Chapter 6 Numerical solution of eigenvalue problems

6.1 Theoretical foundations

Definition 6.1 Eigenvalue, eigenvector, spectrum

Let $A \in \mathbb{C}^{n \times n}$. A number $\lambda \in \mathbb{C}$ is called an eigenvalue of A, if there exists a vector $x \in \mathbb{C}^n$, $x \neq 0$ such that

$$(*) \quad A\,x \;=\; \lambda\, x\,.$$

The vector x is called an eigenvector associated with λ. The set of all eigenvalues λ of A is said to be the spectrum of A and will be denoted by $\sigma(A)$.
An immediate consequence of $(*)$ is the following:

$$\lambda \in \sigma(A) \quad \Longleftrightarrow \quad N(A - \lambda I) \;:=\; \{x \in \mathbb{C}^n \setminus \{0\} \mid (A - \lambda I)x = 0\} \;\neq\; \emptyset \quad \Longleftrightarrow \quad \det\,(A - \lambda I) \;=\; 0\,.$$

Definition 6.2 Characteristic polynomial, multiplicities

The polynomial $\varphi_A \in P_n(\mathbb{C})$ given by

$$\varphi_A(\lambda) \;=\; \det\,(A \;-\; \lambda\, I) \quad , \quad \lambda \in \mathbb{C}$$

is called the characteristic polynomial of A.

If $\sigma(A) = \{\lambda_1, \ldots \lambda_k\}$, $\lambda_i \neq \lambda_j$, $1 \leq i \neq j \leq k$, then φ_A can be written as

$$\varphi_A(\lambda) \;=\; (-1)^n \, (\lambda - \lambda_1)^{m_1} \, (\lambda - \lambda_2)^{m_2} \ldots (\lambda - \lambda_k)^{m_k} \quad , \quad m_i \in \mathbb{N} \, , \; 1 \leq i \leq k \, , \; \sum_{i=1}^{k} m_i \;=\; n \, .$$

The number m_i is called the algebraic multiplicity of the eigenvalue λ_i and will be denoted by $\sigma(\lambda_i) := m_i$. On the other hand,

$$\rho(\lambda_i) \;:=\; \dim \left(N(A \;-\; \lambda_i \, I\right)$$

is said to be the geometric multiplicity of λ_i. There holds

$$1 \;\leq\; \rho(\lambda_i) \;\leq\; \sigma(\lambda_i) \;\leq\; n \, .$$

Definition 6.3 Similarity transformation
Assume that $T \in C^{n \times n}$ is regular. Then, the mapping

$$A \;\longmapsto\; B \;:=\; T^{-1} \, A \, T$$

is called a similarity transformation. The matrices A and B are said to be similar.

Lemma 6.4 Invariance properties of similarity tranbsformations
For $A \in C^{n \times n}$, the spectrum $\sigma(A)$, the characteristic polynomial φ_A and the multiplicities $\rho(\lambda)$, $\sigma(\lambda)$, $\lambda \in \sigma(A)$, are invariant with respect to similarity transformations.

Similarity transformations are used to transform matrices to a simpler form, called canonical forms. From a numerical point of view, we are particularly interested in unitary transformations U (i.e., $U^H = U^{-1}$) whose spectral condition number $\kappa_2(U)$ is equal to 1.

Theorem 6.5 The Schur normal form
For each matrix $A \in \mathbb{C}^{n\times n}$ with eigenvalues $\lambda_i \in \mathbb{C}$, $1 \leq i \leq n$, there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that

$$U^H\,A\,U = \begin{pmatrix} \lambda_1 & \star & \cdot & \cdot & \cdot & \star \\ 0 & \lambda_2 & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \star \\ 0 & & & & 0 & \lambda_n \end{pmatrix}.$$

Proof: Induction on n:

(i) 1.Step: $n = 1$ (obvious!)

(ii) 2.Step: Assume that the assertion holds true for $A \in \mathbb{C}^{n-1\times n-1}$.

(iii) 3.Step: Let $A \in \mathbb{C}^{n\times n}$ and $\lambda_1 \in \sigma(A)$ with eigenvector $x_1 \in \mathbb{C}^n$, $\|x_1\|_2 = 1$. Then, there exist vectors $x_2, ..., x_n \in \mathbb{C}^n$, such that $x_1, ..., x_n$ is an orthonormal basis of $\mathbb{C}^n$.

Then, for the unitary matrix $X = (x_1|...|x_n)$ with column vectors x_i , $1 \leq i \leq n$, there holds:

$$X^H\, A\, X\, e_1 \;=\; X^H\, A\, x_1 \;=\; \lambda_1\, X^H\, x_1 \;=\; \lambda_1\, e_1\ ,$$

whence

$$X^H\, A\, X \;=\; \begin{pmatrix} \lambda_1 & a \\ 0 & A_1 \end{pmatrix}\ ,$$

where $A_1 \in C^{n-1 \times n-1}$ and $a^H \in C^{n-1}$.
By assumption, there exists a unitary matrix $U_1 \in C^{(n-1)\times(n-1)}$ such that

$$U_1^H\, A_1\, U_1 \;=\; \begin{pmatrix} \lambda_2 & \star & \cdot & \cdot & \star \\ & \cdot & & & \cdot \\ & & & \cdot & \star \\ 0 & & & & \lambda_n \end{pmatrix}\ .$$

The matrix

$$U \;=\; X \begin{pmatrix} 1 & | & 0 \\ - & - & - \\ 0 & | & U_1 \end{pmatrix}$$

is the unitary matrix we are looking for, since

$$U^H \, A \, U \;=\; \begin{pmatrix} 1 & 0 \\ 0 & U_1^H \end{pmatrix} X^H \, A \, X \begin{pmatrix} 1 & 0 \\ 0 & U_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & U_1^H \end{pmatrix} \begin{pmatrix} \lambda_1 & a \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U_1 \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda_1 & aU_1 \\ 0 & U_1^H A_1 U_1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & \star & \cdot & \star \\ & \cdot & & \cdot \\ & & \cdot & \star \\ & & & \lambda_n \end{pmatrix} .$$

Corollary 6.6 Hermitian matrices

Any Hermitian matrix $A \in \mathbb{C}^{n \times n}$, i.e., $A = A^H$, has n real eigenvalues $\lambda_i \ , \ 1 \leq i \leq n$, with pairwise orthogonal eigenvectors $x_i \ , \ 1 \leq i \leq n$. If $U \in \mathbb{C}^{n \times n} \ , \ U = (x_1|...|x_n)$, denotes the unitary matrix having the eigenvectors as column vectors, there holds:

$$U^H \, A \, U \ = \begin{pmatrix} \lambda_1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & \lambda_n \end{pmatrix} .$$

Proof: The proof follows readily from Theorem 6.5 due to the fact that $U^H A U$ is Hermitian, if A is Hermitian:

$$(U^H \, A \, U)^H \ = \ U^H \, A^H \, (U^H)^H \ = \ U^H \, A \, U \, .$$

Corollary 6.7 Characterization of the eigenvalues of Hermitian matrices
If $A \in \mathbb{C}^{n \times n}$ is a Hermitian matrix, there holds:

$$(+) \quad \lambda_{max}(A) \; = \; \max_{x \in \mathbb{C}^n \setminus \{0\}} \frac{x^H A x}{x^H x} \; ,$$

$$(++) \quad \lambda_{min}(A) \; = \; \min_{x \in \mathbb{C}^n \setminus \{0\}} \frac{x^H A x}{x^H x} \; ,$$

where $\lambda_{max}(A) := \max \{\lambda | \lambda \in \sigma(A)\}$, $\lambda_{min}(A) := \min \{\lambda | \lambda \in \sigma(A)\}$.
In particular, a Hermitian matrix A is positive definite (positive semidefinite), if and only if $\sigma(A) \subset (0, \infty)$ $(\sigma(A) \subset [0, \infty))$.

Proof: Assume $x \in \mathbb{C}^n$, $x \neq 0$, and $y = U^H x = (\eta_1, ..., \eta_n)^T$, where $U^H A U = \Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_n)$.
Then, there holds:

$$\frac{x^H A x}{x^H x} \; = \; \frac{(x^H U) U^H A u (U^H x)}{(x^H U)(U^H x)} \; = \; \frac{y^H \Lambda y}{y^H y} \; = \; \frac{\sum_i \lambda_i |\eta_i|^2}{\sum_i |\eta_i|^2} \; \leq \; \lambda_{max}(A) \; .$$

If we choose x as the eigenvector associated with $\lambda_{max}(A)$, we have equality. The proof of $(++)$ is done analogously replacing A by $-A$.

Definition 6.8 Normal matrices

A matrix $A \in \mathbb{C}^{n \times n}$ is called normal, if

$$A^H A \;=\; A A^H \,.$$

Theorem 6.9 Characterization of normal matrices

A matrix $A \in \mathbb{C}^{n \times n}$ is normal if and only if there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ such that

$$U^H \, A \, U \;=\; \begin{pmatrix} \lambda_1 & & 0 \\ & \cdot & \\ 0 & & \lambda_n \end{pmatrix} \,.$$

Normal matrices have n pairwise orthogonal eigenvectors x_i , $1 \le i \le n$, forming the columns of the unitary matrix U.

Proof: Let $A \in \mathbb{C}^{n \times n}$ be normal. According to Theorem 6.5 there exists a unitary matrix U such that

$$U^H \, A \, U \;=\; \begin{pmatrix} \lambda_1 & \star & \cdot & \star \\ & & \cdot & \cdot \\ 0 & & & \lambda_n \end{pmatrix} \;=:\; R \;=\; (r_{ij})_{i,j=1}^{n} \,.$$

Since $A^H A = A A^H$, it follows that

$$R^H R = U^H A^H U U^H A U = U^H A^H A U = U^H A A^H U = U^H A U U^H A^H U = R R^H .$$

Now, we have

$$(R^H R)_{11} = \bar{\lambda}_1 \lambda_1 = |\lambda_1|^2 ,$$
$$(R R^H)_{11} = |\lambda_1|^2 + \sum_{j=2}^{n} |r_{1j}|^2 ,$$

and hence, $r_{1j} = 0 \ , \ 2 \le j \le n$.
In much the same way we show $r_{ij} = 0 \ , \ i+1 \le j \le n \ , \ 2 \le i \le n$.
On the other hand, assume that A is unitarily diagonalizable. Setting $\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_n)$, we have:

$$A^H A = U \Lambda^H U^H U \Lambda U^H = U \Lambda^H \Lambda U^H = U \Lambda U^H U \Lambda^H U^H = A A^H .$$

6.2 Condition of the eigenvalue problem

As far as the condition of the eigenvalue problem

$$(*) \quad A\,x \;=\; \lambda\,x\,.$$

is concerned, we study the case of an algebraically simple eigenvalue $\lambda \in \sigma(A)$, i.e., $\sigma(\lambda) = 1$. For this purpose, we introduce the notion of the Gâteaux derivative of a mapping $J : V \to C$, where V is a linear space:

Definition 6.10 Gâteaux derivative
Let V be a linear space, $J : V \to C$ and $u \in V$. Then, J is said to be Gâteaux differentiable in $u \in V$, if for all $v \in V$ the following limit exists:

$$J'(u;v) \;=\; \lim_{t \to 0_+} \frac{J(u+tv) \;-\; J(u)}{t}$$

Moreover, if there is a mapping $J'(u) : V \to C$ such that

$$J'(u)v \;=\; J'(u;v) \quad , \quad v \in V\,,$$

$J'(u)$ is referred to as the Gâteaux derivative of J in u.

We assume $V = \mathbb{C}^{n \times n}$ and consider the mapping

$$\begin{aligned} \lambda \, : \, \mathbb{C}^{n\times n} &\rightarrow \mathbb{C} \\ A &\longmapsto \lambda(A) \, , \end{aligned}$$

where $\lambda(A) \in \sigma(A)$ and $\sigma(\lambda(A)) = 1$.

Theorem 6.11 Gâteaux derivative of algebraically simple eigenvalues
Assume $\lambda : \mathbb{C}^{n\times n} \rightarrow \mathbb{C}$ with $A \in \mathbb{C}^{n\times n} \longmapsto \lambda(A)$, where $\lambda(A) \in \sigma(A)$ such that $\sigma(\lambda(A)) = 1$. Moreover, let $x \in \mathbb{C}^n \setminus \{0\}$ be an eigenvector of A associated with $\lambda(A)$ and let $y \in \mathbb{C}^n \setminus \{0\}$ be an eigenvector of A^H associated with $\bar{\lambda}(A)$, i.e., $Ax = \lambda(A)x$ and $A^H y = \bar{\lambda}(A)y$.
Then, the Gâteaux derivative of λ in A is given by

$$\begin{aligned} \lambda'(A) \, : \, \mathbb{C}^{n\times n} &\rightarrow \mathbb{C} \\ C &\longmapsto \lambda'(A)C \; = \; \frac{< Cx, y >}{< x, y >} \, , \end{aligned}$$

where $< \cdot, \cdot >$ denotes the inner product in $\mathbb{C}^n$.

Proof: We consider

$$\lim_{t \to 0_+} \frac{\lambda(A + tC) \; - \; \lambda(A)}{t} \; .$$

Since $\lambda(A)$ is a simple zero of the characteristic polynomial φ_A, there holds:

$$0 \; \neq \; \varphi_A'(\lambda(A)) \; = \; \frac{\partial}{\partial \lambda} \varphi_{A+tC}(\lambda)|_{t=0} \; .$$

According to the Implicit Function Theorem there exists a continuously differentiable function

$$\lambda \; : \; (-\varepsilon, \varepsilon) \; \to \; C \quad , \quad t \; \longmapsto \; \lambda(t)$$

with the property that $\lambda(0) = \lambda(A)$ and that $\lambda(t)$ is a simple eigenvalue of $A + tC$, i.e., $\lambda(t) = \lambda(A + tC)$ and $\sigma(\lambda(A + tC)) = 1$.
Moreover, there exists a continuously differentiable function

$$x \; : \; (-\varepsilon, \varepsilon) \; \to \; C^n \quad , \quad t \; \longmapsto \; x(t)$$

such that $x(0) = x$ and $x(t)$ is an eigenvector of $A + tC$ associated with $\lambda(A + tC)$.

Hence, we obtain

$$\lim_{t \to 0_+} \frac{\lambda(A + tC) \ - \ \lambda(A)}{t} \ = \ \lim_{t \to 0_+} \frac{\lambda(t) \ - \ \lambda(0)}{t} \ = \ \lambda'(0) \ .$$

For the computation of $\lambda'(0)$, we differentiate the equation

$$(A \ + \ t \ C) \ x(t) \ = \ \lambda(t) \ x(t)$$

with respect to t in $t = 0$ and thus get

$$[\ \frac{d}{dt} \ (A \ + \ t \ C) \ ] x(t)|_{t=0} \ + \ (A \ + \ t \ C) \ x'(t)|_{t=0} \ = \ \lambda'(t) \ x(t)|_{t=0} \ + \ \lambda(t) \ x'(t)|_{t=0} \quad \Longrightarrow$$

$$C \ x \ + \ A \ x'(0) \ = \ \lambda(A) \ x'(0) \ + \ \lambda'(0) \ x \ .$$

Scalar multiplication by y yields

$$< Cx, y > \ + \ < Ax'(0), y > = \ \lambda(A) \ < x'(0), y > \ + \lambda'(0) \ < x, y > \ .$$

Now, we have

$$< Ax'(0), y > = < x'(0), A^H y > = < x'(0), \bar{\lambda}(A) y > = \ \lambda(A) \ < x'(0), y > \quad \Longrightarrow \quad \lambda'(0) \ = \ \frac{< Cx, y >}{< x, y >} \ .$$

Corollary 6.12 Absolute and relative condition

Under the assumptions of Theorem 6.11 let α denote the angle between the eigenvectors x of A and y of A^H. Then, for the absolute resp. relative condition of an algebraically simple eigenvalue $\lambda(A) \in \sigma(A)$ there holds:

$$(\star) \quad \kappa_{\mathrm{abs}} = \|\lambda'(A)\| = \frac{\|x\|_2 \|y\|_2}{|<x,y>|} = \frac{1}{|\cos\alpha|} ,$$

$$(\diamond) \quad \kappa_{\mathrm{rel}} = \frac{\|A\|_2}{|\lambda(A)|} \|\lambda'(A)\| = \frac{\|A\|_2}{|\lambda(A) \cos\alpha|} .$$

Proof: According to Theorem 6.11 we have

$$\|\lambda'(A)\| = \sup_{C \neq 0} \frac{|\lambda'(A)C}{\|C\|_2} = \sup_{C \neq 0} \frac{|<Cx,y> / <x,y>|}{\|C\|_2} .$$

Further, the Cauchy-Schwarz inequality yields

$$|<Cx,y>| \leq \|Cx\|_2 \, \|y\|_2 \leq \|C\|_2 \, \|x\|_2 \, \|y\|_2 ,$$

where equality holds if and only if $C = yx^H$ which implies $(\star)$.

Corollary 6.13 Condition of the eigenvalue problem for normal matrices

For normal matrices there holds

$$\kappa_{\mathrm{abs}} \ = \ 1 \ ,$$

i.e., the eigenvalue problem for normal matrices is well conditioned.

6.3 Direct and inverse vector iteration

Given a matrix $A \in \mathbb{C}^{n \times n}$ and starting from an initial vector $x^{(0)} \in \mathbb{C}^n$, the direct vector iteration (von Mises method, power method) computes a sequence of iterates according to

$$(*) \quad x^{(k+1)} = A\, x^{(k)} \quad , \quad k \geq 0 .$$

The motivation is that during the iteration a dominant eigenvalue λ of A prevails and the sequence $(x^{(k)})_{\mathbb{N}}$ converges to an eigenvector associated with this eigenvalue.

Theorem 6.14 Direct vector iteration

Let $A \in \mathbb{C}^{n \times n}$ be diagonalizable and let $\lambda_1 \in \sigma(A)$ be the dominant eigenvalue, i.e.,

$$(+) \quad |\lambda_1| > |\lambda_2| > \ldots > |\lambda_n| .$$

Moreover, let $x^{(0)} \in \mathbb{C}^n$ be an initial vector such that

$$(++) \quad < x^{(0)}, \varphi > \neq 0 \quad , \quad \varphi \in N(A - \lambda_1 I)$$

and let $\|\cdot\|$ be a vector norm on $\mathbb{C}^n$. Then, there holds

(i) The sequence $(y^{(k)})_{\mathbb{N}}$, where $y^{(k)} := x^{(k)}/\|x^{(k)}\|$ and $x^{(k+1)} = Ax^{(k)}$, $k \geq 0$, converges to an eigenvector φ_1 of A associated with the eigenvalue λ_1 satisfying $\|\varphi_1\| = 1$.

(ii) The sequence $(\mu_k)_{\mathbb{N}}$, where $\mu_k := \|x^{(k+1)}\|/\|x^{(k)}\|$, $k \geq 0$, converges to $|\lambda_1|$.

Proof: Being diagonalizable, the matrix A has n linearly independent eigenvectors $\varphi_i, 1 \leq i \leq n$, with $A\varphi_i = \lambda_i\varphi_i$ and $\|\varphi_i\| = 1$. The initial vector $x^{(0)}$ admits the representation

$$x^{(0)} = \sum_{i=1}^{n} \alpha_i \, \varphi_i \ ,$$

where $\alpha_i \neq 0$ according to assumption $(++)$. It follows that

$$x^{(k)} = A^k x^{(0)} = \sum_{i=1}^{n} \alpha_i \, \lambda_i^k \, \varphi_i = \alpha_1 \, \lambda_1^k \, \underbrace{[\, \varphi_1 + \sum_{i=2}^{n} \frac{\alpha_i}{\alpha_1} \, (\frac{\lambda_i}{\lambda_1})^k \, \varphi_i \,]}_{=:z^{(k)}} \ .$$

In view of $|\lambda_i| < |\lambda_1|$, $2 \leq i \leq n$, we have $\lim\limits_{k\to\infty} z^{(k)} = \varphi_1$, and hence (observe $x^{(k+1)} = \alpha_1\lambda_1^{k+1}z^{(k+1)}$),

$$y^{(k)} = \frac{x^{(k)}}{\|x^{(k)}\|} = \frac{z^{(k)}}{\|z^{(k)}\|} \to \pm\varphi_1 \quad (k \to \infty) \quad , \quad \mu_k = \frac{\|x^{(k+1)}\|}{\|x^{(k)}\|} = |\lambda_1| \, \frac{\|z^{(k+1)}\|}{\|z^{(k)}\|} \to |\lambda_1| \quad (k \to \infty) \ .$$

Remark 6.15 Generalization of the convergence result

Theorem 6.14 is also valid in case of an r-fold dominant eigenvalue and for non-diagonalizable matrices.

The proof of Theorem 6.14 shows that the speed of convergence of the direct vector iteration depends on the quotient $|\lambda_2/\lambda_1|$. Hence, the convergence is slow, if λ_1 and λ_2 are close in absolute value. This disadvantage is circumvented when using the inverse vector iteration (Wielandt's method): We assume that we know an approximation $\bar{\lambda}$ of an eigenvalue $\lambda_i \in \sigma(A)$, $1 \leq i \leq n$, such that

$$(\star) \quad |\bar{\lambda} - \lambda_i| < |\bar{\lambda} - \lambda_j| \quad , \quad 1 \leq j \neq i \leq n \, .$$

Under the assumption $(\star)$, it follows that $(\bar{\lambda} - \lambda_i)^{-1}$ has the largest magnitude in absolute value of the eigenvalues of the matrix $(A - \bar{\lambda} I)^{-1}$ so that it is natural to apply the vector iteration to this matrix:

$$(\diamond) \quad (A - \bar{\lambda}\, I)\, x^{(k+1)} = x^{(k)} \quad , \quad k \geq 0 \, ,$$

where $x^{(0)}$ is chosen such that $< x^{(0)}, \varphi > \neq 0$, $\varphi \in N(A - \lambda_i I)$.

Each iteration requires the solution of a linear system with $A - \bar{\lambda}I$ as the coefficient matrix. Therefore, the LR decomposition of this matrix only has to be performed once.
The rate of convergence of the inverse vector iteration is given by

$$\max_{j \neq i} \frac{|\lambda_i - \bar{\lambda}|}{|\lambda_j - \bar{\lambda}|} < 1 ,$$

which is the smaller the better $\bar{\lambda}$ approximates λ_i.

The fact that the coefficient matrix $A - \bar{\lambda}I$ is nearly singular for $\bar{\lambda} \approx \lambda_i$ does not cause any trouble, since we only determine the direction of an eigenvector which is well conditioned as shown by the following example:

Example 6.16 Condition of the inverse vector iteration
Consider the matrix

$$A = \begin{pmatrix} -1 & 3 \\ -2 & 4 \end{pmatrix}$$

with the eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 2$.

Using $\bar{\lambda} = 1 - \varepsilon \ , \ 0 < \varepsilon \ll 1,$ as an approximation of $\lambda_1 = 1$, we obtain the nearly singular matrix

$$A \ - \ \bar{\lambda}I \ = \begin{pmatrix} -2+\varepsilon & 3 \\ -2 & 3+\varepsilon \end{pmatrix} .$$

We compute

$$(A \ - \ \bar{\lambda}I)^{-1} \ = \frac{1}{\varepsilon^2+\varepsilon} \begin{pmatrix} 3+\varepsilon & -3 \\ 2 & -2+\varepsilon \end{pmatrix} .$$

When determining $y^{(k)} = x^{(k)}/\|x^{(k)}\| \ , \ x^{(k)} = (A - \bar{\lambda}I)^{-1}x^{(k-1)},$, the factor $1/(\varepsilon^2+\varepsilon)$ is cancelled. Hence, the computation of a direction of the solution of $(A - \bar{\lambda}I)x = b$ is well conditioned.

Remark 6.17 Computational work for the inverse vector iteration

The LR decomposition of $A - \bar{\lambda}I$ and the solution of a staggered system at each iteration is computationally cheap, if A is a tridiagonal matrix ($O(n)$ essential operations).
Therefore, the next sections deal with unitary transformations of Hermitian matrices to tridiagonal form. In this context, we will also encounter methods for the simultaneous computation of all eigenvalues.

6.4 The methods of Givens and Jacobi

6.4.1 The Givens method

The Givens method enables to transform a Hermitian matrix $A \in \mathbb{C}^{n\times n}$ to tridiagonal form $A^{(m)}$ by a finite number of unitary transformations

$$A^{(0)} := A \to A^{(1)} \to \ldots \to A^{(m)} \quad , \quad A^{(k)} = U_k^H \, A^{(k-1)} \, U_k \quad , \quad 1 \le k \le m$$

where the unitary matrices U_k are of the form $U_k = \Omega_{ij}$ with

$$\Omega_{ij} = \begin{pmatrix} 1 & & & & & & & 0 \\ & \ddots & & & & & & \\ & & \cos\varphi & & & -\exp(-i\psi)\sin\varphi & & \\ & & & 1 & & & & \\ & & & & \ddots & & & \\ & & & & 1 & & & \\ & & \exp(i\psi)\sin\varphi & & & \cos\varphi & & \\ & & & & & & \ddots & \\ 0 & & & & & & & 1 \end{pmatrix}$$

For the description of the Givens method, we restrict ourselves to the case of a symmetric matrix $A \in \mathbb{R}^{n \times n}$, where we may choose $\psi = 0$. The application of Ω_{ij} corresponds to a rotation by the angle φ in the plane spanned by the unit vectors e_i and e_j.
In particular, we have:

(•) Left multiplication of A by Ω_{ij}^T changes the rows i and j,

(•) Right multiplication of A by Ω_{ij} changes the columns i and j.

1. Step of the transformation: We choose $U_1 := \Omega_{23}$ such that $a_{31}^{(1)} = a_{13}^{(1)} = 0$.
Setting $c := \cos\varphi$ and $s := \sin\varphi$, we obtain:

$$a_{i2}^{(1)} = a_{2i}^{(1)} = c\, a_{2i} + s\, a_{3i} \quad , \quad a_{i3}^{(1)} = a_{3i}^{(1)} = -s a_{2i} + c\, a_{31} \quad , \quad i \neq 2, 3 \, ,$$
$$a_{22}^{(1)} = c^2\, a_{22} + s^2\, a_{33} + 2\, c\, s\, a_{32} \quad , \quad a_{33}^{(1)} = s^2\, a_{22} + c^2\, a_{33} - 2\, c\, s\, a_{23} \, ,$$
$$a_{32}^{(1)} = a_{23}^{(1)} = -\, c\, s\, (a_{22} - a_{33}) + (c^2 - s^2)\, a_{32} \, .$$

Obviously, we get $a_{31}^{(1)} = a_{13}^{(1)} = 0$ for

$$c := \begin{cases} a_{21}/\sigma & , \quad \sigma \neq 0 \\ 1 & , \quad \sigma = 0 \end{cases} \quad , \quad s := \begin{cases} a_{31}/\sigma & , \quad \sigma \neq 0 \\ 0 & , \quad \sigma = 0 \end{cases} \quad , \quad \text{where} \quad \sigma := (a_{21}^2 + a_{31}^2)^{1/2} \, .$$

2. Step of the transformation: We choose $U_2 = \Omega_{24}$ and determine the angle of rotation φ such that $a_{41}^{(2)} = a_{14}^{(2)} = 0$.

In the general case, the Givens rotations $U_k = \Omega_{ij}$ are successively chosen as the rotations

$$
\begin{array}{cccccc}
\Omega_{23} & \Omega_{24} & \cdot & \cdot & \cdot & \Omega_{2n} \\
 & \Omega_{34} & \cdot & \cdot & \cdot & \Omega_{2n} \\
 & & \cdot & & & \cdot \\
 & & & \cdot & & \cdot \\
 & & & & & \Omega_{n-1,n}
\end{array}
$$

6.4.2 The Jacobi method

The Jacobi method realizes the iterative transformation of a Hermitian matrix $A \in \mathbb{C}^{n \times n}$ by transformations $A^{(k)} = U_k^H A^{(k-1)} U_k \ , \ k \geq 1 \ , \ A^{(0)} := A$, with unitary matrices $U_k = \Omega_{ij}$ such that

$$A^{(k)} \ \rightarrow \ \Lambda \ = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad (k \rightarrow \infty)$$

where $\lambda_i \in \sigma(A) \ , \ 1 \leq i \leq n$.
Again, we restrict ourselves to the case of a symmetric matrix $A \in \mathbb{R}^{n \times n}, A^T = A$.
The transformation

$$A^{(k)} \ = \ (a_{rs}^{(k)})_{r,s=1}^n \ \rightarrow \ A^{(k+1)} \ = \ A^{(k+1)} \ = \ \Omega_{ij}^T \ A^{(k)} \ \Omega_{ij} \ = \ (a_{rs}^{(k+1)})_{rs=1}^n$$

requires the computation of $c := \cos\varphi$ and $s := \sin\varphi$ such that $a_{ij}^{(k+1)} = 0$.

As we already know from Givens' method, the elements of $A^{(k)}$ are changed in the rows and columns i and j:

$$A^{(k+1)} = \begin{pmatrix} a_{11}^{(k)} & \cdots & a_{1i}^{(k)} & \cdots & a_{1j}^{(k)} & \cdots & a_{1n}^{(k)} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{i1}^{(k)} & \cdots & a_{ii}^{(k)} & \cdots & a_{ij}^{(k)} & \cdots & a_{in}^{(k)} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{j1}^{(k)} & \cdots & a_{ji}^{(k)} & \cdots & a_{jj}^{(k)} & \cdots & a_{jn}^{(k)} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}^{(k)} & \cdots & a_{ni}^{(k)} & \cdots & a_{nj}^{(k)} & \cdots & a_{nn}^{(k)} \end{pmatrix} .$$

We obtain:

$$a_{ri}^{(k+1)} = a_{ir}^{(k+1)} = c\, a_{ri}^{(k)} + s\, a_{rj^{(k)}} \quad , \quad a_{rj}^{(k+1)} = a_{jr}^{(k+1)} = -\, s a_{ri}^{(k)} + c\, a_{rj} \quad , \quad r \neq i, j \,,$$

$$a_{ii}^{(k+1)} = c^2\, a_{ii}^{(k)} + s^2\, a_{jj}^{(k)} + 2\, c\, s\, a_{ij}^{(k)} \quad , \quad a_{jj}^{(k+1)} = s^2\, a_{ii}^{(k)} + c^2\, a_{jj}^{(k)} - 2\, c\, s\, a_{ij}^{(k)} \,,$$

$$a_{ij}^{(k+1)} = a_{ji}^{(k+1)} = -\, c\, s\, (a_{ii}^{(k)} - a_{jj}^{(k)}) + (c^2 - s^2)\, a_{ij}^{(k)} \,.$$

Observing the relation

$$\tan 2\varphi = 2 \tan\varphi / (1 - \tan^2\varphi)$$

for the rotation angle φ, the condition $a_{ij}^{(k+1)} = a_{ji}^{(k+1)} = 0$ results after multiplication of the corresponding equation by $2/c^2$ in

$$\tan 2\varphi = \frac{2\, a_{ij}^{(k)}}{a_{ii}^{(k)} - a_{jj}^{(k)}} \quad , \quad |\varphi| \leq \pi/4 \,.$$

Since for $a_{ii}^{(k)} \approx a_{jj}^{(k)}$ cancellation may occur, we use instead a numerically more stable formula: We set $\vartheta := \mathrm{cotan} 2\varphi$ with $\vartheta := (a_{ii}^{(k)} - a_{jj}^{(k)})/(2 a_{ij}^{(k)})$ and compute $t := \tan\varphi$ as the smallest root in absolute value of the quadratic equation

$$t^2 + 2\, t\, \vartheta - 1 = 0 \,.$$

As a measure for the convergence of Jacobi's method we use the sum of the squares of the off-diagonal elements

$$S(A^{(k)}) \; = \; \sum_{r \neq s} |a_{rs}^{(k)}|^2 \, .$$

By means of the explicit formulas for $a_{rs}^{(k+1)}$ we compute:

$$0 \; \leq \; S(A^{(k+1)} \; = \; S(A^{(k)}) \; - \; 2\, |a_{ij}^{(k)}|^2 \; < \; S(A^{(k)}) \quad , \quad \text{if} \quad a_{ij}^{(k)} \neq 0 \, .$$

Consequently, the sequence $(S(A^{(k)}))_{\mathbb{N}}$ is a monotonically decreasing sequence of nonnegative numbers and hence, it is convergent. If the rotations Ω_{ij} are performed according to

$$\begin{array}{cccc} \Omega_{12} & \Omega_{13} & \cdot \;\; \cdot & \Omega_{1n} \\ & \Omega_{23} & & \Omega_{2n} \\ & & \cdot & \cdot \\ & & & \Omega_{n-1,n} \quad , \end{array}$$

and cyclically repeated, we can show:

Theorem 6.18 Convergence of Jacobi's method

Assume $A \in \mathbb{R}^{n \times n}$ to be symmetric. If the Jacobi method is realized using cyclic repetition of the rotations as indicated before, there holds: $\lim\limits_{k \to \infty} S(A^{(k)}) \; = \; 0$.

Corollary 6.19 Order of convergence of Jacobi's method

Assume that the symmetric matrix $A \in \mathbb{R}^{n \times n}$ only has algebraically simple eigenvalues. Then, Jacobi's method is quadratically convergent:

$$S(A^{(k+N)} \leq \Delta^{-1} S(A^{(k)})^2 ,$$

where $N := (n-1)n/2$ and $\Delta := \min_{i \neq j} |\lambda_i(A) - \lambda_j(A)|$.

Remark 6.20 Disadvantages of Jacobi's method

(i) Specific matrix structures (e.g., band structures) are not invariant.

(ii) Despite the quadratic convergence, Jacobi's method can not compete with the QR method which is (locally) convergent of third order.

6.5 The Lanczos method

The Lanczos method is an iterative procedure for the approximate computation of the eigenvalues of a Hermitian matrix which only needs one matrix-vector multiplication per iteration.

Reminder: If $A \in \mathbb{C}^{n\times n}$ is Hermitian and $\lambda_{max}(A) := \max\{\lambda \in \sigma(A)\}$, $\lambda_{min}(A) := \min\{\lambda \in \sigma(A)\}$, then Corollary 6.7 tells us

$$\lambda_{max}(A) \;=\; \max_{x\neq 0} \frac{< Ax, x >}{< x, x >} \quad , \quad \lambda_{min}(A) \;=\; \min_{x\neq 0} \frac{< Ax, x >}{< x, x >} \; .$$

Definition 6.21 Rayleigh quotient

Assume $A \in \mathbb{C}^{n\times n}$ to be Hermitian. The function $\mu : \mathbb{C}^n \to \mathbb{R}$ given by

$$\mu(x) \;=\; \frac{< Ax, x >}{< x, x >} \quad , \quad x \neq 0$$

is called the Rayleigh quotient of A.

Corollary 6.7 allows the following characterization of all eigenvalues:

Theorem 6.22 Characterization of the eigenvalues of Hermitian matrices

Let $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$ be the eigenvalues of a Hermitian matrix $A \in \mathbb{C}^{n\times n}$ with associated eigenvectors $\varphi_1, \varphi_2, \ldots, \varphi_n$. Then, there holds

$$\lambda_i = \min_{x \in \mathrm{span}\{\varphi_i,\ldots,\varphi_n\}\setminus\{0\}} \mu(x) = \max_{x \in \mathrm{span}\{\varphi_1,\ldots,\varphi_i\}\setminus\{0\}} \mu(x) .$$

Idea: Approximation of $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ by the solution of the optimization problems with respect to a sequence of subspaces $V_1 \subset V_2 \subset \ldots \subset \mathbb{C}^n$. For this purpose, the Krylov subspaces $K_m(x^{(0)}, A), m \le n,$ with respect to an initial vector $x^{(0)} \in \mathbb{C}^n$ are well suited:

$$K_m(x^{(0)}, A) := \mathrm{span}\{x^{(0)}, Ax^{(0)}, \ldots, A^{m-1}x^{(0)}\} .$$

We expect that the extremals

$$\lambda_{\max}^{(m)} := \max_{x \in K_m(x^{(0)},A)\setminus\{0\}} \mu(x) \le \lambda_{\max}(A) \quad , \quad \lambda_{\min}^{(m)} := \min_{x \in K_m(x^{(0)},A)\setminus\{0\}} \mu(x) \ge \lambda_{\min}(A)$$

approximate $\lambda_{\max}(A), \lambda_{\min}(A)$ for increasing m.

Idea: Construction of an orthonormal basis $v_1, ..., v_m$ of $K_m(x^{(0)}, A)$ such that the orthogonal matrix $Q_m := (v_1|...|v_m)$ enables the transformation to a tridiagonal matrix according to

$$Q_m^T \, A \, Q_m \;=\; T_m$$

Theorem 6.23 Lanczos algorithm

An orthonormal basis of the m-dimensional Krylov subspace $K_m(x^{(0)}, A)$ can be computed by the following three-term recursion:

$$\begin{aligned} w_{k+1} \;&=\; (A \;-\; \alpha_k \, I) \, v_k \;-\; \beta_k \, v_{k-1} \quad , \quad 1 \leq k \leq m-1 \, , \\ \alpha_k \;&=\; < Av_k, v_k > \, , \; \beta_{k+1} \;= < w_{k+1}, w_{k+1} >^{1/2} \, , \\ v_{k+1} \;&=\; w_{k+1}/\beta_{k+1} \, , \end{aligned}$$

where $v_0 := 0$, $v_1 := x^{(0)}/\|x^{(0)}\|_2$, $\beta_1 := 1$.

Proof: Induction on m.

(i) 1.Step: $m = 1$ (obvious!)

(ii) 2.Step: Assume that the assertion holds true for $m - 1$.

(iii) 3.Step: In view of $w_m - Av_{m-1} \in K_m(x^{(0)}, A)$ and the induction assumption $K_{m-1}(x^{(0)}, A) = \mathrm{span}\{v_1, ..., v_{m-1}\}$ there holds:

$$w_m \; - \; Av_{m-1} \; = \; \sum_{j=1}^{m-1} c_j \, v_j \quad , \quad c_j \; = < w_m - Av_{m-1}, v_j > \; .$$

The requirement $w_m \perp \mathrm{span}\{v_1, ..., v_{m-1}\}$ leads to

$$c_j \; = \; - \; < Av_{m-1}, v_j > = \; - \; < v_{m-1}, Av_j > \; .$$

Sine $Av_j \in K_{j+1}(x^{(0)}, A) = \mathrm{span}\{v_1, ..., v_{j+1}\}$ and $v_{m-1} \perp \mathrm{span}\{v_1, ..., v_{m-2}\}$, it follows that

$$c_j \; = \; 0 \quad , \quad 1 \leq j \leq m-3 \; .$$

For $j = m-1$ and $j = m-2$ we obtain

$$\begin{aligned} c_{m-1} \; &= \; - \; < Av_{m-1}, v_{m-1} > = \; - \; \alpha_{m-1} \; , \\ c_{m-2} \; &= \; - \; < v_{m-1}, Av_{m-2} > = \; - \; < v_{m-1}, w_{m-1} - \sum_{j=1}^{m-2} c_j \, v_j > = \\ &= \; - \; < v_{m-1}, w_{m-1} > = \; - \; \beta_{m-1} \; < v_{m-1}, v_{m-1} > = \; - \; \beta_{m-1} \end{aligned}$$

and hence,

$$w_m \; = \; (A \; - \; \alpha_{m-1} \, I)v_{m-1} \; - \; \beta_{m-1} \, v_{m-2} \; .$$

Corollary 6.24 Monotonicity of the eigenvalue approximations

Under the assumptions of Theorem 6.23 there holds

$$\lambda_{\max}^{(m)} = \lambda_{\max}(T_m) \quad , \quad \lambda_{\min}^{(m)} = \lambda_{\min}(T_m) \; ,$$

and by means of $K_m(x^{(0)}, A) \subset K_{m+1}(x^{(0)}, A)$ we obtain

$$\lambda_{\max}(A) \;\geq\; \lambda_{\max}^{(m+1)} \;\geq\; \lambda_{\max}^{(m)} \quad , \quad \lambda_{\min}(A) \;\leq\; \lambda_{\min}^{(m+1)} \;\leq\; \lambda_{\min}^{(m)} \; .$$

Remark 6.25 Restart of the Lanczos algorithm

(i) In general, we do not have $\lambda_{\max}^{(n)} = \lambda_{\max}(A)$, $\lambda_{\min}^{(n)} = \lambda_{\min}(A)$, since $K_n(x^{(0)}, A) \neq \mathbb{R}^n$. There may exist some $m* < n$ satisfying $\beta_{m*+1} = 0$ such that the three-term recursion terminates. In this case, one should restart with a vector $\tilde{x}^{(0)} \in K_{m*}(x^{(0)}, A)^{\perp}$.

(ii) The extreme eigenvalues $\lambda_{\max}^{(m)}$ and $\lambda_{\min}^{(m)}$ usually converge 'fast' to $\lambda_{\max}(A)$ resp. $\lambda_{\min}(A)$ such that relatively few iterations ($m \ll n$) are required to obtain reasonable approximations.

6.6 The QR method

First, we consider the following case:
$A \in \mathbb{C}^{n \times n}$ is diagonalizable with an eigenvector basis $\varphi_1, ..., \varphi_n$ associated with the eigenvalues $\lambda_1, ..., \lambda_n$.

Interpretation of the vector iteration as an iteration on subspaces:

Choose a one-dimensional subspace $S = \mathrm{span}\{x_1\}$ and iterate according to

$$S\,,\ AS\,,\ A^2S\,,\ A^3S\,,\ldots$$

The sequence $(A^kS)_{k \geq 0}$ converges to the eigenspace

$$T \ = \ \mathrm{span}\,\{\varphi_1\} \quad \Longrightarrow \quad d(A^kS, T) \ \to \ 0 \ \ (k \to \infty)\,,$$

where $d(\cdot, \cdot)$ measures the distance of the subspaces:

$$d(A^kS, T) \ = \ \sup_{s \in A^kS} \inf_{t \in T} \frac{\|s \ - \ t\|_2}{\|s\|_2}\,.$$

Generalization: In case $|\lambda_m| > |\lambda_{m+1}|$, $m < n$, consider the subspaces

$$T = \text{span}\,\{\varphi_1, ..., \varphi_m\} \quad , \quad U = \text{span}\,\{\varphi_{m+1}, ..., \varphi_n\} \, .$$

Definition 6.26 Dominant and co-dominant subspace

The A-invariant subspaces T and U are called the dominant resp. the co-dominant subspace.

Theorem 6.27 Subspace iteration

Assume that $S \subset \mathbb{C}^n$, $\dim S = m < n$, satisfying $S \cap U = \{0\}$. Then, there exists a constant $C > 0$ such that

$$d(A^k S, T) \;\le\; C \, \left|\frac{\lambda_{m+1}}{\lambda_m}\right|^k \, .$$

Simultaneous vector iteration: Choose a basis of S and iterate simultaneously on the basis vectors:

$$S \;=\; \text{span}\,\{q_1^{(0)}, ..., q_m^{(0)}\} \quad \Longrightarrow \quad A^k S \;=\; \text{span}\,\{A^k q_1^{(0)}, ..., A^k q_m^{(0)}\} \, .$$

All sequences $(A^k q_i^{(0)})_{k \ge 0}$, $1 \le i \le m$, converge to $\text{span}\{\varphi_1\}$.

Therefore: For $k \gg 1$, the basis $\{A^k q_1^{(0)}, ..., A^k q_m^{(0)}\}$ is ill conditioned.

Remedy: Orthonormalization

If $q_1^{(k)},...,q_m^{(k)}$ is an orthonormal basis of A^kS, compute $Aq_1^{(k)},...,Aq_m^{(k)}$ and find an orthonormal basis $q_1^{(k+1)},...,q_m^{(k+1)}$ of $A^{k+1}S$ by orthonormalization of $Aq_1^{(k)},...,Aq_m^{(k)}$.

Simultaneous iteration with respect to a complete set of orthonormal vectors $q_1^{(0)},...,q_n^{(0)}$:
For $m = 1,..,n-1$, consider

$$\left.\begin{array}{rcl} S_m &:=& \text{span}\,\{q_1^{(0)},...,q_m^{(0)}\} \quad , \quad T_m \,:=\, \text{span}\,\{\varphi_1,...,\varphi_m\}\ , \\ U_m &:=& \text{span}\,\{\varphi_{m+1},...,\varphi_n\} \quad , \quad S_m \cap U_m \,=\, \{0\}\ ,\ |\lambda_m| > |\lambda_{m+1}| \end{array}\right\} \quad \Longrightarrow \quad A^kS_m \to T_m \ (k \to \infty)$$

$$\Longrightarrow \quad q_1^{(k)},...,q_n^{(k)} \to \text{orthonormal basis } q_1,...,q_n \text{ with } T_m = \text{span}\,\{q_1,...,q_m\}\ .$$

Denoting by $\hat{Q}_k$ the unitary matrix $\hat{Q}_k = (q_1^{(k)}|...|q_n^{(k)})$, we have:

$$A^{(k)} = \hat{Q}_k^H\, A\, \hat{Q}_k \to \begin{pmatrix} \lambda_1 & \star & \cdot & \star \\ & \cdot & & \\ & & \cdot & \star \\ 0 & & & \lambda_n \end{pmatrix} .$$

Reminder: If $A \in \mathbb{C}^{n \times n}$ is regular, there exist a unitary matrix Q and an upper triangular matrix R such that $r_{ii} > 0\ ,\ 1 \leq i \leq n,$ and $A = QR$. Q and R are uniquely determined.

QR algorithm:

Simultaneous iteration in matrix form with orthonormalization by QR decomposition:

Setting $\hat{Q}_k = (q_1^{(k)}|\ldots|q_n^{(k)})$, define

$$D_{k+1} := A\,\hat{Q}_k = (Aq_1^{(k)} \mid \ldots \mid Aq_n^{(k)})\ .$$

Orthonormalization of D_{k+1} by QR decomposition:

$$D_{k+1} = \hat{Q}_{k+1}\,R_{k+1}\ .$$

For the convergence analysis, compute

$$A^{(k)} = \hat{Q}_k^H\,A\,Q_k$$

and check the deviation from upper triangular form.

QR method: Simultaneous iteration with $\hat{Q}_0 = (e_1|\ldots|e_n) = I$:

1.Step:

$$\left.\begin{aligned} D_1 &= A\,\hat{Q}_0 = A \\ D_1 &= \hat{Q}_1\,R_1 \end{aligned}\right\} \Longrightarrow \left\{\begin{aligned} A &= Q_1\,R_1 \quad \text{with } Q_1 := \hat{Q}_1 \\ A^{(1)} &= \hat{Q}_1^H\,A\,Q_1 = R_1\,Q_1 \end{aligned}\right.$$

2.Step:

$$\left.\begin{aligned} D_2 &= A\,\hat{Q}_1 = Q_1\,A^{(1)} \\ A^{(1)} &= Q_2\,R_2 \end{aligned}\right\} \Longrightarrow \left\{\begin{aligned} D_2 &= \underbrace{(Q_1\,Q_2)}_{=:\,\hat{Q}_2}\,R_2 \\ A^{(2)} &= \hat{Q}_2^H\,A^{(1)}\,Q_2 = R_2\,Q_2 \end{aligned}\right.$$

In general: Set $A^{(0)} := A$. For $k = 1, 2, \ldots$, compute

$$\begin{aligned} A^{(k-1)} &= Q_k\,R_k\,, \\ A^{(k)} &= R_k\,Q_k\,. \end{aligned}$$

Theorem 6.28 Convergence of the QR method

Assume that the matrix $A \in C^{n \times n}$ has eigenvalues that are different in absolute value

$$|\lambda_1| > |\lambda_2| > ... > |\lambda_n| .$$

Moreover, assume that for U with $A = U^{-1} \Lambda U$, $\Lambda = \mathrm{diag}(\lambda_i)$ there exists an LR decomposition. Then, there exist phase matrices

$$S_k = \mathrm{diag}\,(\exp(i\varphi_1^{(k)}), ..., \exp(i\varphi_n^{(k)})) ,$$

such that for the matrices $A^{(k)}, Q_k$ and R_k of the QR method there holds:

$$\lim_{k\to\infty} S_{k-1}^H \, Q_k \, S_k = I \quad , \quad \lim_{k\to\infty} S_k^H \, R_k \, S_{k-1} = \begin{pmatrix} \lambda_1 & \star & \cdot & \star \\ & \cdot & & \\ & & \cdot & \\ 0 & & & \lambda_n \end{pmatrix} ,$$

$$\lim_{k\to\infty} a_{jj}^{(k)} = \lambda_j \quad , \quad 1 \le j \le n \quad , \quad |a_{jj}^{(k)}| = O(|\frac{\lambda_i}{\lambda_j}|^k) , \; i > j .$$

Theorem 6.29 Conjugate complex eigenvalues

In contrast to the assumption of Theorem 6.28, assume that λ_r, λ_{r+1} is a pair of conjugate complex eigenvalues ($\lambda_{r+1} = \bar{\lambda}_r$), i.e.,

$$|\lambda_1| > |\lambda_2| > \ldots > |\lambda_r| = |\lambda_{r+1}| > \ldots > |\lambda_n| .$$

Then, there holds

$$\lim_{k\to\infty} a_{ij}^{(k)} = 0 \quad , \quad i,j \notin \{r, r+1\} \; , \; i > j \; ,$$

$$\lim_{k\to\infty} a_{jj}^{(k)} = \lambda_j \quad , \quad j \notin \{r, r+1\} \; ,$$

For $k \to \infty$, the eigenvalues of the 2×2 matrices

$$\begin{pmatrix} a_{rr}^{(k)} & a_{r,r+1}^{(k)} \\ a_{r+1,r}^{(k)} & a_{r+1,r+1}^{(k)} \end{pmatrix}$$

converge to λ_r, λ_{r+1}, although the matrices themselves diverge.

Remark 6.30 Shift techniques

The convergence of the QR method depends on the separation of the eigenvalues. Analogous to the inverse iteration, the convergence can be improved by shift techniques.

Assume that τ_k is an approximation of $\lambda \in \sigma(A)$ such that after a possible reordering of the eigenvalues we have

$$|\lambda_1 - \tau_k| \geq |\lambda_2 - \tau_k| \geq \ldots \geq |\lambda_{n-1} - \tau_k| \gg |\lambda_n - \tau_k| \,.$$

QR method with shift:

Choose $A^{(0)} := A$. For $k = 1, 2 \ldots$, compute

$$\begin{aligned} A^{(k-1)} - \tau_{k-1} I &= Q_k \, R_k \,, \\ A^{(k)} &= R_k \, Q_k + \tau_{k-1} I \,. \end{aligned}$$

Strategy for the computation of the shift parameter τ_k:

Consider the 'bottom part' of the matrix $A^{(k)}$

$$\begin{pmatrix} a^{(k)}_{n-1,n-1} & a^{(k)}_{n-1,n} \\ a^{(k)}_{n,n-1} & a^{(k)}_{nn} \end{pmatrix}$$

and choose τ_k as that eigenvalue of this 2×2 matrix which is closer to $a^{(k)}_{nn}$.

6.7 Estimation of eigenvalues

We distinguish between exclusion theorems and inclusion theorems.
Exclusion theorems specify regions of the complex plane that do not contain an eigenvalue (the complement then contains all eigenvalues). In case of inclusion theorems, regions are determined that contain at least one eigenvalue.

Theorem 6.31 Exclusion theorem of Hirsch
Let $A \in \mathbb{C}^{n \times n}$, $\| \cdot \|$ a vector norm on $\mathbb{C}^n$ and $\|A\| = \sup_{\|x\|=1} \|Ax\|$ the associated matrix norm.

Then, there holds:

$$\rho(A) = \max_{1 \leq i \leq n} |\lambda_i| \leq \|A\| ,$$

where $\lambda_i \in \sigma(A)$, $1 \leq i \leq n$.

Proof: Assume that $\lambda \in \sigma(A)$ with associated eigenvector $x \in \mathbb{C}^n \setminus \{0\}$. Then, there holds

$$|\lambda| \, \|x\| = \|\lambda \, x\| = \|Ax\| \leq \|A\| \, \|x\| .$$

Theorem 6.32 Exclusion theorem of Bauer and Fike

Assume that $A, B \in \mathbb{C}^{n\times n}$ and $\lambda \in \sigma(A) \setminus \sigma(B)$. Then, there holds

$$1 \le \|(\lambda I - B)^{-1}(A - B)\| \le \|(\lambda I - B)^{-1}\| \, \|A - B\| .$$

Proof: If $x \in \mathbb{C}^n \setminus \{0\}$ is an eigenvector associated with $\lambda \in \sigma(A)$, we have

$$(A - B)\, x = (\lambda I - B)\, x \quad \Longrightarrow \quad (\lambda I - B)^{-1}(A - B)\, x = x .$$

Now, let us choose the maximum norm $\|\cdot\|_\infty$ as vector norm. Then, the corresponding matrix norm is the row-sum norm. For $B := \mathrm{diag}(A)$, Theorem 6.32 yields:

Theorem 6.33 Gerschgorin's theorem

Let $A = (a_{ij})_{i,j=1}^n \in \mathbb{C}^{n\times n}$. Then, the union of all discs

$$K_i := \{ \mu \in \mathbb{C} \mid |\mu - a_{ii}| \le \sum_{j\ne i} |a_{ij}| \}$$

contains all eigenvalues of A, i.e., $\sigma(A) \subset \cup_{i=1}^n K_i$.

Corollary 6.34 Strengthened version of Gerschgorin's theorem

Under the assumptions of Theorem 6.33 there holds: If the union

$$M_1 := \cup_{j=1}^{k} K_{i_j}$$

of k discs K_{i_j}, $1 \le j \le k$, is disjoint with respect to the union M_2 of the remaining discs, then M_1 contains exactly k eigenvalues and M_2 contains exactly $n - k$ eigenvalues.

Example 6.35 Example illustrating Gerschgorin's theorem

$$A = \begin{pmatrix} 1 & 0.1 & -0.1 \\ 0 & 2 & 0.4 \\ -0.2 & 0 & 3 \end{pmatrix} \quad , \quad \begin{aligned} K_1 &= \{ \mu \mid |\mu - 1| \le 0.2 \} \\ K_2 &= \{ \mu \mid |\mu - 2| \le 0.4 \} \\ K_3 &= \{ \mu \mid |\mu - 3| \le 0.2 \} \end{aligned}$$

Theorem 6.36 Inclusion theorem

Let $A \in \mathbb{C}^{n\times n}$ be normal and $p \in P_n(\mathbb{C})$ as well as $x \in \mathbb{C}^n \setminus \{0\}$. Then, there exists an eigenvalue $\lambda \in \sigma(A)$ such that

$$|p(\lambda)| \le \frac{\|p(A)\,x\|_2}{\|x\|_2}\,.$$

Proof: Since A is normal, there exists a unitary matrix U such that $A = U^H \Lambda U$, where $\Lambda := \mathrm{diag}(\lambda_i)$, $\lambda_i \in \sigma(A)$, $1 \le i \le n$. Hence, we have:

$$\frac{\|p(A)\,x\|_2}{\|x\|_2} = \frac{\|U^H\,p(\Lambda)\,Ux\|_2}{\|x\|_2} = \frac{\|p(\Lambda)\,Ux\|_2}{\|Ux\|_2} \ge \min_{1\le i\le n} |p(\lambda_i)|\,.$$

Corollary 6.37 Theorem of Weinstein

Assume that $A \in \mathbb{C}^{n\times n}$ is normal, let $\mu : \mathbb{C}^n \to \mathbb{C}$ be the Rayleigh quotient and $x \in \mathbb{C}^n \setminus \{0\}$. Then, there holds

$$\sigma(A) \cap \left\{ \lambda \;\middle|\; |\lambda - \mu(x)| \le \sqrt{\frac{< Ax, Ax > \; - \mu(x) \; < x, Ax >}{< x, x >}} \right\} \ne \emptyset\,.$$

Proof: In Theorem 6.36 we choose $p \in P_1(\mathbb{C})$ such that $p(\lambda) := \lambda - \mu(x)$.