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

Using Complex Shifts in Rayleigh Quotient Iteration to Compute Interior Eigenvalues

Bachelor Thesis

Author: Nils Friess

Supervisor: Prof. Dr. Robert Scheichl

Submitted: April 28, 2020

Abstract This thesis presents a novel iterative method for computing interior eigenvalues of real symmetric matrices.

Zusammenfassung In dieser Arbeit wird eine neue iterative Methode zur Berechnung innerer Eigenwerte reller symmetrischer Matrizen vorgestellt.

Contents

Todo list

1 Introduction

{chapter:intro}

This thesis introduces a novel shift-and-invert-type method for computing eigenvalues and eigenvectors of real symmetric matrices that can be seen as a modification of *Rayleigh Quotient Iteration (RQI)*. Numerical examples suggest that our method overcomes some of the drawbacks of RQI at the expense of slower convergence.

The thesis is structured as follows. In this chapter, we discuss a model problem from physics before we collect all the necesarry definitions and results from numerical linear algebra that will be needed later. We finish this introductory chapter by reviewing some of the most basic iterative methods for computing eigenvectors.

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 it so popular, namely the local cubic convergence.

In the third chapter we introduce our novel method that we call *Complex Rayleigh Quotient Iteration*. After motivating the method we carefully study different numerical examples to better understand the behaviour of the method.

1.1 Motivation

The following example introduces a physical problem that requires the solution of an eigenvalue problem. There is a plethora of problems which lead to eigenvalue problems in almost all of the natural sciences but also in finance and other social sciences (see, for example,).

References for eigenvalue problems

{ex:modelproblem}

Example 1.1 (Photonic Crystal Fibres)**.**

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

{sec:symmetric:eige

In Example 1.1 a problem of the form

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

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

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

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 $\sigma(\boldsymbol{A})$.

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 holds 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, altough being of theoretical significance, cannot be used to calculate eigenvalues numerically for two reasons. First, the computation of the

coefficients of the polynomial is not stable [4, 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. 20, 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 [7]. 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*

$$\boldsymbol{A}\mathbf{v} = \lambda \boldsymbol{M}\mathbf{v}\,. \tag{1.2}$$ {eq:eigvalproblem:g

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. Citation needed.

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 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., $\boldsymbol{A} = \boldsymbol{A}^* \coloneqq \overline{\boldsymbol{A}}^{\mathsf{T}}$, 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.

{prop:eigval:facts}

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 [5, 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

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

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

[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.

(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 $\boldsymbol{A} - \mu\boldsymbol{I}$.

□

In the following we restrict our attention to the *symmetric eigenvalue problem*, i. e., we want to find solutions of equation (1.1) given $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is a real symmetric matrix. The (real) eigenvalues and associated (real) eigenvectors of $\boldsymbol{A}$ are denoted by $\lambda_1, \ldots, \lambda_n$ and $\mathbf{v}_1, \ldots, \mathbf{v}_n$, respectively. Since for any eigenvector $\mathbf{v}$ also any scalar multiple of $\mathbf{v}$ is also an eigenvector, we assume that the eigenvectors are normalised w. r. t. the Euclidean norm, so that

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

Note, that 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. To indicate that eigenvalues belong to a particular matrix $\boldsymbol{M}$ we sometimes write $\lambda(\boldsymbol{M})$ or $\lambda_j(\boldsymbol{M})$. If not stated otherwise, for the remainder of the thesis $\boldsymbol{A}$ denotes a symmetric $n \times n$ matrix with eigenvalues $\lambda_j(\boldsymbol{A}) = \lambda_j$ and corresponding eigenvectors $\mathbf{v}_j$ for $j = 1, \ldots, n$.

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 neighbors are very close. These eigenvalues will always be of interior type and we do usually not have any a priori knowledge about their location. Also, the matrices are assumed to be large such that 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 fail in this case when the gap between the wanted eigenvalue and adjacent eigenvalues is too small.

1.2.2 Iterative methods for eigenvalue problems

{sec:iterative:algo

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 defintion 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 $\mathbb{R}^n$ that converges to $\mathbf{z} \in \mathbb{R}^n$.

(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 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 identites are convenient

$$\begin{aligned} \cos \angle(\mathbf{x}, \mathbf{y}) &= \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|}{\|\mathbf{x}\| \|\mathbf{y}\|} , \\ \sin \angle(\mathbf{x}, \mathbf{y}) &\coloneqq \sqrt{1 - \cos^2 \angle(\mathbf{x}, \mathbf{y})} , \\ \tan \angle(\mathbf{x}, \mathbf{y}) &\coloneqq \frac{\sin \angle(\mathbf{x}, \mathbf{y})}{\cos \angle(\mathbf{x}, \mathbf{y})} . \end{aligned}$$

Power method

The *power method* is probably the simplest iterative method for computing eigenvectors. It is based on generating the sequence $\mathbf{x}^{(k)} \coloneqq \boldsymbol{A}^k \mathbf{x}^{(0)}$ 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 under- and overflow errors $\mathbf{x}^{(k)}$ is normalised at each step. Altough the vectors are usually normalised w. r. t. the Euclidean norm, in Algorithm 1.7 we normalise by ensuring that the largest component of the current approximation is equal to one (i. e., we normalise w. r. t. to the maximum norm) as proposed in [18]. 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 choosen, the method converges linearly with convergence rate

$$\rho = \frac{|\lambda_{n-1}|}{|\lambda_n|}\,. \tag{1.3}$$ {eq:convergence:pow

Thus, the method can be very slow if the distance between the eigenvalues λ_n and λ_{n-1} is very small. For more details, see [18, pp. 85 sqq.].

Algorithm 1.7: Power method

begin
  Choose nonzero initial vector $\mathbf{x}^{(0)}$
  for $k = 1, 2, \ldots$ *until convergence* **do**
    $\mathbf{x}^{(k)} = \frac{1}{\alpha^{(k)}} \boldsymbol{A}\mathbf{x}^{(k-1)}$
    /* $\alpha^{(k)}$ is the component of $\boldsymbol{A}\mathbf{x}^{(k-1)}$ with the maximum modulus */

{alg:power:method}

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.

(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 ensures that $\mathbf{x}^{(k)}$ is normalised. 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 decomposition, if applicable) and solve the system $(\boldsymbol{A} - \sigma\boldsymbol{I})\mathbf{x}^{(k)} = \mathbf{x}^{(k-1)}$ for $\mathbf{x}^{(k)}$, reducing the complexity from $\mathcal{O}(n^3)$ to $\mathcal{O}(n^2)$ at each step. We summarise the results in Algorithm 1.8 (there, we normalise w. r. t. the Euclidean norm).

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

{alg:sii}

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 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. Clearly, a shift that 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. In practice — contrary to the expectations — the method has shown to produce good approximations even in these cases and it took several years from when inverse iteration was introduced in the 1940s until Wilkinson established that this is merely a problem in exact arithmetic. He elucidated that this poses no problem when the linear system is solved numerically using a backwards stably method, see [22, pp. 621–630] and [16]. 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 again this poses no problem in practice.

Note that we did not specify the "until convergence" criteria in neither of the algorithms above. We postpone this discussion until Section 2.1.

Example 1.9. Consider the matrix

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 6 & 8 \end{pmatrix}.$$

Example of Power method and inverse iteration

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. [4, 6, 19].

2 Classic Rayleigh Quotient Iteration

{chapter:rqi}

In this chapter we thoroughly introduce the 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 disadvantages and shortcomings.

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

{sec:rqi:rq}

In Chapter 1 we briefly introduced iterative methods for computing eigenpairs. In essence, RQI is shifted inverse iteration where the shift is replaced by the *Rayleigh quotient* at each step. As we will see, the Rayleigh Quotient can be used to obtain good estimates of an eigenvalue given an approximation of an eigenvector.

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

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

[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 $\rho_{\boldsymbol{A}}(\mathbf{x})$.

We begin by discussing some basic facts.

{lem:rq:properties}

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

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

{eq:rq:innerprod}

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)

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

□

Altough 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 $\boldsymbol{A}$, then

$$\|\boldsymbol{A}\mathbf{v} - \lambda\mathbf{v}\|^2 = 0\,.$$

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

$$\hat{\lambda} = \arg\min_{\mu\in\mathbb{C}} \|\boldsymbol{A}\hat{\mathbf{v}} - \mu\hat{\mathbf{v}}\|^2\,.$$

This is a linear least squares problem in μ with normal equations (cf. [3, p. 106])

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

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

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

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" [15, p. 77].

{lem:rq:quadestimat

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*

$$|\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}) - \lambda| \leq \|\boldsymbol{A} - \lambda\boldsymbol{I}\| \sin^2\angle(\mathbf{x},\mathbf{v}) \leq \|\boldsymbol{A} - \lambda\boldsymbol{I}\| \left(\frac{\|\mathbf{x} - \alpha\mathbf{v}\|}{\|\mathbf{x}\|}\right)^2 \quad \text{for all } \alpha \in \mathbb{C}\,.$$

The proof is given in Appendix **??**. □

This result is often paraphrased as "the Rayleigh Quotient is a *quadratically accurate* estimate of an eigenvalue" (see for example [20, p. 204]) and is often

[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.

presented in the following shorter form cite examples

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

If $\boldsymbol{A}$ is non-normal, the Rayleigh quotient is still an estimate of order one, i. e.,

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

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

{alg:rqi}

We have yet to define what we mean by "until convergence" (cf. the algorithms persented in Section 1.2.2). 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 `tol` is a user-given error tolerance. 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. A rigorous justfication is given in [18, pp. 59 sqq.]. Here, we give only some important results without proof for the case when $\boldsymbol{A}$ is symmetric. A popular result, usually referred to as the *Bauer-Fike theorem* (see, e. g. [18, p. 59]) states that there exists an eigenvalue λ of $\boldsymbol{A}$ such that

Rewrite

$$\left|\lambda - \mu^{(k)}\right| \leq \kappa(\boldsymbol{Q})\left\|\mathbf{r}^{(k)}\right\|, \tag{2.2}$$

{eq:bauer-fike}

where $\boldsymbol{Q}$ is a matrix that transforms $\boldsymbol{A}$ into diagonal form (i. e., $\boldsymbol{Q}$ is such that $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$ is diagonal) and $\kappa(\boldsymbol{Q}) \coloneqq \|\boldsymbol{Q}\|\left\|\boldsymbol{Q}^{(-1)}\right\|$ is the *condition number* of $\boldsymbol{Q}$. Such a matrix exists due to the *spectral theorem* [5, Theorem 18 and Corollary, p. 314]. Since $\boldsymbol{A}$ is real and symmetric $\boldsymbol{Q}$ is an *orthogonal* matrix. In particular $\|\boldsymbol{Q}\| = \|\boldsymbol{Q}^{-1}\| = 1$ and thus

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

Therefore, if the stopping criterion is fulfilled, we have $|\lambda - \mu^{(k)}| < \texttt{tol}$. For the eigenvector one can show [18, 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 associtated with λ and δ is the distance from $\mu^{(k)}$ to the rest of the spectrum, i. e.,

$$\delta \coloneqq \min_i \left\{|\mu^{(k)} - \lambda_i| \ : \ \lambda_i \neq \lambda\right\}.$$

These results hold if $\mu^{(k)} = (\mathbf{x}^{(k)})^* \boldsymbol{A}(\mathbf{x}^{(k)})$ (Rayleigh Quotient).

Stopping criteria in Power method and SIT

2.2 History and recent developments

Now that we have defined the Rayleigh Quotient and Rayleigh Quotient Iteration we give an overview on 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 on some important milestones within this 60 years. Large parts of this overview are based on [19].

Page

1894 — Lord Rayleigh In the second edition of his book titled "The Theory of Sound" John William Strutt, 3rd Baron Rayleigh [17, p. 110], proposed the following iteration for imporoving an approximate eigenvector $\mathbf{x}^{(0)}$:

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

{eq:rayleigh:iterat

where $\mathbf{e}_1$ denotes the first natural coordinate vector, i. e., the first column of the $n \times n$ identity matrix and $\mathbf{x}^{(i)}$ and $\mathbf{x}^{(i+1)}$ denote the current and next iterate, respectively. To be precise, Lord Rayleigh considered the *generalised eigenvalue problem*

$$\boldsymbol{A}\mathbf{v} = \lambda \boldsymbol{M}\mathbf{v} \tag{2.4}$$

{eq:general:eig}

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 [8] 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 $\boldsymbol{A}$ (given 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 Crandall [2] suggests

Full name

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

{eq:unnormalised:rq

Actually, Crandall also considered the generalised eigenproblem (2.4) but for our comparitive 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 (wrong) assumption that the sequence of vectors $\mathbf{x}^{(k)}$ converges, Crandall establishes cubic convergence for this sequence. To see why the assumption is wrong, we assume the contrary, i. e., suppose $\mathbf{x}^{(i)} \longrightarrow \mathbf{v}_k$ for some $k = 1, \ldots, n$. From (2.5) we have

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

i. e., $\mathbf{v}_k$ 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.

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" [9–14]. 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

$$\text{Solve} \quad (\boldsymbol{A} - \mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(i)})\boldsymbol{I})\mathbf{x}^{(i+1)} = \boldsymbol{\eta}\,, \quad \boldsymbol{\eta} \neq \mathbf{0}\,. \tag{2.6}$$

{eq:rqi:ostrowski:f

He rigorously establishes a *quadratic* convergence rate for the sequence $\{\mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(i)})\}$. He then refers to a paper of Wielandt [21] and his *fractional* or *broken iteration* (German: *gebrochene Iteration*). Inspired by Wielandt's method he suggests replacing the fixed vector **eta** in the right hand side of (2.6) to the solution from the previous step

$$(\boldsymbol{A} - \mathcal{R}_{\boldsymbol{A}}(\mathbf{x}^{(i)})\boldsymbol{I})\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)}\,, \tag{2.7}$$

{eq:rqi:unnormalise

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.,

$$\frac{\mu_{i+1} - \lambda}{(\mu_i - \lambda)^3} \longrightarrow \gamma \quad \text{as } i \to \infty\,, \tag{2.8}$$

{eq:rqi:cubic}

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 (2.7) is the same algorithm previously proposed by Crandall given in Equation (2.5). 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." [9, p. 241].[3]

In the beginning of the second paper [10] Ostrowski discusses this in more detail. 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. This was all shown under the assumption that the matrix $\boldsymbol{A}$ is real symmetric and Ostrowski mentions that all results remain valid in the complex Hermitian case.

The third paper [11] of the series addresses the non-symmetric (or non-Hermitian) case for which Ostrowski is also able to define a method that attains local cubic convergence. 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.

Read Ostrowski's other papers

This concludes the overview on the development of RQI from the first mention of the Rayleigh Quotient by Lord Rayleigh to the first definition of RQI and the first rigorous proof of what makes it so appealing, namely the (local) cubic convergence.

2.2.2 Further developments and recent contributions

2.3 Convergence Analysis

{sec:convergence}

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.

As we have seen in the previous section, the first rigorous proof of the cubic convergence of RQI was given by Ostrowski [9] but over time a number of

[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 (2.7) and the asymptotic formula (46) he references is given in (2.8).

different, simpler proofs were presented. The one we give below follows closely the one given by Demmel [3, pp. 215 sq.].

{thm:rqi:cubic}

Theorem 2.5 (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 two 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 $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ 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

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

[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.

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

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

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. Derive this

Proof of Theorem 2.5. Assume that the eigenvalue that is computed is *simple* (cf. Definition **??**) 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 $\mathbf{\Lambda}$ we obtain

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

where $\eta \coloneqq -2\mathbf{e}_1^\mathsf{T}\mathbf{\Lambda}\mathbf{d}^{(i)} - (\mathbf{d}^{(i)})^\mathsf{T}\mathbf{\Lambda}\mathbf{d}^{(i)} = \lambda_1\epsilon^2 - (\mathbf{d}^{(i)})^\mathsf{T}\mathbf{\Lambda}\mathbf{d}^{(i)}$. 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

$$|\eta| \leq |\lambda_1|\epsilon^2 + \|\mathbf{\Lambda}\|\|\mathbf{d}^{(i)}\|^2 \leq |\lambda_{\max}|\epsilon^2 + \|\mathbf{\Lambda}\|\epsilon^2 = 2\|\mathbf{\Lambda}\|\epsilon^2 \tag{2.9}$$ {eq:eta:bound}

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

$$\begin{aligned}&\mathbf{y}^{(i+1)} = (\mathbf{\Lambda} - \mu^{(i)})^{-1}\mathbf{x}^{(i)} \\ &= \left(\frac{x_1^{(i)}}{\lambda_1 - \mu^{(i)}}, \frac{x_2^{(i)}}{\lambda_2 - \mu^{(i)}}, \cdots, \frac{x_n^{(i)}}{\lambda_n - \mu^{(i)}}\right)^\mathsf{T} && (\text{since } (\mathbf{\Lambda} - \mu^{(i)}\boldsymbol{I})^{-1} = \operatorname{diag}\left(\frac{1}{(\lambda_j - \mu^{(i)})}\right)) \\ &= \left(\frac{1 + d_1^{(i)}}{\lambda_1 - \mu^{(i)}}, \frac{d_2^{(i)}}{\lambda_2 - \mu^{(i)}}, \cdots, \frac{d_n^{(i)}}{\lambda_n - \mu^{(i)}},\right)^\mathsf{T} && (\text{since } \mathbf{x}^{(i)} = \mathbf{e}_1 + \mathbf{d}^{(i)}) \\ &= \left(\frac{1 - \epsilon^2/2}{\eta}, \frac{d_2^{(i)}}{\lambda_2 - \lambda_1 + \eta}, \cdots, \frac{d_n^{(i)}}{\lambda_n - \lambda_1 + \eta},\right)^\mathsf{T} && (\text{since } \mu^{(i)} = \lambda_1 - \eta,\ d_1^{(i)} = -\epsilon^2/2) \\ &= \frac{1 - \epsilon^2/2}{\eta}\left(1, \frac{d_2^{(i)}\eta}{(1 - \epsilon^2/2)(\lambda_2 - \lambda_1 + \eta)}, \cdots, \frac{d_n^{(i)}\eta}{(1 - \epsilon^2/2)(\lambda_n - \lambda_1 + \eta)},\right)^\mathsf{T} \\ &=: \frac{1 - \epsilon^2/2}{\eta}\left(\mathbf{e}_1 + \hat{\mathbf{d}}^{(i+1)}\right).\end{aligned}$$

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

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

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

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

we can bound the denominators of $\hat{\mathbf{d}}^{(i+1)}$ using $|\lambda_j - \lambda_1 + \eta| \geq \delta - |\eta|$, and in conjunction with (2.9) 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. □

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 [15, 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 particularized 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 choosen as the shift), see for example [20, p. 208] or [1, pp. 89 sq.].

3 Complex Rayleigh Quotient Iteration

In this chapter we introduce *Complex Rayleigh Quotient Iteration* (CRQI[1]). This is a novel shift and invert type algorithm similar to classic RQI that overcomes some of its disadvantages. .

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

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

Due to the pairwise orthogonality of the eigenvectors this implies

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

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

[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*.

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 neighboring eigenvalues is increased. 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 this matrix 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.

Rewrite

The idea is now to apply Rayleigh Quotient iteration to the matrix

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

{eq:a:tilde}

instead of $\boldsymbol{A}$, where $\gamma > 0$ is positive real number and i denotes the imaginary unit. To get an intuition on why we choose this matrix, we first note that the matrix $\boldsymbol{I} - \mathbf{u}\mathbf{u}^\mathsf{T}$, interpreted as a linear mapping, defines the orthogonal projection onto the span of $\mathbf{u}$. Of course, this works for any vector $\mathbf{v} \in \mathbb{R}^n$ and can be seen as follows. Given a vector $\mathbf{x} \in \mathbb{R}^n$, the projection of $\mathbf{x} \in \mathbb{R}^n$ onto the span of $\mathbf{v}$ is

$$\frac{\langle \mathbf{v}, \mathbf{x} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}\,,$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product on $\mathbb{R}^n$. Thus, the projection onto the orthogonal complement of the span of $\mathbf{v}$ is

$$\mathbf{x} - (\mathbf{v}^\mathsf{T}\mathbf{x})\mathbf{v} = \mathbf{x} - (\mathbf{v}\mathbf{v}^\mathsf{T})\mathbf{x} = (\boldsymbol{I} - \mathbf{v}\mathbf{v}^\mathsf{T})\mathbf{x}\,.$$

Look up projection formulas, search references and correct derviation

Therefore, a vector $\mathbf{x}$ that is almost parallel to $\mathbf{u}$ will bareley "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

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

{eq:a:tilde:multortl

Now recall that $\mathbf{u}$ is assumed to be a good approximation of an eigenvector $\mathbf{v}_k$ of $\boldsymbol{A}$. This also implies that the orthogonal subspace of the span of $\mathbf{u}$ approximates the orthogonal subspace spanned by $\mathbf{v}_k$. However, the latter is the space spanned by the remaining eigenvectors of $\boldsymbol{A}$, since the eigenvectors are pairwise orthogonal. 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 $\lambda_j - \gamma i$ due to (3.3). The eigenvalue corresponding to $\mathbf{v}_k$ would then be approximately equal to λ_k since $\tilde{\boldsymbol{A}}\mathbf{v}_k \approx \boldsymbol{A}\mathbf{v}_k = \lambda_k\mathbf{v}_k$.

To make this intuition more quantitative, we first need the following result where we replace $\mathbf{u}$ by the exact target eigenvector $\mathbf{v}_k$.

{lem:eigs:atilde0}

Lemma 3.1. *The matrix $\tilde{\boldsymbol{A}}^{(0)} \coloneqq \boldsymbol{A} - \gamma i(\boldsymbol{I} - \mathbf{v}_k\mathbf{v}_k^\mathsf{T})$ has the same eigenvectors as $\boldsymbol{A}$ with corresponding eigenvalues $\lambda_j(\tilde{\boldsymbol{A}}^{(0)}) = \lambda_j - \gamma i$ for $j \neq k$ and $\lambda_k(\tilde{\boldsymbol{A}}^{(0)}) = \lambda_k$.*

Introduce eigenvalue notation in chapter 1

Proof. First consider $j \neq k$. We have

$$(\boldsymbol{A} - \gamma i(\boldsymbol{I} - \mathbf{v}_k\mathbf{v}_k^\mathsf{T}))\mathbf{v}_j = \boldsymbol{A}\mathbf{v}_j - \gamma i\mathbf{v}_j + \gamma i\mathbf{v}_k\underbrace{\mathbf{v}_k^\mathsf{T}\mathbf{v}_j}_{=0} = \lambda_j\mathbf{v}_j - \gamma i\mathbf{v}_j = (\lambda_j - \gamma i)\mathbf{v}_j\,.$$

For $j = k$ we have $\mathbf{v}_k^\mathsf{T}\mathbf{v}_j = \mathbf{v}_k^\mathsf{T}\mathbf{v}_k = 1$ and thus

$$(\boldsymbol{A} - \gamma i(\boldsymbol{I} - \mathbf{v}_k\mathbf{v}_k^\mathsf{T}))\mathbf{v}_k = \lambda_k\mathbf{v}_k - \gamma i\mathbf{v}_k + \gamma i\mathbf{v}_k = \lambda_k\mathbf{v}_k\,.$$

□

We now decompose $\tilde{\boldsymbol{A}}$ into the sum $\tilde{\boldsymbol{A}} = \tilde{\boldsymbol{A}}^{(0)} + \tilde{\boldsymbol{A}}^{(1)}$, where $\tilde{\boldsymbol{A}}^{(0)}$ is defined

in Lemma 3.1 and $\tilde{\boldsymbol{A}}^{(1)}$ is given by

$$\tilde{\boldsymbol{A}}^{(1)} \coloneqq \gamma i(\mathbf{u}\mathbf{u}^{\mathsf{T}} - \mathbf{v}_k\mathbf{v}_k^{\mathsf{T}})\,.$$

Weyl's perturbation inequality then gives the bound

$$\max_{k=1}^{n} |\lambda_k(\tilde{\boldsymbol{A}}) - \lambda_k(\tilde{\boldsymbol{A}}^{\mathbf{0}})| \leq \left\|\tilde{\boldsymbol{A}}^{(1)}\right\|. \tag{3.4}$$ {eq:weyl}

To compute the norm on the right hand side of (3.4) we make use of the following well-known facts from linear algebra.

Proposition 3.2. *Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ be a real symmetric matrix. Then*

(a) The sum of the eigenvalues of $\boldsymbol{A}$ is equal to the trace *of $\boldsymbol{A}$ that is defined as the sum of its diagonal entries, i. e., ,*

$$\sum_{j=1}^{n} \lambda_j(\boldsymbol{A}) = \sum_{k=1}^{n} \boldsymbol{A}_{kk}\,.$$

Bibliography

[1] 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 p. 27).

[2] S. 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. DOI: 10.1098/rspa.1951.0129 (cit. on p. 21).

[3] James W. Demmel. *Applied Numerical Linear Algebra*. Philadelphia: Society for Industrial and Applied Mathematics, 1997 (cit. on pp. 17, 24).

[4] 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. 7, 14).

[5] Kenneth M. Hoffman and Ray A. Kunze. *Linear Algebra*. Prentice-Halll Inc., Englewood Cliffs, New Jersey, 1971 (cit. on pp. 8, 19).

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

[7] 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)* 1846.30 (Jan. 1846), pp. 51–94 (cit. on p. 7).

[8] Walter Kohn. "A Variational Iteration Method for Solving Secular Equations". In: *The Journal of Chemical Physics* 17.7 (1949), pp. 670–670. DOI: 10.1063/1.1747366 (cit. on p. 21).

[9] 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. 22, 23).

[10] 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. 22, 23).

[11] 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. 22, 23).

[12] 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. 22).

[13] 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. 22).

[14] 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. 22).

[15] Beresford N. Parlett. *The Symmetric Eigenvalue Problem*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1998 (cit. on pp. 17, 27).

[16] G. Peters and J. H. Wilkinson. "Inverse Iteration, Ill-Conditioned Equations and Newton's Method". In: *SIAM Review* 21.3 (1979), pp. 339–360 (cit. on p. 14).

[17] 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).

[18] Yousef Saad. *Numerical Methods for Large Eigenvalue Problems*. Society for Industrial and Applied Mathematics, 2011 (cit. on pp. 11, 12, 19).

[19] 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 (Jan. 2018), pp. 3–55 (cit. on pp. 14, 20).

[20] Lloyd N. Trefethen and David Bau. *Numerical linear algebra*. Other titles in applied mathematics. Society for Industrial and Applied Mathematics, 1997 (cit. on pp. 7, 17, 27).

[21] 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. 22).

[22] J. H. Wilkinson. *The Algebraic Eigenvalue Problem*. Monographs On Numerical Analysis. Oxford: Clarendon Press, 1965 (cit. on p. 14).