in Equation (1.1) can be generalised in multiple ways. Many problems from physics lead to the *generalised eigenvalue problem*

$$\boldsymbol{A}\mathbf{v} = \lambda \boldsymbol{M}\mathbf{v}\,. \tag{1.2}$$

In our case, we have $\boldsymbol{M} = \boldsymbol{I}$, the identity matrix. Many of the numerical algorithms for solving eigenvalue problems of the form (1.1) can be modified to solve (1.2); often certain assumptions have to be posed on $\boldsymbol{M}$ such as positive definiteness.

Since matrices can be seen as representations of linear operators on finite-dimensional vector spaces, we can define eigenvalue problems for linear operators on more general spaces, that are possibly of infinite dimension. The eigenvectors are then usually called *eigenfunctions*. Other generalisations include the *quadratic eigenvalue problem*

$$(\lambda^2 \boldsymbol{A}_2 + \lambda \boldsymbol{A}_1)\mathbf{v} = \boldsymbol{A}_0\mathbf{v}\,,$$

with matrix coefficients $\boldsymbol{A}_0, \boldsymbol{A}_1, \boldsymbol{A}_2 \in \mathbb{C}^{n \times n}$ or more general *nonlinear eigenproblems*

$$\boldsymbol{Q}(\lambda)\mathbf{v} = \mathbf{0}\,,$$

where $\boldsymbol{Q}(\lambda)$ is a nonlinear matrix-valued function. In this thesis, we only consider problems of the form (1.1).

In the following proposition we collect some basic facts on eigenvalues and eigenvectors. The results are shown under the assumption that $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is a complex Hermitian matrix, i.e., $\boldsymbol{A} = \boldsymbol{A}^* := \overline{\boldsymbol{A}}^\mathsf{T}$, where the bar denotes the complex conjugate. If $\boldsymbol{A}$ is a real matrix, we have $\overline{\boldsymbol{A}} = \boldsymbol{A}$ and thus the following facts hold in particular for real symmetric matrices.

Proposition 1.4. *Let $\boldsymbol{A} = \boldsymbol{A}^* \in \mathbb{C}^{n \times n}$ be a Hermitian matrix. Denote by $\lambda_1, \lambda_2, \ldots, \lambda_n$ the eigenvalues[1] of $\boldsymbol{A}$ with associated eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$.*

(i) All eigenvalues of $\boldsymbol{A}$ are real.

(ii) There exists an orthonormal basis of $\mathbb{C}^n$ consisting of eigenvectors of $\boldsymbol{A}$. If $\boldsymbol{A}$ is a real symmetric matrix, the eigenvectors form an orthonormal basis of $\mathbb{R}^n$.

(iii) If $\boldsymbol{A}$ is non-singular the eigenvalues of $\boldsymbol{A}^{-1}$ are given by $\lambda_1^{-1}, \ldots, \lambda_n^{-1}$ with eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

(iv) Let $\mu \in \mathbb{R}$ an arbitrary scalar. The eigenvalues of $\boldsymbol{A} - \mu \boldsymbol{I}$ are $\lambda_i - \mu$ with eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Proof. Both (i) and (ii) are well-known results from linear algebra and the proofs can be found in most standard literature (see for example [9, Theorem 18 and Corollary, p. 314]).

(iii) Suppose $\boldsymbol{A}$ is invertible and let $(\lambda, \mathbf{v})$ be an eigenpair of $\boldsymbol{A}$ (note that since $\boldsymbol{A}$ is non-singular we have $\lambda \neq 0$). Then

$$\boldsymbol{A}\mathbf{v} = \lambda\mathbf{v} \quad \Leftrightarrow \quad \boldsymbol{A}^{-1}\boldsymbol{A}\mathbf{v} = \lambda\boldsymbol{A}^{-1}\mathbf{v} \quad \Leftrightarrow \quad \lambda^{-1}\mathbf{v} = \boldsymbol{A}^{-1}\mathbf{v},$$

---

[1] Of course, the eigenvalues need not be distinct. But since the eigenvalues of $\boldsymbol{A}$ are the roots of the n-degree polynomial $\chi_{\boldsymbol{A}}(t)$, when counting these roots with their multiplicity, this polynomial has n roots over $\mathbb{C}$. Thus, we can label the eigenvalues from 1 to n.

hence $(\lambda^{-1}, \mathbf{v})$ is an eigenpair of $\boldsymbol{A}^{-1}$.

(iv) For $\mu \in \mathbb{R}$ and $(\lambda, \mathbf{v})$ an eigenpair we have

$$\boldsymbol{A}\mathbf{v} = \lambda\mathbf{v} \quad \Leftrightarrow \quad \boldsymbol{A}\mathbf{v} - \mu\mathbf{v} = \lambda\mathbf{v} - \mu\mathbf{v} \quad \Leftrightarrow \quad (\boldsymbol{A} - \mu\boldsymbol{I})\mathbf{v} = (\lambda - \mu)\mathbf{v},$$

hence $(\lambda - \mu, \mathbf{v})$ is an eigenpair of $\boldsymbol{A} - \mu\boldsymbol{I}$.

$\square$

In the following, we restrict our attention to the *symmetric eigenvalue problem*, i. e., we want to find solutions of equation (1.1) assuming $\boldsymbol{A}$ is a real symmetric matrix and thus, if not stated otherwise, for the remainder of the thesis $\boldsymbol{A}$ denotes a matrix of this type. The (real) eigenvalues are denoted by $\lambda_j(\boldsymbol{A}) = \lambda_j$ with corresponding (real) eigenvectors $\mathbf{v}_j$. Since any scalar multiple of an eigenvector is also an eigenvector, we assume that they are normalised w. r. t. the Euclidean norm, so that

$$\|\mathbf{v}_i\| := \|\mathbf{v}_i\|_2 := \sqrt{\mathbf{v}_i^\mathsf{T}\mathbf{v}_i} = 1 \quad \text{for all } i = 1, \ldots, n.$$

Due to Proposition 1.4 (ii) we have

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \mathbf{v}_i^\mathsf{T}\mathbf{v}_j = 0 \quad \text{for } i \neq j,$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product on $\mathbb{R}^n$. Since all eigenvalues are real we can label them in increasing order of magnitude

$$|\lambda_1| \leq |\lambda_2| \leq \ldots \leq |\lambda_n|.$$

The eigenvalues λ_1 and λ_n are called *extreme* eigenvalues. The remaining eigenvalues $\lambda_2, \ldots, \lambda_{n-1}$ are called *interior* eigenvalues.

In addition to the assumption that the matrix we work with is real and symmetric, we are mainly interested in cases in which the target eigenvalue and

its neighbours are very close. We usually do not have any a priori knowledge about their location. Also, the matrices are assumed to be large such that the computation of the complete set of eigenpairs is too expensive. However, we assume a good approximation of the target eigenvector is available. As we will see, traditional methods do not perform well when the gap between the wanted eigenvalue and adjacent eigenvalues is too small.

1.2.2 Iterative methods for eigenvalue problems

With the necessary facts from linear algebra at hand we can introduce some simple iterative methods for computing eigenpairs of symmetric matrices. We are always interested in how fast these methods produce good approximations of eigenvectors or eigenvalues (or both). The following definition provides us with a notion of the speed at which a sequence converges to its limit.

Definition 1.5 (Order of Convergence)**.** Let $(\mathbf{x}^{(k)})_{k \in \mathbb{N}}$ be a sequence in $\mathbb{C}^n$ that converges to $\mathbf{z} \in \mathbb{C}^n$.

(i) The sequence is said to converge *linearly* to $\mathbf{z}$, if there exists a constant $0 < \rho < 1$ such that
$$\lim_{k \to \infty} \frac{\left\| \mathbf{x}^{(k+1)} - \mathbf{z} \right\|}{\left\| \mathbf{x}^{(k)} - \mathbf{z} \right\|} < \rho \,,$$
where ρ is called the *rate of convergence.*

(ii) The sequence *converges with order q to* $\mathbf{z}$ for $q > 1$ if
$$\lim_{k \to \infty} \frac{\left\| \mathbf{x}^{(k+1)} - \mathbf{z} \right\|}{\left\| \mathbf{x}^{(k)} - \mathbf{z} \right\|^q} < M \,,$$
for some $M > 0$. In particular, convergence with order

- $q = 2$ is called *quadratic convergence,*
- $q = 3$ is called *cubic convergence*

etc.

In some cases, in particular for sequences that approximate eigenvectors, the convergence behaviour is best studied in terms of the *error angle* between $\mathbf{x}^{(k)}$ and $\mathbf{z}$.

Definition 1.6 (Angle)**.** The *angle* between two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ is defined as

$$\angle(\mathbf{x}, \mathbf{y}) = \arccos \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|}{\|\mathbf{x}\| \|\mathbf{y}\|} \,.$$

Often, the following identities are convenient

$$\sin \angle(\mathbf{x}, \mathbf{y}) \coloneqq \sqrt{1 - \cos^2 \angle(\mathbf{x}, \mathbf{y})}, \qquad \tan \angle(\mathbf{x}, \mathbf{y}) \coloneqq \frac{\sin \angle(\mathbf{x}, \mathbf{y})}{\cos \angle(\mathbf{x}, \mathbf{y})} \,.$$

To see why also the angle can be used to measure the convergence speed, suppose $\mathbf{x}^{(k)}$ converges to the unit vector $\mathbf{z}$. Let $\mathbf{u}^{(k)}$ be the unit vector that lies in the span of $\mathbf{x}^{(k)}$ and $\mathbf{z}$ and is orthogonal to $\mathbf{z}$ and denote by $\phi^{(k)} = \angle(\mathbf{x}^{(k)}, \mathbf{z})$ the error angle between the current vector iterate and the limit. Now, write the vector iterate $\mathbf{x}^{(k)}$ as

$$\mathbf{x}^{(k)} = \mathbf{z} \cos \phi^{(k)} + \mathbf{u}^{(k)} \sin \phi^{(k)} \,.$$

We temporarily drop the superscripts and write $\mathbf{x} = \mathbf{x}^{(k)}$, $\mathbf{u} = \mathbf{u}^{(k)}$ and $\phi = \phi^{(k)}$. Then, using the identities $\sin^2(\phi/2) = \frac{1 - \cos \phi}{2}$, $\sin^2 \phi + \cos^2 \phi = 1$ and the Pythagorean theorem we obtain

$$\begin{aligned}
\|\mathbf{x} - \mathbf{z}\|^2 &= \|\mathbf{z} \cos \phi + \mathbf{u} \sin \phi - \mathbf{z}\|^2 \\
&= \|\mathbf{z}(\cos \phi - 1)\|^2 + \|\mathbf{u} \sin \phi\|^2 \\
&= (\cos \phi - 1)^2 + 1 - \cos^2 \phi \\
&= 2(1 - \cos \phi) = 4 \sin^2(\phi/2) \,.
\end{aligned}$$

Thus, convergence orders w. r. t. the norm imply the same convergence orders in

terms of the error angles and vice verca. Note that we did assume convergence of the sequence.

Power method

The *power method* is one of the oldest iterative methods for computing eigenvectors. It is based on generating the sequence $\mathbf{x}^{(k)} := \boldsymbol{A}^k \mathbf{x}^{(0)}$ where $\mathbf{x}^{(0)}$ is a non-zero unit vector. Of course, $\boldsymbol{A}^k$ does not have to be computed explicitly at each step since

$$\boldsymbol{A}^k \mathbf{x} = \boldsymbol{A}(\boldsymbol{A}(\dots \boldsymbol{A}(\boldsymbol{A}\mathbf{x})\dots)).$$

To prevent underflow and overflow errors, $\mathbf{x}^{(k)}$ is normalised at each step. In Algorithm 1.7 we normalise by ensuring that the largest component of the current approximation is equal to one. Of course, other norms can be used. The sequence $\mathbf{x}^{(k)}$ converges to the eigenvector associated with the eigenvalue λ_n under the assumptions that λ_n is dominant (i.e., $|\lambda_n|$ is strictly greater than $|\lambda_{n-1}|$) and that the starting vector $\mathbf{x}^{(0)}$ has a non-vanishing component in the direction of $\mathbf{v}_n$. The advantage of normalising w.r.t. the maximum norm is that the largest component of $|\boldsymbol{A}\mathbf{x}^{(k-1)}|$ converges to the eigenvalue λ_n. Regardless of the normalisation chosen, the method converges linearly with convergence rate

$$\rho = \frac{|\lambda_{n-1}|}{|\lambda_n|}. \tag{1.3}$$

Thus, the method can be very slow if the distance between the eigenvalues λ_n and λ_{n-1} is very small. For a detailed convergence proof, see [25, pp. 86 sq.].

Besides the possible slow convergence rate, the power method will always converge to an eigenvector associated with the dominant eigenvalue λ_n. In many applications, however, one already has a good approximation of another eigenvalue and wants to compute an eigenvector it belongs to. The following method allows for such computations.

Algorithm 1.7: Power method

begin

    Choose nonzero initial vector $\mathbf{x}^{(0)}$

    for $k = 1, 2, \ldots$ *until convergence* **do**

        $\mathbf{x}^{(k)} = \dfrac{1}{\alpha^{(k)}} \boldsymbol{A} \mathbf{x}^{(k-1)}$

        `/*` $\alpha^{(k)}$ `is the component of` $\boldsymbol{A}\mathbf{x}^{(k-1)}$ `with the maximum`

          `modulus` `*/`

(Shifted) Inverse Iteration

The *inverse iteration* is the power method applied to $\boldsymbol{A}^{-1}$ (provided that the inverse exists). Due to Proposition 1.4 (iii) this will produce a sequence of vectors $\mathbf{x}^{(k)}$ converging to the eigenvector associated to the eigenvalue that is smallest in modulus λ_1. Combining this idea with Proposition 1.4 (iv) yields the *shifted inverse iteration*. There, the iterates are defined by

$$\mathbf{x}^{(k)} = \beta(\boldsymbol{A} - \sigma \boldsymbol{I})^{-1} \mathbf{x}^{(k-1)} \,,$$

where β is responsible for normalising $\mathbf{x}^{(k)}$. The smallest eigenvalue in modulus of the shifted matrix $\boldsymbol{A} - \sigma \boldsymbol{I}$ is the eigenvalue of $\boldsymbol{A}$ that is closest to σ. Hence, this method converges to an eigenvector associated with this eigenvalue. Of course, the inverse need not be computed explicitly. Instead, before the loop we can compute the LU decomposition of $\boldsymbol{A} - \sigma \boldsymbol{I}$ (or any other factorisation, if applicable) and solve the system $(\boldsymbol{A} - \sigma \boldsymbol{I})\mathbf{x}^{(k)} = \mathbf{x}^{(k-1)}$ for $\mathbf{x}^{(k)}$. At each step then, only one backward and one forward substitution is required, reducing the complexity from $\mathcal{O}(n^3)$ to $\mathcal{O}(n^2)$. We summarise the results in Algorithm 1.8 (there, we normalise w. r. t. the Euclidean norm).

Since this is essentially the power method (applied to the inverse of $\boldsymbol{A} - \sigma \boldsymbol{I}$) this algorithm still converges linearly. However, if we denote by μ_1 the eigenvalue that is closest to the shift σ and by μ_2 the one that is the next closest one, the

---

Algorithm 1.8: Shifted inverse iteration

Input: Nonzero unit vector $\mathbf{x}^{(0)}$, shift $\sigma \in \mathbb{R}$
Compute $\boldsymbol{LU}$ decomposition $\boldsymbol{A} - \sigma \boldsymbol{I} = \boldsymbol{LU}$
for $k = 1, 2, \ldots$ *until convergence* **do**
    Solve $(\boldsymbol{A} - \sigma \boldsymbol{I})\tilde{\mathbf{x}}^{(k)} = \mathbf{x}^{(k-1)}$ for $\tilde{\mathbf{x}}^{(k)}$
    $\mathbf{x}^{(k)} \leftarrow \tilde{\mathbf{x}}^{(k)}/\|\tilde{\mathbf{x}}^{(k)}\|$

---

eigenvalue of largest modulus of $(\boldsymbol{A} - \sigma \boldsymbol{I})^{-1}$ is $1/(\mu_1 - \sigma)$ and (1.3) suggests that the convergence rate is

$$\rho = \frac{|\mu_1 - \sigma|}{|\mu_2 - \sigma|} \, .$$

Therefore, the method is often used to compute an eigenvector of $\boldsymbol{A}$ if a good approximation of the corresponding eigenvalue is already available. Note, however, that a shift which is very close to an eigenvalue produces a very ill-conditioned linear system. In the case when σ is exactly an eigenvalue the system is even singular and one might expect inverse iteration to fail in these cases since, in general, it is impossible to solve ill-conditioned systems accurately. Despite this seemingly sincere problem, in practice it was observed that the method produces good approximations even in these cases. According to Parlett [21, pp. 84 sq.], it was Wilkinson who elucidated why the ill-conditioning is not a problem in most cases. Suppose $\sigma \approx \lambda$ where λ is an eigenvalue of $\boldsymbol{A}$ with corresponding eigenvector $\mathbf{v}$. Wilkinson illustrated that although $\tilde{\mathbf{x}}^{(k)}$ may be far from $\mathbf{v}$, the normalised solution $\mathbf{x}^{(k)} = \tilde{\mathbf{x}}^{(k)}/\|\tilde{\mathbf{x}}^{(k)}\|$ will not be far from $\mathbf{v}$, when the system is solved backwards-stably, for more details see [30, pp. 621–630], [21, pp. 68–71] and [23]. This will become important again later when we discuss Rayleigh Quotient Iteration. There, the system that is solved gets increasingly ill-conditioned at each step but for the same reason as above, in practice this poses no problem.

Note that we did not specify the "until convergence" criteria in neither of

the algorithms above. We postpone this discussion until Section 2.1.

At each step in the shifted inverse iteration, better approximations for the target eigenvector are computed. One could try to use these approximations to replace occasionally the shift by an approximation of the corresponding eigenvalue. There are different techniques to obtain such estimates, e.g. the *Wielandt Shifted Inverse Iteration* or the *Rayleigh Quotient Iteration*, the latter of which is rigorously studied in the next chapter. For further discussion on the variants and developments of these so called *shift and invert* techniques see, e.g. [7, 10, 27].

2 Classic Rayleigh Quotient Iteration

In this chapter we thoroughly introduce Rayleigh Quotient Iteration (or RQI, for short) and discuss some important results. Among others, we give a proof of the (local) cubic convergence of RQI. We provide numerical examples for the presented results and discuss some disadvantages.

As mentioned above, we assume that we have given an approximation of an eigenvector of a real symmetric matrix $\boldsymbol{A}$ and we want to compute the exact eigenvector and corresponding eigenvalue.

2.1 The Rayleigh Quotient

In Chapter 1 we briefly introduced some simple iterative eigenvalue methods. In essence, RQI is shifted inverse iteration where the shift is replaced by the *Rayleigh quotient* at each step.

Definition 2.1 (Rayleigh Quotient)**.** Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$. The mapping

$$\mathcal{R}_{\boldsymbol{A}} : \mathbb{C}^n \setminus \{\mathbf{0}\} \to \mathbb{C}, \qquad \mathbf{x} \mapsto \frac{\mathbf{x}^* \boldsymbol{A} \mathbf{x}}{\mathbf{x}^* \mathbf{x}}$$

is called the *Rayleigh quotient*[1] corresponding to the matrix $\boldsymbol{A}$.

Note that for real matrices, we have

$$\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) = \frac{\mathbf{x}^\mathsf{T} \boldsymbol{A} \mathbf{x}}{\mathbf{x}^\mathsf{T} \mathbf{x}}$$

We begin by discussing some basic facts.

---

[1]Other notations that are popular in the literature include $R_{\boldsymbol{A}}(\mathbf{x})$, $R(\boldsymbol{A}, \mathbf{x})$, $r_{\boldsymbol{A}}(\mathbf{x})$, $\sigma_{\boldsymbol{A}}(\mathbf{x})$ or $\rho_{\boldsymbol{A}}(\mathbf{x})$.

Lemma 2.2. *Let* $\mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$, $0 \neq \alpha, \beta \in \mathbb{C}$ *and* $\boldsymbol{A} \in \mathbb{C}^{n \times n}$.

(i) If $(\lambda, \mathbf{v})$ *is an eigenpair of* $\boldsymbol{A}$, *then* $\mathcal{R}_{\boldsymbol{A}}(\mathbf{v}) = \lambda$.

(ii) $\mathcal{R}_{\beta \boldsymbol{A}}(\alpha \mathbf{x}) = \beta \mathcal{R}_{\boldsymbol{A}}(\mathbf{x})$ *(Homogeneity)*

(iii) $\mathcal{R}_{\boldsymbol{A} - \alpha \boldsymbol{I}}(\mathbf{x}) = \mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) - \alpha$ *(Translation invariance)*

Proof.   (i) We can write the Rayleigh Quotient as

$$\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) = \frac{\langle \mathbf{x}, \boldsymbol{A}\mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle}, \tag{2.1}$$

where $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* \mathbf{y}$ denotes the Euclidean inner product on $\mathbb{C}^n$. Due to the linearity in the second argument we obtain

$$\mathcal{R}_{\boldsymbol{A}}(\mathbf{v}) = \frac{\langle \mathbf{v}, \boldsymbol{A}\mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} = \frac{\langle \mathbf{v}, \lambda\mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} = \lambda \frac{\langle \mathbf{v}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} = \lambda.$$

(ii) By again writing the Rayleigh Quotient as in (2.1) and using the semi-linearity in the first and linearity in the second argument of the inner product, we obtain

$$\mathcal{R}_{\beta \boldsymbol{A}}(\alpha \mathbf{x}) = \frac{\langle \alpha \mathbf{x}, \beta \boldsymbol{A}(\alpha \mathbf{x}) \rangle}{\langle \alpha \mathbf{x}, \alpha \mathbf{x} \rangle} = \beta \frac{\overline{\alpha} \alpha \langle \mathbf{x}, \boldsymbol{A}\mathbf{x} \rangle}{\overline{\alpha} \alpha \langle \mathbf{x}, \mathbf{x} \rangle} = \beta \mathcal{R}_{\boldsymbol{A}}(\mathbf{x}).$$

(iii)

$$\mathcal{R}_{\boldsymbol{A} - \alpha \boldsymbol{I}}(\mathbf{x}) = \frac{\mathbf{x}^*(\boldsymbol{A} - \alpha \boldsymbol{I})\mathbf{x}}{\mathbf{x}^* \mathbf{x}} = \frac{\mathbf{x}^* \boldsymbol{A} \mathbf{x} - \alpha \mathbf{x}^* \mathbf{x}}{\mathbf{x}^* \mathbf{x}} = \mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) - \alpha.$$

$\square$

Although the Rayleigh Quotient might look arbitrary at first sight, it occurs naturally as the solution of a least squares minimisation problem. First note that if $(\lambda, \mathbf{v})$ is an eigenpair of $\boldsymbol{A}$

$$\| \boldsymbol{A}\mathbf{v} - \lambda\mathbf{v} \| = 0.$$

Now, suppose $\hat{\mathbf{v}}$ is an approximation for $\mathbf{v}$ and we want to find the best approximation $\hat{\lambda}$ for λ in the sense that

$$\hat{\lambda} = \arg\min_{\mu \in \mathbb{C}} \|\boldsymbol{A}\hat{\mathbf{v}} - \mu\hat{\mathbf{v}}\| \,.$$

This is a linear least squares problem in μ with normal equations (see, e.g. [6, Theorem 6.12, p. 362])

$$(\hat{\mathbf{v}}^*\hat{\mathbf{v}})\,\mu = \hat{\mathbf{v}}^*\boldsymbol{A}\hat{\mathbf{v}}$$

and dividing by $(\hat{\mathbf{v}}^*\hat{\mathbf{v}})$ yields the solution

$$\mu = \frac{\hat{\mathbf{v}}^*\boldsymbol{A}\hat{\mathbf{v}}}{\hat{\mathbf{v}}^*\hat{\mathbf{v}}} = \mathcal{R}_{\boldsymbol{A}}(\hat{\mathbf{v}})\,.$$

That is, the Rayleigh Quotient is the choice of μ that minimises the *residual* norm for the eigenvalue problem. The following result specifies how good of an estimate the Rayleigh Quotient is. According to Parlett this is the property to which "the phenomenal convergence rate [of RQI] can be attributed" [21, p. 77]. We postpone the proof of this statement until when we discuss the convergence behaviour of RQI.

Lemma 2.3 (Eigenvalue estimate)**.** *Let* $\mathbf{x} \in \mathbb{C}^n$ *be an approximation of an eigenvector* $\mathbf{v}$ *of a normal[2] matrix* $\boldsymbol{A}$ *with corresponding eigenvalue* λ*. Then*

$$|\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) - \mathcal{R}_{\boldsymbol{A}}(\mathbf{v})| = |\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) - \lambda| = \mathcal{O}\left(\|\mathbf{x} - \mathbf{v}\|^2\right)\,.$$

This result is often paraphrased as "the Rayleigh Quotient is a *quadratically accurate* estimate of an eigenvalue" (see for example [28, p. 204]). If $\boldsymbol{A}$ is non-normal, the Rayleigh quotient is still an estimate of order one, i.e.,

$$|\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) - \mathcal{R}_{\boldsymbol{A}}(\mathbf{v})| = \mathcal{O}(\|\mathbf{x} - \mathbf{v}\|)\,.$$

---

[2]A matrix $\boldsymbol{A}$ is said to be *normal* if $\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^*$. Note that for complex Hermitian (or real symmetric) matrices we have $\boldsymbol{A} = \boldsymbol{A}^*$, hence Hermitian (and thus symmetric) matrices are normal.

We now have a method that allows us to obtain an estimation of an *eigenvalue* from an *eigenvector*. With the Shifted Inverse Iteration (Algorithm 1.8) we have the converse: a method for obtaining an *eigenvector* estimate from an *eigenvalue* estimate. Rayleigh Quotient Iteration is essentially a combination of those two methods where each step consists of one step of Shifted Inverse Iteration and the computation of the Rayleigh Quotient.

Algorithm 2.4: Rayleigh Quotient Iteration

Input: Nonzero unit vector $\mathbf{x}^{(0)}$

$\mu^{(0)} \leftarrow \left(\mathbf{x}^{(0)}\right)^* \boldsymbol{A} \mathbf{x}^{(0)}$

for $k = 1, 2, \ldots$ *until convergence* **do**

    Solve $\left(\boldsymbol{A} - \mu^{(k)} \boldsymbol{I}\right) \mathbf{y}^{(k)} = \mathbf{x}^{(k-1)}$ for $\mathbf{y}^{(k)}$

    $\mathbf{x}^{(k)} \leftarrow \mathbf{y}^{(k)} / \left\|\mathbf{y}^{(k)}\right\|$

    $\mu^{(k)} \leftarrow \left(\mathbf{x}^{(k)}\right)^* \boldsymbol{A} \mathbf{x}^{(k)}$

We have yet to define what we mean by "until convergence" in this Algorithm but also in the simple vector iterations from Chapter 1. Now that we the defined the Rayleigh Quotient, we can define the following stopping criterion. Run the iteration until

$$\left\|\mathbf{r}^{(k)}\right\| = \left\|\boldsymbol{A}\mathbf{x}^{(k)} - \mu^{(k)}\mathbf{x}^{(k)}\right\| < \texttt{tol}\,,$$

where $\mathbf{r}^{(k)} := \boldsymbol{A}\mathbf{x}^{(k)} - \mu^{(k)}\mathbf{x}^{(k)}$ is called the *residual vector* and $\texttt{tol}$ is a user-given error tolerance. In Algorithm 1.7 (Power method) and Algorithm 1.8 (Inverse Iteration) the Rayleigh Quotient of the current vector iterate $\mu^{(k)} = (\mathbf{x}^{(k)})^* \boldsymbol{A} \mathbf{x}^{(k)}$ is not computed. Therefore, for this *residual-based error control* to work, the computation of $\mu^{(k)}$ has to be added to these algorithms.

Obviously, if for some k the tuple $(\mu^{(k)}, \mathbf{x}^{(k)})$ is an eigenpair we have $\mathbf{r}^{(k)} = \mathbf{0}$. For approximate eigenpairs we expect a small residual to imply small errors in these approximations. Details on this residual-based error control can be found in [25, Section 3.2] or [3, Section 5.2]. Here, we give only some important

results without proof for the case when $\boldsymbol{A}$ is Hermitian. A popular result, usually referred to as the *Bauer-Fike theorem* for Hermitian matrices (see, e. g. [25, p. 59]) states that there exists an eigenvalue λ of $\boldsymbol{A}$ such that

$$\left|\lambda - \mu^{(k)}\right| \leq \left\|\mathbf{r}^{(k)}\right\|. \tag{2.2}$$

Thus, if the stopping criterion is fulfilled, we have $|\lambda - \mu^{(k)}| < \texttt{tol}$. For the eigenvector one can show [25, p. 63] that the following bound holds

$$\sin \angle(\mathbf{x}^{(k)}, \mathbf{v}) \leq \frac{\left\|\mathbf{r}^{(k)}\right\|}{\delta},$$

where $\mathbf{v}$ is an eigenvector associated with λ and δ is the distance from $\mu^{(k)}$ to the rest of the spectrum, i. e.,

$$\delta := \min_{i} \left\{ |\mu^{(k)} - \lambda_i| \ : \ \lambda_i \neq \lambda \right\}.$$

2.2 History and recent developments

Now that we have defined the Rayleigh Quotient and Rayleigh Quotient Iteration we give an overview of the historic developments of RQI. We also discuss recent contributions that are relevant to this thesis. Some of the results, mainly the ones that are concerned with the convergence of RQI, are discussed in more detail in Section 2.3.

2.2.1 Chronology of Rayleigh Quotient iteration

It took about 60 years from the first mention of what is now called the Rayleigh quotient until RQI was fully defined as it is given in Algorithm 2.4. In this section we give an overview of some important milestones within this 60 years. Large parts of this overview are based on [27].

Add pages

1894 — Lord Rayleigh   In the second edition of his book titled "The Theory of Sound" John William Strutt, 3rd Baron Rayleigh [24, p. 110], proposed the following iteration for improving an approximate eigenvector $\mathbf{x}^{(0)}$:

$$\text{Solve} \quad (\boldsymbol{A} - \mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(i)})\boldsymbol{I})\,\mathbf{x}^{(i+1)} = \mathbf{e}_1\,, \tag{2.3}$$

where $\mathbf{e}_1$ denotes the first natural coordinate vector, i.e., the first column of the identity matrix and $\mathbf{x}^{(i)}$ and $\mathbf{x}^{(i+1)}$ denote the current and next iterate, respectively. In fact, Lord Rayleigh considered the generalised eigenvalue problem and so in his text, the iteration reads

$$\text{Solve} \quad (\boldsymbol{A} - \mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(i)})\boldsymbol{M})\,\mathbf{x}^{(i+1)} = \mathbf{e}_1\,.$$

1949 — Kohn   In a letter to the editor Walter Kohn [12] suggest the following iteration

$$\text{Solve} \quad (\boldsymbol{A} - \mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(i)})\boldsymbol{I})\mathbf{x}^{(i+1)} = \mathbf{e}_k\,,$$

where $\mathbf{e}_k$ is *any* of the natural coordinate vectors. Without a rigorous proof Kohn argues that $\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(i)})$ converges quadratically to an eigenvalue of $\boldsymbol{A}$ (provided that $\mathbf{x}^{(0)}$ is sufficiently close to an eigenvector of $\boldsymbol{A}$). Despite the similarity to (2.3), Kohn does not mention Lord Rayleigh's method and it is not known whether or not he was aware of it.

1951 — Crandall   In a text communicated to the Royal society of London Stephen Crandall [4] suggests

$$\text{Solve} \quad (\boldsymbol{A} - \mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(i)})\boldsymbol{I})\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)}\,. \tag{2.4}$$

Actually, Crandall also considered the generalised eigenproblem but for our comparative purposes it is sufficient to consider the case $\boldsymbol{M} = \boldsymbol{I}$. Note, that this algorithm is RQI without the normalisation step. Based on the assumption that the sequence of vectors $\mathbf{x}^{(k)}$ converges, Crandall establishes cubic convergence

in the this sequence and the approximate eigenvalue sequence. To see why the assumption is wrong, we assume the contrary, i. e., suppose that the sequence converges to an eigenvector, say $\mathbf{v}_k$. From (2.4) we have

$$\boldsymbol{A}\mathbf{v}_k - \mathcal{R}_{\boldsymbol{A}}(\mathbf{v}_k)\mathbf{v}_k = \mathbf{v}_k \qquad \Leftrightarrow \qquad \boldsymbol{A}\mathbf{v}_k = (1 + \mathcal{R}_{\boldsymbol{A}}(\mathbf{v}_k))\mathbf{v}_k \,,$$

and so $\mathbf{v}_k$ is an eigenvector of $\boldsymbol{A}$ with corresponding eigenvalue $1 + \mathcal{R}_{\boldsymbol{A}}(\mathbf{v}_k)$. Since we know that for any eigenvector, the value of $\mathcal{R}_{\boldsymbol{A}}(\mathbf{v}_k)$ is the eigenvalue it belongs to we have

$$\mathcal{R}_{\boldsymbol{A}}(\mathbf{v}_k) = 1 + \mathcal{R}_{\boldsymbol{A}}(\mathbf{v}_k)$$

which is a contradiction.

Also, Crandall establishes what is usually referred to as *r*-cubic convergence. This is a weaker notion of convergence than the one defined earlier, which is called *q*-cubic convergence (see [27, Appendix A.1] for a detailed definition of these notions of convergence).

1957 − **59 — Ostrowski** Alexander Ostrowski published a series of six papers titled "On the Convergence of the Rayleigh Quotient Iteration for the Computation of the Characteristic Roots and Vectors. I-VI" [13–18]. We mention the titles here, as they represent the first mention of the term *Rayleigh Quotient Iteration.*

In the first paper the author suggests the iteration

$$\text{Solve} \quad (\boldsymbol{A} - \mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(i)})\boldsymbol{I})\mathbf{x}^{(i+1)} = \boldsymbol{\eta} \,, \quad \boldsymbol{\eta} \neq \boldsymbol{0} \,. \tag{2.5}$$

He rigorously establishes a *quadratic* convergence rate for the sequence of Rayleigh Quotients $\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(i)})$. He then refers to a paper of Wielandt [29] and his *fractional* or *broken iteration* (German: *gebrochene Iteration*). Inspired by Wielandt's method he suggests replacing the fixed vector $\boldsymbol{\eta}$ in the right hand

side of (2.5) to the solution from the previous step

$$(\boldsymbol{A} - \mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(i)})\boldsymbol{I})\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} \,, \tag{2.6}$$

starting with an arbitrary non-zero vector $\mathbf{x}^{(0)}$. He then gives a rigorous proof of the local *cubic* convergence of the sequence of Rayleigh quotients $\mu_i \coloneqq \mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(i)})$, i. e.,

$$\frac{\mu_{i+1} - \lambda}{(\mu_i - \lambda)^3} \longrightarrow \gamma \quad \text{as } i \to \infty \,, \tag{2.7}$$

where λ is an eigenvalue of $\boldsymbol{A}$ and γ is a positive constant. Local convergence here means that $\mathbf{x}^{(0)}$ is assumed to be near the eigenvector corresponding to λ.

Note that (2.6) is the same algorithm previously proposed by Crandall given in Equation (2.4). Ostrowski was not aware of Crandall's method; however, while the first paper was in press the following note was added:

> "Professor G. Forsythe has directed my attention to a paper by S. H. Crandall [...]. In particular, Professor Crandall establishes the *cubic character* of convergence of ξ_x in the rule (28), (29). However he does not arrive at our asymptotic formula (46), which is the principal result of our paper." [13, p. 241].[3]

In the beginning of the second paper [14] Ostrowski discusses this in more detail and remarks that Crandall proofed the *r*-cubic convergence of the sequence of eigenvalue iterates, while he showed *q*-cubic convergence.

More importantly, he also points out in §21 of the text that in order to assure convergence in the vector iterates (and not just the Rayleigh Quotients) one needs to normalise the vectors. With this small yet important modification of Crandall's algorithm he fully defined RQI as it is known today.

The third paper [15] of the series addresses the non-symmetric case for which Ostrowski is also able to define a method that attains local cubic convergence.

---

[3]Here, ξ_x denotes the x-th iterate of the approximate eigenvector, i. e., in our notation $\mathbf{x}^{(x)}$. The rule (28), (29) in Ostrowski's paper corresponds to our equation (2.6) and the asymptotic formula (46) he references is given in (2.7).

This method uses a generalised notion of the Rayleigh quotient and comes at the expense of solving two linear systems at each step instead of one.

This concludes the overview of the development of RQI from Lord Rayleigh's iteration that lacks the changing right hand side and the normalisation of the vector iterate to the complete definition of RQI by Ostrowski.

Check content of papers 4-6 of Ostrowski

2.2.2 Further developments and recent contributions

We now give an overview of the developments and contributions after the introduction of RQI in 1958. We also look at some recent contributions.

After Ostrowski's proof of the local cubic convergence, it took another ten years until a first important result concerning the *global* convergence behaviour of RQI was presented. In 1968 Parlett and Kahan [22] showed that RQI applied to symmetric matrices converges for almost all starting vectors and in 1974 Parlett [20] generalised the result for the case when $\boldsymbol{A}$ is normal. A slightly more concise version of the proof has been again published in Parlett's book [21].

Of course, not only is it of interest *that* the method converges but also *which* eigenpair it converges to. The following example demonstrates that the convergence behaviour of RQI might sometimes be unexpected.

Example 2.5. This example is based on an example from [19, p. 254]. Let $\boldsymbol{A} = \mathrm{diag}(1,\ 2,\ 4)$ with eigenvalues $\lambda_1 = 1,\ \lambda_2 = 2,\ \lambda_3 = 4$ and corresponding eigenvectors $\mathbf{e}_i$, the columns of the identity matrix.

(a) First consider RQI started with the vector

$$\mathbf{x}^{(0)} = \begin{pmatrix} 0.8163392507169525 \\ -0.0004821161298470036 \\ 0.5775725022046341 \end{pmatrix}.$$

This produces the eigenpair $(1, \mathbf{e}_1)$ although

$$\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(0)}) = 2.000770218344729\,.$$

Note, however, that computing the angle between $\mathbf{x}^{(0)}$ and the eigenvectors gives

$$\angle(\mathbf{x}^{(0)}, \mathbf{e}_1) \approx 0.61575\,, \quad \angle(\mathbf{x}^{(0)}, \mathbf{e}_2) \approx 1.57128\,, \quad \angle(\mathbf{x}^{(0)}, \mathbf{e}_3) \approx 0.95504\,,$$

that is, the angle between the initial vector and the resulting eigenvector is smallest.

(b) Now take
$$\mathbf{x}^{(0)} = (0.74278,\ 0.55709,\ 0.37139)^{\mathsf{T}}$$

the angles are

$$\angle(\mathbf{x}^{(0)}, \mathbf{e}_1) \approx 0.73358\,, \quad \angle(\mathbf{x}^{(0)}, \mathbf{e}_2) \approx 0.97992\,, \quad \angle(\mathbf{x}^{(0)}, \mathbf{e}_3) \approx 1.19029$$

and the initial shift is $\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(0)}) \approx 1.7241$. In this case, RQI converges to the eigenpair $(2, \mathbf{e}_2)$, i.e., not the eigenvector that has the smallest angle with $\mathbf{x}^{(0)}$.

This example shows that there seems to be no obvious way how the computed eigenpair depends on $\mathbf{x}^{(0)}$ or $\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(0)})$, and indeed the characterisation of the global convergence behaviour of RQI is not straightforward. In contrast, the local convergence behaviour is better understood. Ostrowski [13] defined explicit estimates of *convergence neighbourhoods* for the shift, i.e., intervals around an eigenvalue in which convergence to this eigenvalue is assured. However, these intervals depend on quantities that are not known beforehand so they are of little practical use.

Efforts have been made to obtain local convergence regions with as little knowledge about the spectrum as possible. For example, Beattie and Fox [2] derive conditions under which convergence is assured to be in a given interval assuming the number of eigenvalues contained in the interval is known. Szyld [26] suggests combining shifted inverse iteration and RQI to ensure that the computed eigenvalue lies in a given interval. He obtains criteria to switch

back and forth between the two algorithms to benefit from their respective advantages (inverse iteration is guarantueed to converge to the nearest eigenvalue but convergence is merely linear; RQI posseses local cubic convergence but the global convergence behaviour is possibly erratic).

The task of identifying global convergence regions in terms of the initial vector seems to be more difficult. Parlett noted in 1980 that "There appears to be no simple description of how $\mathbf{v}$ depends on $\mathbf{x}^{(0)}$" [21, p. 82] and it looks as if this statement still holds true today. Pantazis and Szyld [19] and Batterson and Smillie [1] studied *basins of attraction*, i. e., regions in the unit sphere from which RQI will converge to a specific eigenvector. They did, however, only consider the three-dimensional case.

Besides the unpredictablity of the outcome of RQI another major drawback is its high cost. At every iteration a linear system has to be solved and since the system changes at each step one cannot factorize the matrix beforehand as was the case in inverse iteration. There are several obvious possiblities to reduce the cost such as to change the shift only occasionally. Sometimes, inverse iteration is run a fixed number of steps before changing the shift to the Rayleigh Quotient and Szyld's paper [26] that was mentioned above can also be interpreted as a method to reduce the computational cost for large problems.

> Discuss recent contributions (e. g. , inexact RQI, Multistep RQI?, sparse PCA using RQI?)

2.3 Convergence Analysis

We mentioned earlier the local cubic convergence of RQI. In this section we give a proof of this property and we also discuss some results concerning the global convergence behaviour. We do not give a full proof of the global convergence, since parts of it are very technical and, at least for our purposes, it does not give much insight in the method.

As we have seen in the previous section, the first rigorous proof of the cubic convergence of RQI was given by Ostrowski [13] but over time a number of different, simpler proofs were presented. The one we give below follows closely

the one given by Demmel [5, pp. 215 sq.].

Theorem 2.6 (Cubic convergence)**.** *Rayleigh Quotient Iteration is locally cubically convergent.*

Before giving the proof, we make some remarks. *Locally* here means that

1. it is assumed that the sequence $(\mathbf{x}^{(i)})_i$ does converge to an eigenvector of $\boldsymbol{A}$ and

2. there is a finite number of iterations for which convergence might not be cubic (the *preasymptotic phase*). For some $j \in \mathbb{N}$, however, the j-th iterate is a sufficiently good approximation of the target eigenvector, such that the sequence of subsequent iterates does converge cubically.

We defined the notion of cubic convergence in the first chapter, nonetheless we remark again here that this means the number of correct digits *triples* at each step once the error is small enough. And even if the method might not converge cubically from the beginning, in practice the preasymptotic phase rarely takes more than three steps.

Some of the proofs in the literature are given under the assumption that $\boldsymbol{A}$ is diagonal. To see why no generality is lost with this assumption, consider the eigendecomposition of $\boldsymbol{A}$ and write $\boldsymbol{Q}^\mathsf{T}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{\Lambda}$, where $\boldsymbol{Q}$ is an orthogonal matrix consisting of the eigenvectors of $\boldsymbol{A}$ as its columns and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix of eigenvalues. We discard the superscripts for a moment and change variables in RQI[4] to $\hat{\mathbf{x}} \coloneqq \boldsymbol{Q}^\mathsf{T}\mathbf{x}$ and $\hat{\mathbf{y}} \coloneqq \boldsymbol{Q}^\mathsf{T}\mathbf{y}$. Then

$$\mu^{(i)} = \mu = \mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) = \frac{\mathbf{x}^\mathsf{T}\boldsymbol{A}\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}} = \frac{\hat{\mathbf{x}}^\mathsf{T}\boldsymbol{Q}^\mathsf{T}\boldsymbol{A}\boldsymbol{Q}\hat{\mathbf{x}}}{\hat{\mathbf{x}}^\mathsf{T}\boldsymbol{Q}^\mathsf{T}\boldsymbol{Q}\hat{\mathbf{x}}} = \frac{\hat{\mathbf{x}}^\mathsf{T}\boldsymbol{\Lambda}\hat{\mathbf{x}}}{\hat{\mathbf{x}}^\mathsf{T}\hat{\mathbf{x}}} = \mathcal{R}_{\boldsymbol{\Lambda}}(\hat{\mathbf{x}}) = \mathcal{R}_{\boldsymbol{\Lambda}}(\hat{\mathbf{x}}^{(i)}),$$

and $\boldsymbol{Q}\hat{\mathbf{y}}^{(i+1)} = (\boldsymbol{A} - \mu^{(i)}\boldsymbol{I})^{-1}\boldsymbol{Q}\hat{\mathbf{x}}^{(i)}$. Hence,

$$\hat{\mathbf{y}}^{(i+1)} = \boldsymbol{Q}^\mathsf{T}(\boldsymbol{A} - \mu^{(i)}\boldsymbol{I})^{-1}\boldsymbol{Q}\hat{\mathbf{x}}^{(i)} = (\boldsymbol{Q}^\mathsf{T}\boldsymbol{A}\boldsymbol{Q} - \mu^{(i)}\boldsymbol{I})^{-1}\hat{\mathbf{x}}^{(i)} = (\boldsymbol{\Lambda} - \mu^{(i)}\boldsymbol{I})^{-1}\hat{\mathbf{x}}^{(i)},$$

---

[4]Recall the notation from Algorithm 2.4 where $\mathbf{y} = \mathbf{y}^{(i)}$ denotes the unnormalised i-th iterate and $\mathbf{x} = \mathbf{x}^{(i)}$ denotes the same iterate after normalisation.

Citation [Parlett]

where we used $\boldsymbol{Q}^{\mathsf{T}} = \boldsymbol{Q}^{-1}$, the orthogonality of $\boldsymbol{Q}$. We see that running RQI with $\boldsymbol{A}$ and $\mathbf{x}^{(0)}$ is equivalent to running RQI with $\boldsymbol{\Lambda}$ and $\hat{\mathbf{x}}^{(0)}$. Thus, we assume that $\boldsymbol{A} = \boldsymbol{\Lambda}$ is already diagonal which in particular implies that the eigenvectors of $\boldsymbol{A}$ are $\mathbf{e}_i$, the natural coordinate vectors.

Proof of Theorem 2.6. Assume that the eigenvalue that is computed is *simple* and suppose without loss of generality that $\mathbf{x}^{(i)}$ converges to $\mathbf{e}_1$. Remember that we assume that the current iterate is a sufficiently good estimate of $\mathbf{e}_1$, such that for some i we can write $\mathbf{x}^{(i)} = \mathbf{e}_1 + \mathbf{d}^{(i)}$ with $\|\mathbf{d}^{(i)}\| = \epsilon \ll 1$. To show cubic convergence, we have to verify that

$$\lim_{i \to \infty} \frac{\|\mathbf{x}^{(i+1)} - \mathbf{e}_1\|}{\|\mathbf{x}^{(i)} - \mathbf{e}_1\|^3} < M \,,$$

for some positive constant M. We know that $\|\mathbf{x}^{(i)} - \mathbf{e}_1\|^3 = \epsilon^3$, hence it suffices to show that $\|\mathbf{x}^{(i+1)} - \mathbf{e}_1\| = \mathcal{O}(\epsilon^3)$. In other words we have to show that $\mathbf{x}^{(i+1)} = \mathbf{e}_1 + \mathbf{d}^{(i+1)}$ with $\|\mathbf{d}^{(i+1)}\| = \mathcal{O}(\epsilon^3)$.

Since the vectors are normalised at each step we have

$$\begin{aligned}
1 = (\mathbf{x}^{(i)})^{\mathsf{T}}(\mathbf{x}^{(i)}) = (\mathbf{e}_1 + \mathbf{d}^{(i)})^{\mathsf{T}}(\mathbf{e}_1 + \mathbf{d}^{(i)}) &= \mathbf{e}_1^{\mathsf{T}}\mathbf{e}_1 + 2\mathbf{e}_1^{\mathsf{T}}\mathbf{d}^{(i)} + (\mathbf{d}^{(i)})^{\mathsf{T}}\mathbf{d}^{(i)} \\
&= 1 + 2d_1^{(i)} + \epsilon^2 \,.
\end{aligned}$$

where $d_1^{(i)}$ denotes the first component of the vector $\mathbf{d}^{(i)}$. Rewriting gives $d_1^{(i)} = -\epsilon^2/2$ and using the symmetry[5] of $\boldsymbol{\Lambda}$ we obtain

$$\begin{aligned}
\mu^{(i)} = (\mathbf{x}^{(i)})^{\mathsf{T}}\boldsymbol{\Lambda}\mathbf{x}^{(i)} = (\mathbf{e}_1 + \mathbf{d}^{(i)})^{\mathsf{T}}\boldsymbol{\Lambda}(\mathbf{e}_1 + \mathbf{d}^{(i)}) \\
= \mathbf{e}_1^{\mathsf{T}}\boldsymbol{\Lambda}\mathbf{e}_1 + 2\mathbf{e}_1^{\mathsf{T}}\boldsymbol{\Lambda}\mathbf{d}^{(i)} + (\mathbf{d}^{(i)})^{\mathsf{T}}\boldsymbol{\Lambda}\mathbf{d}^{(i)} = \lambda_1 - \eta \,,
\end{aligned}$$

---

[5]For a symmetric matrix $\boldsymbol{B} = \boldsymbol{B}^{\mathsf{T}}$ holds

$$\mathbf{u}^{\mathsf{T}}\boldsymbol{B}\mathbf{w} = (\mathbf{u}^{\mathsf{T}}\boldsymbol{B}\mathbf{w})^{\mathsf{T}} = \mathbf{w}^{\mathsf{T}}(\mathbf{u}^{\mathsf{T}}\boldsymbol{B})^{\mathsf{T}} = \mathbf{w}^{\mathsf{T}}\boldsymbol{B}^{\mathsf{T}}\mathbf{u} = \mathbf{w}^{\mathsf{T}}\boldsymbol{B}\mathbf{u} \,.$$

where $\eta := -2\mathbf{e}_1^\mathsf{T}\mathbf{\Lambda}\mathbf{d}^{(i)} - (\mathbf{d}^{(i)})^\mathsf{T}\mathbf{\Lambda}\mathbf{d}^{(i)} = \lambda_1\epsilon^2 - (\mathbf{d}^{(i)})^\mathsf{T}\mathbf{\Lambda}\mathbf{d}^{(i)}$. Using the fact that the spectral norm of a symmetric matrix is equal to its spectral radius, i.e., the absolute value of the largest eigenvalue $\lambda_{\max}$, we can bound $|\eta|$ as follows

$$|\eta| \leq |\lambda_1|\epsilon^2 + \|\mathbf{\Lambda}\|\|\mathbf{d}^{(i)}\|^2 \leq |\lambda_{\max}|\epsilon^2 + \|\mathbf{\Lambda}\|\epsilon^2 = 2\|\mathbf{\Lambda}\|\epsilon^2 \qquad (2.8)$$

and we see that $\mu^{(i)} = \lambda_1 - \eta = \lambda_1 + \mathcal{O}(\epsilon^2)$ (note that this proves Lemma 2.3). We obtain

$$\mathbf{y}^{(i+1)} = (\mathbf{\Lambda} - \mu^{(i)})^{-1}\mathbf{x}^{(i)}$$

$$= \left(\frac{x_1^{(i)}}{\lambda_1 - \mu^{(i)}}, \frac{x_2^{(i)}}{\lambda_2 - \mu^{(i)}}, \cdots, \frac{x_n^{(i)}}{\lambda_n - \mu^{(i)}}\right)^\mathsf{T} \qquad \left(\text{since } (\mathbf{\Lambda} - \mu^{(i)}\mathbf{I})^{-1} = \mathrm{diag}\left(\frac{1}{(\lambda_j - \mu^{(i)})}\right)\right)$$

$$= \left(\frac{1 + d_1^{(i)}}{\lambda_1 - \mu^{(i)}}, \frac{d_2^{(i)}}{\lambda_2 - \mu^{(i)}}, \cdots, \frac{d_n^{(i)}}{\lambda_n - \mu^{(i)}},\right)^\mathsf{T} \qquad (\text{since } \mathbf{x}^{(i)} = \mathbf{e}_1 + \mathbf{d}^{(i)})$$

$$= \left(\frac{1 - \epsilon^2/2}{\eta}, \frac{d_2^{(i)}}{\lambda_2 - \lambda_1 + \eta}, \cdots, \frac{d_n^{(i)}}{\lambda_n - \lambda_1 + \eta},\right)^\mathsf{T} \qquad (\text{since } \mu^{(i)} = \lambda_1 - \eta, \ d_1^{(i)} = -\epsilon^2/2)$$

$$= \frac{1 - \epsilon^2/2}{\eta}\left(1, \frac{d_2^{(i)}\eta}{(1 - \epsilon^2/2)(\lambda_2 - \lambda_1 + \eta)}, \cdots, \frac{d_n^{(i)}\eta}{(1 - \epsilon^2/2)(\lambda_n - \lambda_1 + \eta)},\right)^\mathsf{T}$$

$$=: \frac{1 - \epsilon^2/2}{\eta}\left(\mathbf{e}_1 + \hat{\mathbf{d}}^{(i+1)}\right).$$

If we denote by δ the gap between λ_1 and the rest of the spectrum, i.e.,

$$\delta := \min_{j=2}^{n}|\lambda_j - \lambda_1|,$$

we can bound the denominators of $\hat{\mathbf{d}}^{(i+1)}$ using $|\lambda_j - \lambda_1 + \eta| \geq \delta - |\eta|$, and in conjunction with (2.8) we get

$$\left\|\hat{\mathbf{d}}^{(i+1)}\right\| \leq \frac{\|\mathbf{d}^{(i)}\||\eta|}{(1 - \epsilon^2/2)(\delta - |\eta|)} \leq \frac{2\|\Lambda\|\epsilon^3}{(1 - \epsilon^2/2)(\delta - 2\|\Lambda\|\epsilon^2)}$$

or $\|\hat{\mathbf{d}}^{(i+1)}\| = \mathcal{O}(\epsilon^3)$. Finally, since $\mathbf{x}^{(i+1)} = \mathbf{e}_1 + \mathbf{d}^{(i+1)} = \mathbf{y}^{(i+1)}/\|\mathbf{y}^{(i+1)}\|$ and

$$
\frac{\mathbf{y}^{(i+1)}}{\|\mathbf{y}^{(i+1)}\|} = \frac{\frac{1-\epsilon^2/2}{\eta}\left(\mathbf{e}_1 + \hat{\mathbf{d}}^{(i+1)}\right)}{\left\|\frac{1-\epsilon^2/2}{\eta}\left(\mathbf{e}_1 + \hat{\mathbf{d}}^{(i+1)}\right)\right\|} = \frac{\mathbf{e}_1 + \hat{\mathbf{d}}^{(i+1)}}{\left\|\mathbf{e}_1 + \hat{\mathbf{d}}^{(i+1)}\right\|}
$$

we see that also $\mathbf{d}^{(i+1)} = \mathcal{O}(\epsilon^3)$, which concludes the proof. $\qquad\square$

We see that the property that the Rayleigh Quotient yields quadratically accurate approximations of eigenvalues is indeed crucial in the derivation of the cubic convergence. Some other proofs, for example the one given by Parlett [21, p. 77], proof convergence of the *error angle* $\phi^{(k)} = \angle(\mathbf{x}^{(k)}, \mathbf{v})$, i.e.,

$$
\lim_{k\to\infty}\left|\frac{\phi^{(k+1)}}{\left(\phi^{(k)}\right)^3}\right| \leq 1\,.
$$

In some texts that study inverse iteration in a more particularised fashion, cubic convergence is proofed by combining the quadratic accuracy of the Rayleigh Quotient together with convergence results from inverse iteration (since a single step of RQI is equivalent to a single step of inverse iteration with the Rayleigh quotient of the current iterate vector chosen as the shift), see for example [28, p. 208] or [3, pp. 89 sq.].

Global convergence behaviour

3 Complex Rayleigh Quotient Iteration

In this chapter we introduce *Complex Rayleigh Quotient Iteration* (CRQI[1]). This is a novel shift-and-invert algorithm similar to classic RQI. Numerical examples suggest that that this new method overcomes some of the disadvantages of classic RQI.

Finish intro for chapter

3.1 Motivation

As was the case in the previous chapters we fix a real symmetric matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$. Recall that since the eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of $\boldsymbol{A}$ form an orthonormal basis of $\mathbb{R}^n$ we can write every $\mathbf{u} \in \mathbb{R}^n$ as

$$\mathbf{u} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$$

for certain $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$. Suppose now, that $\mathbf{u}$ is a good approximation for one of the eigenvectors, say for $\mathbf{v}_k$. Then

$$\alpha_k \approx 1 \qquad \text{and} \qquad \alpha_j \approx 0 \ \text{ for } \ j \neq k \,.$$

Due to the pairwise orthogonality of the eigenvectors this implies

$$\mathbf{u}^\mathsf{T} \mathbf{v}_j \approx \begin{cases} 1 & \text{if } j = k \,, \\ 0 & \text{if } j \neq k \,. \end{cases} \tag{3.1}$$

---

[1]We abbreviate the classic Rayleigh Quotient Iteration that was discussed in the previous chapter by *RQI* or *classic RQI* and the method introduced in this chapter by *CRQI*.

As we have seen before, when using classic RQI, even good approximations of eigenvectors can lead to convergence to the wrong eigenpair when the gap between the target eigenvalue and eigenvalues nearby is very small. The main idea of CRQI is now to use the approximation $\mathbf{u}$ of $\mathbf{v}_k$ to perturb the linear system that is solved at each step in RQI in such a way that the distance between the target eigenvalue λ_k and the neighbouring eigenvalues is increased. In other words, we try to isolate the target eigenvalue. Of course this perturbed linear system will lead to wrong solutions and so we will "decrease" this perturbation successively until we arrive at the unperturbed problem.

We make use of the fact that all eigenvalues of $\boldsymbol{A}$ are real and perturb $\boldsymbol{A}$ in such a way that the eigenvalues are "raised" into the complex plane. Of course, we do not want to raise them all equally but rather in such a way that the Euclidean distance between the target eigenvalue and the other eigenvalue is increased. This is done by incorporating the approximation $\mathbf{u}$ of the target eigenvector $\mathbf{v}_k$ into the working matrix. Let

$$\tilde{\boldsymbol{A}} \coloneqq \boldsymbol{A} - \gamma i (\boldsymbol{I} - \mathbf{u}\mathbf{u}^\mathsf{T}), \tag{3.2}$$

where $\gamma > 0$ is positive real number and i denotes the imaginary unit. Note that the matrix $\boldsymbol{I} - \mathbf{u}\mathbf{u}^\mathsf{T}$ defines the orthogonal projection onto the span of $\mathbf{u}$. Therefore, a vector $\mathbf{x}$ that is almost parallel to $\mathbf{u}$ will barely "see" the imaginary part $\gamma i (\boldsymbol{I} - \mathbf{u}\mathbf{u}^\mathsf{T})$ when multiplied with $\tilde{\boldsymbol{A}}$ and so $\tilde{\boldsymbol{A}}\mathbf{x} \approx \boldsymbol{A}\mathbf{x}$. If, however, the vector $\mathbf{x}$ is almost perpendicular to $\mathbf{u}$ we have

$$\tilde{\boldsymbol{A}}\mathbf{x} = \boldsymbol{A}\mathbf{x} - \gamma i \mathbf{x} - \mathbf{u}\underbrace{\mathbf{u}^\mathsf{T}\mathbf{x}}_{\approx 0} \approx (\boldsymbol{A} - \gamma i \boldsymbol{I})\mathbf{x}. \tag{3.3}$$

Since u approximates $\mathbf{v}_k$, the orthogonal complement of the span of $\mathbf{u}$ approximates the orthogonal complement of the span by $\mathbf{v}_k$. The latter, due to the orthogonality of the eigenvectors, is the subspace spanned by the remaining eigenvectors. Therefore, we expect that the eigenvectors of $\tilde{\boldsymbol{A}}$ are similar to those of $\boldsymbol{A}$ and that the eigenvalues corresponding to eigenvectors $\mathbf{v}_j$, $j \neq k$

to approximately be $\lambda_j - \gamma i$ due to (3.3). The eigenvalue corresponding to $\mathbf{v}_k$ would then be approximately equal to λ_k since $\tilde{\boldsymbol{A}}\mathbf{v}_k \approx \boldsymbol{A}\mathbf{v}_k = \lambda_k\mathbf{v}_k$.

If we would use this matrix in RQI, the results would of course not be the target eigenpair but an eigenpair of $\tilde{\boldsymbol{A}}$. Thus, instead of keeping this matrix fixed, we replace the vector $\mathbf{u}$ by the current iterate $\mathbf{x}^{(k)}$ and the scalar γ by a sequence $\gamma^{(k)}$ that converges to zero. Ideally, in the beginning of the iteration $\gamma^{(k)}$ should be sufficiently large such that the target eigenvalue is properly isolated. As the the iterates get closer to the target eigenpair, $\gamma^{(k)}$ should decrease such that in the end $\tilde{\boldsymbol{A}} \approx \boldsymbol{A}$. An obvious choice is the norm of the current residual $\mathbf{r}^{(k)}$ or related quantities such as the square of the residual norm. How the choice of this shift influences the convergence behaviour is discussed in the next first example of Section 3.3.

3.2 Implementation

Altough we will soon slightly change the matrix $\tilde{\boldsymbol{A}}$ defined in (3.2) again, we summarise the discussion up to this point in Algorithm 3.1.[2]

Strictly speaking, this algorithm is not a Rayleigh Quotient Iteration as the matrix changes at each step. Obviously, the method is strongly linked to RQI so we still refer to it as Complex Rayleigh Quotient Iteration. As mentioned above, this is not the final version as we wish to present it. The following lemma allows for a small simplification of the algorithm.

Lemma 3.2. *Let* $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ *be a nonsingular real symmetric matrix. Let* $\mathbf{u} \in \mathbb{C}^n$ *be a unit vector. We define the matrices*

$$\boldsymbol{B} := \boldsymbol{A} - (\sigma - \gamma i)\boldsymbol{I} \tag{3.4}$$

and

$$\boldsymbol{C} := \boldsymbol{A} - \sigma\boldsymbol{I} + \gamma i(\boldsymbol{I} - \mathbf{u}\mathbf{u}^*)\,, \tag{3.5}$$

> Sign of shift in Algo step should be +?

---

[2]By $\Re(\mathbf{x})$ we denote the real part of the complex vector $\mathbf{x}$.

Algorithm 3.1: Complex Rayleigh Quotient Iteration, First Version

---

Input: Nonzero initial vector $\mathbf{x}^{(0)}$ with $\|\mathbf{x}^{(0)}\| = 1$

begin

    $\mu^{(0)} \leftarrow (\mathbf{x}^{(0)})^* \boldsymbol{A} \mathbf{x}^{(0)}$

    $r^{(0)} \leftarrow \|(\boldsymbol{A} - \mu^{(0)}\boldsymbol{I})\mathbf{x}^{(0)}\|$

    $\gamma^{(0)} \leftarrow r^{(0)}$

    for $k = 1, 2, \ldots$ *until convergence* **do**

        $\tilde{\boldsymbol{A}} \leftarrow \boldsymbol{A} - \gamma^{(k)} i (\boldsymbol{I} - \mathbf{x}^{(k)}(\mathbf{x}^{(k)})^*)$

        Solve $(\tilde{\boldsymbol{A}} - \mu^{(k)}\boldsymbol{I})\tilde{\mathbf{x}}^{(k+1)} = \mathbf{x}^{(k)}$

        $\mathbf{x}^{(k+1)} \leftarrow \tilde{\mathbf{x}}^{(k+1)}/\|\tilde{\mathbf{x}}^{(k+1)}\|$

        $\mu^{(k+1)} \leftarrow (\mathbf{x}^{(k+1)})^* \boldsymbol{A} \mathbf{x}^{(k+1)}$

        $r^{(k+1)} \leftarrow \|(\boldsymbol{A} - \mu^{(k+1)}\boldsymbol{I})\mathbf{x}^{(k+1)}\|$

        $\gamma^{(k+1)} \leftarrow r^{(k+1)}$

    $\mathbf{x} \leftarrow \Re(\mathbf{x}^{(k+1)})$

    $\mathbf{x} \leftarrow \mathbf{x}/\|\mathbf{x}\|$

---

where i is the imaginary unit and $\sigma, \gamma > 0$ are positive real numbers such that σ is not an eigenvalue of $\boldsymbol{A}$. Then

$$\mathcal{R}_{\boldsymbol{A}}(\boldsymbol{B}^{-1}\mathbf{u}) = \mathcal{R}_{\boldsymbol{A}}(\boldsymbol{C}^{-1}\mathbf{u}). \tag{3.6}$$

Proof. Without loss of generality, we can assume $\sigma = 0$. Otherwise, set $\tilde{\boldsymbol{A}} = \boldsymbol{A} - \sigma\boldsymbol{I}$ and use this matrix instead of $\boldsymbol{A}$ ($\tilde{\boldsymbol{A}}$ is obviously still real and symmetric and invertible since σ is not an eigenvalue of $\boldsymbol{A}$). First, observe that $\boldsymbol{C} = \boldsymbol{B} - \gamma i\mathbf{u}\mathbf{u}^*$. Using the Sherman-Morrison formula [8] and letting $\alpha \coloneqq 1 - \gamma i\mathbf{u}^*\boldsymbol{B}^{-1}\mathbf{u} \in \mathbb{C}$

we obtain

$$
\begin{aligned}
\boldsymbol{C}^{-1}\mathbf{u} &= (\boldsymbol{B} - \gamma i\mathbf{u}\mathbf{u}^*)^{-1}\mathbf{u} \\
&= \left(\boldsymbol{B}^{-1} + \alpha^{-1}\boldsymbol{B}^{-1}\mathbf{u}\gamma i\mathbf{u}^*\boldsymbol{B}^{-1}\right)\mathbf{u} \\
&= \boldsymbol{B}^{-1}\mathbf{u} + \boldsymbol{B}^{-1}\mathbf{u}\underbrace{\alpha^{-1}\gamma i\mathbf{u}^*\boldsymbol{B}^{-1}\mathbf{u}}_{\in\mathbb{C}} \\
&= \boldsymbol{B}^{-1}\mathbf{u}(1 + \alpha^{-1}\gamma i\mathbf{u}^*\boldsymbol{B}^{-1}\mathbf{u}).
\end{aligned}
$$

Thus, the vector $\boldsymbol{C}^{-1}\mathbf{u}$ is a scalar multple of $\boldsymbol{B}^{-1}\mathbf{u}$ and the result follows from Lemma 2.2 (i) (the Homogenity of the Rayleigh Quotient). □

Immediately, the following result follows.

Corollary 3.3. *Algorithm 3.1 produces the same results when $\tilde{\boldsymbol{A}}$ is replaced with $\hat{\boldsymbol{A}} = \boldsymbol{A} + \gamma i\boldsymbol{I}$.*

Proof. □

This result allows for a simplification of the algorithm. We combine the real shift $\mu^{(k)}$ and the imaginary shift $i\gamma^{(k)}$ into a new shift $\sigma^{(k)} := \mu^{(k)} - i\gamma^{(k)}$ and solve the system

$$
(\boldsymbol{A} - \sigma^{(k)}\boldsymbol{I})\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}
$$

for $\mathbf{x}^{(k+1)}$ at each step.

3.3 Numerical Experiments

In this section we discuss different numerical examples to better understand the behaviour of CRQI. Throughout the section we will always compare the method to classic RQI. All experiments were executed in Matlab 9. The criterion for convergence was $\|\mathbf{r}^{(k)}\| = \|\boldsymbol{A}\mathbf{x}^{(k)} - \mu^{(k)}\mathbf{x}^{(k)}\| < 10^{-8}$. The source codes for both classic and complex RQI can be found in the Appendix. The methods are defined slightly differently from the Algorithms given above. They both

[margin note: Where is the non-singularity needed?]

[margin note: Cite matlab]

allow to explicitly set the initial shifts $\mu^{(0)}$ and $\gamma^{(0)}$ to specific values whereas in Algorithm 2.4 and Algorithm 3.1 the initial shifts are always initialised as the Rayleigh Quotient of the initial vector and the initial residual, respectively. In most of the examples below the matrix $\boldsymbol{A}$ was a randomly generated sparse symmetric 200×200 matrix. To simulate that a good approximation of the initial vector is available we computed the full set of eigenvectors collected as columns of the matrix $\boldsymbol{V}$. Next, a weight vector $\mathbf{w}$ of uniformly distributed numbers between 0 and 1 is created. One of the components is set to a higher value than the others, e. g. $w_{50} = 10$ (in most examples the index of the component was also chosen randomly). The initial vector is then set to a weighted linear combination of the eigenvectors, i. e., $\mathbf{x}^{(0)} = \beta \boldsymbol{V} \mathbf{w}$, where β has to be chosen such that $\mathbf{x}^{(0)}$ is normalised. Now, $\mathbf{x}^{(0)}$ is a vector with a strong component in the direction of the target eigenvector and random (smaller) contributions in the directions of the other eigenvectors.

The first example demonstrates the influence of the imaginary shift $\gamma^{(k)}i$ on the convergence speed. We started by running two versions of CRQI. The first uses $\gamma^{(k)} = \|\mathbf{r}^{(k)}\|$, the second uses the square of the residual norm. A plot of the residuals against the iteration number of both these approaches together with the results using classic RQI[3] and a third approach described shortly is given in Figure 3.1.

We observe that in the initial phase of the iteration, the second variant seems to be slower than the first. Altough it is not clearly observable in the figure below, other examples suggested that during the final steps of the iterations the second version was faster than the first. Consequently, by combining both approaches and thus changing the shift adaptively we expect faster convergence. The results are also plotted in Figure 3.1 and are in accordance with the expectation. In

---

[3]Classic RQI is merely included for speed comparison. Since the eigenvalues of the test problem are closely spaced, in almost all of the examples classic RQI failed to converge to the right eigenvalue.
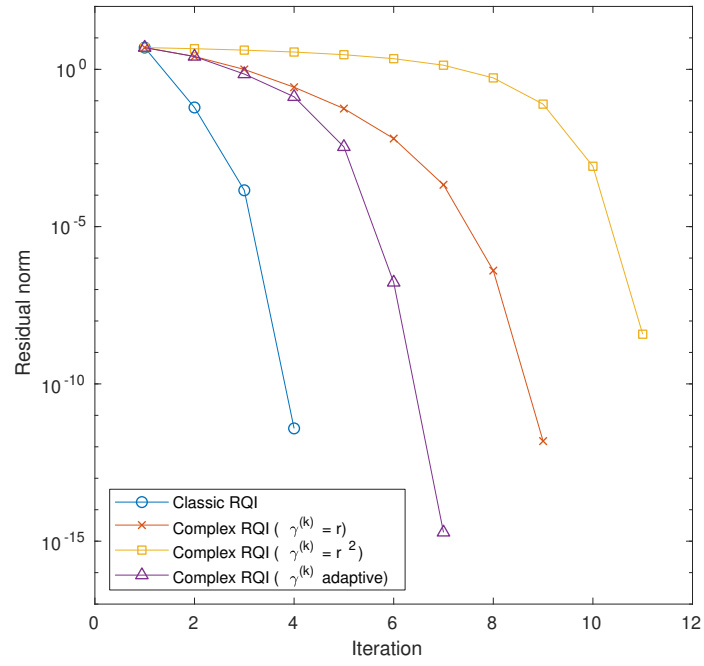
Figure 3.1: Plot of residuals of Classic RQI, Complex RQI with $\gamma^{(k)} = \|\mathbf{r}^{(k)}\|$, Complex RQI with $\gamma^{(k)} = \|\mathbf{r}^{(k)}\|^2$ and Complex RQI with the imaginary shift chosen adaptively (see text). The last variantt seems to combine the adavantages of the second and third alternatives.

this particular case, the shift was changed adaptively according to

$$\gamma = \begin{cases} \|\mathbf{r}\| & \text{if } \|\mathbf{r}\| \geq 1 \,, \\ \|\mathbf{r}\|^2 & \text{if } \|\mathbf{r}\| < 1 \,, \end{cases}$$

where we dropped the superscripts from $\gamma = \gamma^{(k)}$ and $\mathbf{r} = \mathbf{r}^{(k)}$. In the following, when we speak of CRQI we mean CRQI performed with this adaptive choice of imaginary shifts.

Running the same experiment but increasing the component of the initial vector in the direction of the target eigenvector it showed that the number of additional steps required by CRQI diminshed. For eigenvectors that were very close to the target sometimes both classic RQI and complex RQI finished within the same number of iterations and while CRQI generally produced the correct eigenpair, classic RQI usually failed to do so.

Further investigation revealed that the behaviour as in Figure 3.2(b) and Figure 3.2(a)

(a) Small component in target direction, negative eigenvalue

(b) Small component in target direction, positive eigenvalue

(c) Sufficiently big component in target direction
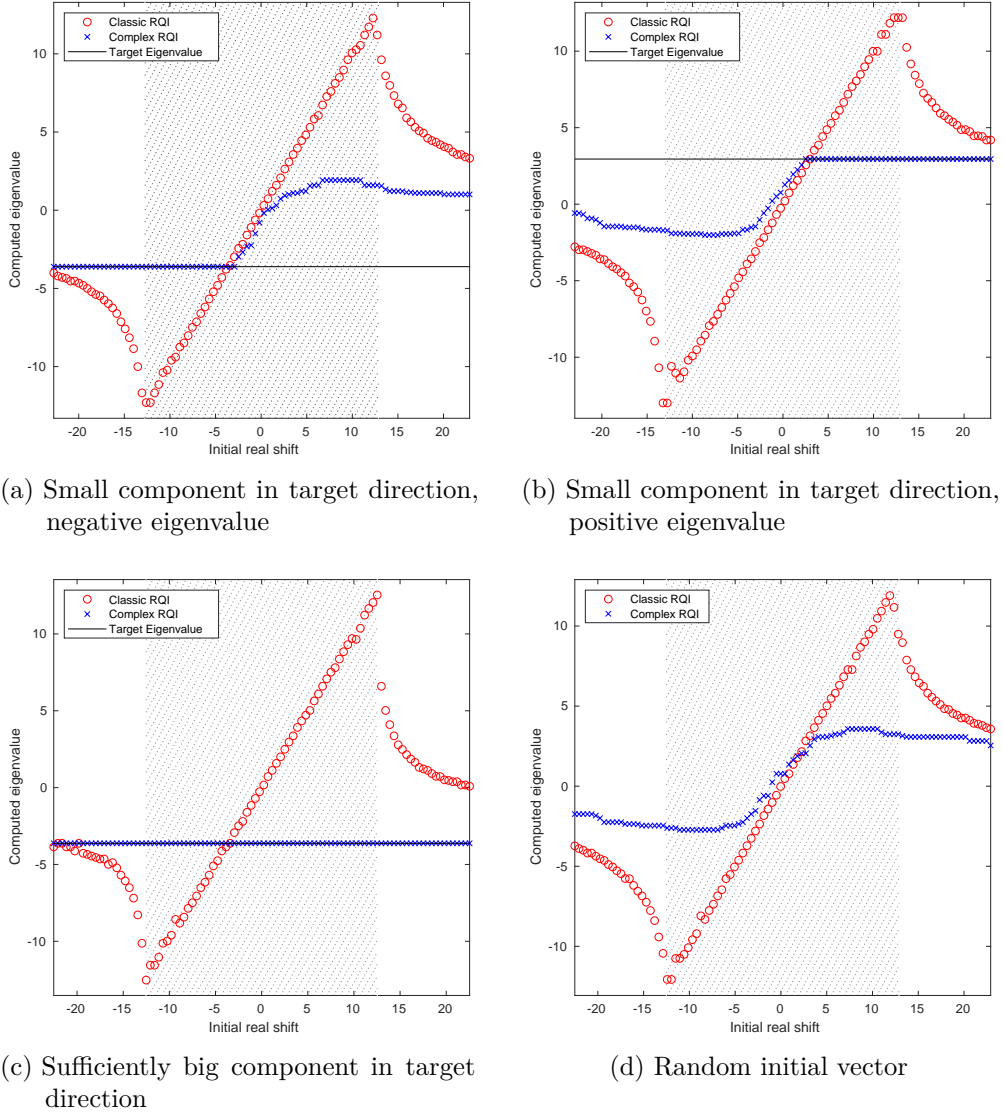
(d) Random initial vector

Figure 3.2: Plot of initial real shift against the computed eigenvalue using classic RQI and CRQI. The dotted area encloses the initial shifts that lie in the spectrum of A. In the first two examples the initial vector had only a small component in the direction of the target eigenvector.

Appendices

A Proofs

Bibliography

[1]   Steve Batterson and John Smillie. "The Dynamics of Rayleigh Quotient Iteration". In: *SIAM Journal on Numerical Analysis* 25.3 (1989), pp. 624–636 (cit. on p. 25).

[2]   Christopher Beattie and David W. Fox. "Localization Criteria and Containment for Rayleigh Quotient Iteration". In: *SIAM Journal on Matrix Analysis and Applications* 10.1 (1989), pp. 80–93 (cit. on p. 24).

[3]   Steffen Börm. *Numerik von Eigenwertaufgaben*. Lecture notes. 2018 (Last accessed: 14 March 2020). URL: `https://www.informatik.uni-kiel.de/~sb/data/Eigenwerte.pdf` (cit. on pp. 18, 29).

[4]   Stephen H. Crandall and Richard Vynne Southwell. "Iterative procedures related to relaxation methods for eigenvalue problems". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 207.1090 (1951), pp. 416–423 (cit. on p. 20).

[5]   James W. Demmel. *Applied Numerical Linear Algebra*. Philadelphia: Society for Industrial and Applied Mathematics, 1997 (cit. on p. 26).

[6]   Stephen H. Friedberg, Arnold J. Insel, and Lawrence E. Spence. *Linear Algebra: Pearson New International Edition*. Fourth Edition. Pearson Education, 2014 (cit. on p. 17).

[7]   Gene H. Golub and Henk A. Van der Vorst. "Eigenvalue computation in the 20th century". In: *Journal of Computational and Applied Mathematics* 123.1-2 (2000), pp. 35–65 (cit. on pp. 6, 14).

[8]   William W. Hager. "Updating the Inverse of a Matrix". In: *SIAM Review* 31.2 (1989), pp. 221–239 (cit. on p. 33).

[9]   Kenneth M. Hoffman and Ray A. Kunze. *Linear Algebra*. Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1971 (cit. on p. 7).

[10] Ilse C. F. Ipsen. "A History Of Inverse Iteration". In: *Helmut Wielandt, Mathematische Werke, Mathematical Works, Volume 2: Linear Algebra and Analysis.* Ed. by Bertram Huppert and Hans Schneider. Berlin, New York: Walter de Gruyter, 1996, pp. 464–472 (cit. on p. 14).

[11] Carl Gustav Jacob Jacobi. "Über ein leichtes Verfahren die in der Theorie der Säcularstörungen vorkommenden Gleichungen numerisch aufzulösen". In: *Journal für die reine und angewandte Mathematik (Crelles Journal)* 30 (1846), pp. 51–94 (cit. on p. 6).

[12] Walter Kohn. "A Variational Iteration Method for Solving Secular Equations". In: *The Journal of Chemical Physics* 17.7 (1949), p. 670 (cit. on p. 20).

[13] Alexander M. Ostrowski. "On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. I". In: *Archive for Rational Mechanics and Analysis* 1.1 (1957), pp. 233–241 (cit. on pp. 21, 22, 24, 25).

[14] Alexander M. Ostrowski. "On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. II". In: *Archive for Rational Mechanics and Analysis* 2.1 (1958), pp. 423–428 (cit. on pp. 21, 22).

[15] Alexander M. Ostrowski. "On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. III". In: *Archive for Rational Mechanics and Analysis* 3.1 (1959), pp. 325–340 (cit. on pp. 21, 22).

[16] Alexander M. Ostrowski. "On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. IV". In: *Archive for Rational Mechanics and Analysis* 3.1 (1959), pp. 341–347 (cit. on p. 21).

[17]   Alexander M. Ostrowski. "On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. V". In: *Archive for Rational Mechanics and Analysis* 3.1 (1959), pp. 472–481 (cit. on p. 21).

[18]   Alexander M. Ostrowski. "On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. VI". In: *Archive for Rational Mechanics and Analysis* 4.1 (1959), pp. 153–165 (cit. on p. 21).

[19]   Ricardo D. Pantazis and Daniel B. Szyld. "Regions of convergence of the Rayleigh quotient iteration method". In: *Numerical Linear Algebra with Applications* 2.3 (1995), pp. 251–269 (cit. on pp. 23, 25).

[20]   Beresford N. Parlett. "The Rayleigh Quotient Iteration and Some Generalizations for Nonnormal Matrices". In: *Mathematics of Computation* 28.127 (1974), pp. 679–693 (cit. on p. 23).

[21]   Beresford N. Parlett. *The Symmetric Eigenvalue Problem*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1998 (cit. on pp. 13, 17, 23, 25, 29).

[22]   Beresford N. Parlett and William Kahan. "On the convergence of a practical QR algorithm". In: *Information Processing, Proceedings of IFIP Congress 1968, Edinburgh, UK, 5-10 August 1968, Volume 1 - Mathematics, Software*. Ed. by A. J. H. Morrel. 1968, pp. 114–118 (cit. on p. 23).

[23]   G Peters and James H. Wilkinson. "Inverse Iteration, Ill-Conditioned Equations and Newton's Method". In: *SIAM Review* 21.3 (1979), pp. 339–360 (cit. on p. 13).

[24]   John William Strutt, Baron Rayleigh. *The Theory of Sound*. 2nd ed. Vol. 1. London and New York: MacMillan and Co, 1894 (cit. on p. 20).

[25]   Yousef Saad. *Numerical Methods for Large Eigenvalue Problems*. Society for Industrial and Applied Mathematics, 2011 (cit. on pp. 11, 18, 19).

[26] Daniel B. Szyld. "Criteria for Combining Inverse and Rayleigh Quotient Iteration". In: *SIAM Journal on Numerical Analysis* 25.6 (1988), pp. 1369–1375 (cit. on pp. 24, 25).

[27] Richard A. Tapia, J. E. Dennis, and Jan P. Schäfermeyer. "Inverse, Shifted Inverse, and Rayleigh Quotient Iteration as Newtons Method". In: *SIAM Review* 60.1 (2018), pp. 3–55 (cit. on pp. 14, 19, 21).

[28] Lloyd N. Trefethen and David Bau. *Numerical linear algebra.* Other titles in applied mathematics. Society for Industrial and Applied Mathematics, 1997 (cit. on pp. 6, 17, 29).

[29] Helmut Wielandt. *Beiträge zur mathematischen Behandlung komplexer Eigenwertprobleme, Teil V: Bestimmung höherer Eigenwerte durch gebrochene Iteration.* Bericht B 44/J/37. Aerodynamische Versuchsanstalt Göttingen, 1944 (cit. on p. 21).

[30] J. H. Wilkinson. *The Algebraic Eigenvalue Problem.* Ed. by E. T. Goodwin and L Fox. Monographs On Numerical Analysis. Oxford: Clarendon Press, 1965 (cit. on p. 13).