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

Bachelor thesis

Rayleigh Quotient iteration with imaginary shifts

Author:
Nils Friess

Supervisor:
Prof. Dr. Robert Scheichl

September 2, 2019

Declaration of Authorship

I, Nils Friess, declare that this thesis titled, "Rayleigh Quotient iteration with imaginary shifts" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

Abstract

Nils Friess

Rayleigh Quotient iteration with imaginary shifts

The Thesis Abstract is written here (and usually kept to just this page). The page is kept centered vertically so can expand into the blank space above the title too. . .

Contents

1 Introduction

We consider the symmetric eigenvalue problem, i. e. given a symmetric matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ are looking for $0 \neq \mathbf{v} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ such that

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

We assume that $\boldsymbol{A}$ is positive definite which implies, since $\boldsymbol{A}$ is also symmetric, that all eigenvalues of $\boldsymbol{A}$ are positive. In the following we are especially interested in finding *interior* eigenpairs of $\boldsymbol{A}$. By this we mean that if the eigenvalues of $\boldsymbol{A}$ are labeled in increasing order of magnitude $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ we are looking for $(\lambda_k, \mathbf{v}_k)$ where $1 < k < n$.

Example 1.1 (Discretization of the LaPlacian)**.** This example describes an eigenvalue problem that naturally arises in the investigation of resonance frequencies[1]. We consider the oscillations of a string of unit length and represent the string by a function

$$u : [0,1] \times \mathbb{R} \to \mathbb{R}, \qquad (x,t) \mapsto u(x,t)$$

where $u(x,t)$ models the displacement of the string at time t in the y direction of a point originally at position x (if we assume the string to lie in the $x-y$-plane). We assume the string is fixed at both of is ends. We want analyse the strings resonance frequencies, i. e. we want to find *standing waves*. To this end, consider the following set of equations

$$\frac{\partial^2 u}{\partial x^2}(x,t) - c^2 \frac{\partial^2 u}{\partial t^2}(x,t) = 0 \qquad \text{for all } t \in \mathbb{R} \text{ and } x \in (0,1)\,, \tag{1.2a}$$

$$u(t,0) = u(t,1) = 0 \qquad \text{for all } t \in \mathbb{R}\,, \tag{1.2b}$$

where $c > 0$ is a paramater describing the string's properties. Equation (1.2a) is called the *wave equation* in one dimension. The boundary conditions (1.2b) encode the fact that the string is fixed at its ends. Since we try to find standing waves of the string we can also To solve this partial differential equation (PDE) we will use use separation of variables to make the following ansatz. Assume the solution is of the form $u(x,t) = X(x)T(t)$. Equation (1.2a) becomes

[1] This example is adapted from [BM12, pp. 1–3]

We can now rewrite this problem as an eigenvalue problem. To this end, we define the differential operator L for $u \in \mathscr{C}^2(0,1)$ by

$$L[u](x) \coloneqq -cu''(x) \quad \text{for all } x \in (0,1)\,.$$

We obtain the following eigenvalue problem

$$L[u_0] = \lambda u_0 \tag{1.3}$$

on the inifinte-dimensional space $\mathscr{C}^2(0,1)$. In order to solve this problem numerically, we have to discretize it. Taylor expansion of u_0 around some $x \in (0,1)$ yields the *central finite difference quotient*

$$u_0''(x) = \frac{u_0(x+h) - 2u_0(x) + u_0(x-h)}{h^2} + O(h^2).$$

If we only consider u_0 in the discrete mesh points $x_i = ih$, $i = 0, \dots, m$, for $h = m^{-1}$ and neglect the $O(h^2)$ term we obtain

$$u_0''(x_i) \approx \frac{1}{h^2}\left(u_0(x_{i-1}) - 2u_0(x_i) + u_0(x_{i+1})\right) \quad \text{for all } i = 1, \dots, m-1\,.$$

This allows us to approximate the left-hand side of (1.3) and we get

$$\frac{c}{h^2}\left(2u_0(x_i) - u_0(x_{i+1}) - u_0(x_{i+1})\right) \approx \lambda u_0(x_i) \quad \text{for all } i = 1, \dots, m-1\,. \tag{1.4}$$

From the boundary conditions in (1.2b) we already know $u_0(x_0) = u_0(x_m) = 0$. This allows us to rewrite (1.3) to obtain the following algebraic eigenvalue problem

$$\boldsymbol{A}\mathbf{u} \coloneqq \frac{c}{h^2}\begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix}\begin{pmatrix} u_0(x_1) \\ u_0(x_2) \\ \vdots \\ u_0(x_{m-1}) \\ u_0(x_m) \end{pmatrix} \approx \lambda \begin{pmatrix} u_0(x_1) \\ u_0(x_2) \\ \vdots \\ u_0(x_{m-1}) \\ u_0(x_m) \end{pmatrix} \tag{1.5}$$

Assumming $c = 1$ the eigenvalues and eigenvectors of $\boldsymbol{A}$ can be explicitly determined (cf. Appendix)

$$\lambda_k = 4m^2 \sin^2\left(\frac{k\pi}{2m}\right) \qquad \text{and} \qquad v_k = \left(\sin\left(\frac{k\pi j}{m}\right)\right)_{j=1}^{m-1}. \tag{1.6}$$

In the following we will introduce a numerical method to iteratively compute approximations of the eigenpairs. We will, however, need approximations of the

eigenvector that is to be computed. To that end we will use the exact values in (1.6) and perturb them. This model problem

1.1 Rayleigh quotient iteration

In this section we introduce the Rayleigh quotient iteration, an iterative method to compute approximations of eigenpairs of a matrix $\boldsymbol{A}$, i. e. to compute (approximate) soltuions of (1.1). We start by revisiting the power method in Algorithm 1.2 before introducing the (shifted) inverse iteration (Algorithm 1.4) which directly leads us to the Rayleigh quotient iteration in Algorithm 1.6.

The power method is based on generating the sequence $\boldsymbol{A}^k\mathbf{v}_0$ where $\mathbf{v}_0$ is a non-zero vector. When normalised appropriately this sequence convergences to the eigenvector associated with the eigenvalue of largest modulus (under certain resonable assumptions, cf. [Saa11, Theorem 4.1, p. 86]). In Algorithm 1.2 we normalise by ensuring that the largest component of the current approximation is equal to one.

Algorithm 1.2: Power method

begin
  Choose nonzero initial vector $\mathbf{v}_0$
  for $k = 1, 2, \ldots$ *until convergence* **do**
    $\mathbf{v}_k = \frac{1}{\alpha_k}\boldsymbol{A}\mathbf{v}_{k-1}$
    `/* ` α_k `is the component of` $\boldsymbol{A}\mathbf{v}_{k-1}$ `with the maximum modulus */`

The next algorithm makes use of the following two facts from linear algebra.

Lemma 1.3. *Let $(\lambda, \mathbf{v})$ be an eigenpair of $\boldsymbol{A}$ and $\sigma \in \mathbb{R}$ an arbitrary scalar. Then*

1. *$\lambda - \sigma$ is an eigenvalue of $\boldsymbol{A} - \sigma\boldsymbol{I}$ with eigenvector $\mathbf{v}$.*
2. *If $\boldsymbol{A}$ is nonsingular then λ^{-1} is an eigenvalue of $\boldsymbol{A}^{-1}$ with eigenvector $\mathbf{v}$.*

Proof. Since $\boldsymbol{A}\mathbf{v} = \lambda\mathbf{v}$, we have $\boldsymbol{A}\mathbf{v} - \sigma\mathbf{v} = \lambda\mathbf{v} - \sigma\mathbf{v}$ and therefore $(\boldsymbol{A} - \sigma\boldsymbol{I})\mathbf{v} = (\lambda - \sigma)\mathbf{v}$, which proofs the first statement. For the second statement, we first remark that since $\boldsymbol{A}$ is nonsingular, $\lambda \neq 0$. Therefore $\boldsymbol{A}^{-1}\boldsymbol{A}\mathbf{v} = \lambda\boldsymbol{A}^{-1}\mathbf{v}$, thus $\lambda^{-1}\mathbf{v} = \boldsymbol{A}^{-1}\mathbf{v}$. □

We now alter the power method by defining the iterate as follows

$$\mathbf{v}_k = \frac{1}{\alpha_k}(\boldsymbol{A} - \sigma\boldsymbol{I})^{-1}\mathbf{v}_{k-1}\,. \tag{1.7}$$

The method will then converge to the eigenvector corresponding to the largest eigenvalue in modulus of $(\boldsymbol{A} - \sigma\boldsymbol{I})^{-1}$, or to the one corresponding to the smallest eigenvalue of $\boldsymbol{A} - \sigma\boldsymbol{I}$. We can make use of this fact to compute interior eigenpairs of $\boldsymbol{A}$. To that end, consider the eigenvalue λ_k, $1 < k < n$. We can choose a constant $\sigma \in \mathbb{R}$

such that

$$\alpha_1 = \frac{1}{\lambda_k - \sigma}$$

is the dominant eigenvalue of $(\boldsymbol{A} - \sigma \boldsymbol{I})^{-1}$. In other words, the method converges to the eigenvector corresponding to the eigenvalue closest to σ. In practice, it is not necessary to compute the inverse of $\boldsymbol{A} - \sigma \boldsymbol{I}$ explicitly. Instead, the $\boldsymbol{LU}$ decomposition is computed once before the iteration. It is then sufficient to solve an upper and lower triangular system in each step, reducing the computational complexity from $O(n^3)$ to $O(n^2)$. The method is summarised in Algorithm 1.4.

Algorithm 1.4: Shifted inverse iteration

begin
  Choose nonzero initial vector $\mathbf{v}_0$
  Compute $\boldsymbol{LU}$ decomposition of $\boldsymbol{A} - \sigma \boldsymbol{I} = \boldsymbol{LU}$
  for $k = 1, 2, \ldots$ *until convergence* **do**
    $\mathbf{v}_k = \frac{1}{\alpha_k}(\boldsymbol{A} - \sigma \boldsymbol{I})^{-1}\mathbf{v}_{k-1} = \frac{1}{\alpha_k}\boldsymbol{U}^{-1}\boldsymbol{L}^{-1}\mathbf{v}_{k-1}$
    /* α_k is the component of $(\boldsymbol{A} - \sigma \boldsymbol{I})^{-1}\mathbf{v}_{k-1}$ with the maximum modulus */

The Rayleigh quotient iteration is now a straightforward modification of this algorithm. Instead of using the same shift throughout every iteration, we replace the shift by an approximation of the eigenvalue, computed from the current iterate. To obtain such an approximation, we try to find $\alpha \in \mathbb{R}$, such that

$$\|\boldsymbol{A}\mathbf{v} - \alpha \mathbf{v}\|_2$$

is minimized. This can be seen as a least squares problem of the form $\mathbf{v}\alpha \approx \boldsymbol{A}\mathbf{v}$. The associated normal equation is given by $\mathbf{v}^\mathsf{T}\mathbf{v}\alpha = \mathbf{v}^\mathsf{T}\boldsymbol{A}\mathbf{v}$, or

$$\alpha = \frac{\mathbf{v}^\mathsf{T}\boldsymbol{A}\mathbf{v}}{\mathbf{v}^\mathsf{T}\mathbf{v}}\,.$$

This result gives rise to the follwing definition.

Definition 1.5 (Rayleigh quotient)**.** Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. The mapping

$$\mathcal{R}_A : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}, \qquad x \mapsto \frac{\mathbf{x}^\mathsf{T}\boldsymbol{A}\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}}$$

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

Note, that for eigenvectors $\mathbf{v}$ of $\boldsymbol{A}$ to the eigenvalue λ the Rayleigh quotient yields

$$\mathcal{R}_A(\mathbf{v}) = \frac{\mathbf{v}^\mathsf{T}\lambda\mathbf{v}}{\mathbf{v}^\mathsf{T}\mathbf{v}} = \lambda\,.$$

We will now alter Algorithm 1.4 by replacing the shift with $\mathcal{R}_A(\mathbf{v}_k)$. Also we replace

the normalisation process by dividing by the norm of the iterate, which is already part of the computation of the Rayleigh quotient. Note that we rewrite the nominator of the Rayleigh quotient by using the inner product $\langle \cdot, \cdot \rangle_2$.

Algorithm 1.6: Rayleigh quotient iteration

begin

Choose initial vector $\mathbf{v}_0$ such that $\|\mathbf{v}_0\|_2 = 1$

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

$\sigma_k = \langle \boldsymbol{A}\mathbf{v}_{k-1}, \mathbf{v}_{k-1} \rangle_2$

$\tilde{\mathbf{v}}_k = (\boldsymbol{A} - \sigma_k \boldsymbol{I})^{-1} \mathbf{v}_{k-1}$

$\mathbf{v}_k = \tilde{\mathbf{v}}_k / \|\tilde{\mathbf{v}}_k\|_2$

A Proofs

Bibliography

[BM12] Steffen Börm and Christian Mehl. *Numerical methods for eigenvalue problems*. Berlin/Boston: Walter de Gruyter, 2012.

[Saa11] Yousef Saad. *Numerical Methods for Large Eigenvalue Problems*. Society for Industrial and Applied Mathematics, 2011.