# 5

# Approximation of Eigenvalues and Eigenvectors

In this chapter we deal with approximations of the eigenvalues and eigenvectors of a matrix $\mathrm{A} \in \mathbb{C}^{n\times n}$. Two main classes of numerical methods exist to this purpose, *partial* methods, which compute the *extremal* eigenvalues of A (that is, those having maximum and minimum module), or *global* methods, which approximate the whole spectrum of A.
It is worth noting that methods which are introduced to solve the matrix eigenvalue problem are not necessarily suitable for calculating the matrix eigenvectors. For example, the *power method* (a partial method, see Section 5.3) provides an approximation to a *particular* eigenvalue/eigenvector pair.

The *QR method* (a global method, see Section 5.5) instead computes the real Schur form of A, a canonical form that displays *all* the eigenvalues of A but *not* its eigenvectors. These eigenvectors can be computed, starting from the real Schur form of A, with an extra amount of work, as described in Section 5.8.2.

Finally, some *ad hoc* methods for dealing effectively with the special case where A is a symmetric $(n \times n)$ matrix are considered in Section 5.10.

## 5.1 Geometrical Location of the Eigenvalues

Since the eigenvalues of A are the roots of the characteristic polynomial $p_{\mathrm{A}}(\lambda)$ (see Section 1.7), iterative methods must be used for their approximation when $n \geq 5$. Knowledge of eigenvalue location in the complex plane can thus be helpful in accelerating the convergence of the process.

A first estimate is provided by Theorem 1.4,

$$|\lambda| \leq \|\mathrm{A}\|, \qquad \forall \lambda \in \sigma(\mathrm{A}), \tag{5.1}$$

for any consistent matrix norm $\|\cdot\|$. Inequality (5.1), which is often quite rough, states that *all* the eigenvalues of A are contained in a circle of radius $R_{\|\mathrm{A}\|} = \|\mathrm{A}\|$ centered at the origin of the Gauss plane.

Another result is obtained by extending Definition 1.23 to complex-valued matrices.

**Theorem 5.1** *If* $\mathrm{A} \in \mathbb{C}^{n\times n}$*, let*

$$\mathrm{H} = \left(\mathrm{A} + \mathrm{A}^H\right)/2 \qquad \textit{and} \qquad i\mathrm{S} = \left(\mathrm{A} - \mathrm{A}^H\right)/2$$

*be the hermitian and skew-hermitian parts of* A*, respectively,* $i$ *being the imaginary unit. For any* $\lambda \in \sigma(\mathrm{A})$

$$\lambda_{min}(\mathrm{H}) \le \mathrm{Re}(\lambda) \le \lambda_{max}(\mathrm{H}), \quad \lambda_{min}(\mathrm{S}) \le \mathrm{Im}(\lambda) \le \lambda_{max}(\mathrm{S}). \tag{5.2}$$

**Proof.** From the definition of H and S it follows that $\mathrm{A} = \mathrm{H} + i\mathrm{S}$. Let $\mathbf{u} \in \mathbb{C}^n$, $\|\mathbf{u}\|_2 = 1$, be the eigenvector associated with the eigenvalue $\lambda$; the Rayleigh quotient (introduced in Section 1.7) reads

$$\lambda = \mathbf{u}^H \mathrm{A}\mathbf{u} = \mathbf{u}^H \mathrm{H}\mathbf{u} + i\mathbf{u}^H \mathrm{S}\mathbf{u}. \tag{5.3}$$

Notice that both H and S are hermitian matrices, whilst $i$S is skew-hermitian. Matrices H and S are thus unitarily similar to a real diagonal matrix (see Section 1.7), and therefore their eigenvalues are real. In such a case, (5.3) yields

$$\mathrm{Re}(\lambda) = \mathbf{u}^H \mathrm{H}\mathbf{u}, \qquad \mathrm{Im}(\lambda) = \mathbf{u}^H \mathrm{S}\mathbf{u},$$

from which (5.2) follows. $\diamond$

An a priori bound for the eigenvalues of A is given by the following result.

**Theorem 5.2 (of the Gershgorin circles)** *Let* $\mathrm{A} \in \mathbb{C}^{n\times n}$*. Then*

$$\sigma(\mathrm{A}) \subseteq \mathcal{S}_{\mathcal{R}} = \bigcup_{i=1}^{n} \mathcal{R}_i, \qquad \mathcal{R}_i = \{z \in \mathbb{C} : \, |z - a_{ii}| \le \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|\}. \tag{5.4}$$

*The sets* $\mathcal{R}_i$ *are called Gershgorin circles.*

**Proof.** Let us decompose A as $\mathrm{A} = \mathrm{D} + \mathrm{E}$, where D is the diagonal part of A, whilst $e_{ii} = 0$ for $i = 1, \ldots, n$. For $\lambda \in \sigma(\mathrm{A})$ (with $\lambda \neq a_{ii}$, $i = 1, \ldots, n$), let us introduce the matrix $\mathrm{B}_\lambda = \mathrm{A} - \lambda\mathrm{I} = (\mathrm{D} - \lambda\mathrm{I}) + \mathrm{E}$. Since $\mathrm{B}_\lambda$ is singular, there exists a non-null vector $\mathbf{x} \in \mathbb{C}^n$ such that $\mathrm{B}_\lambda \mathbf{x} = \mathbf{0}$. This means that $((\mathrm{D} - \lambda\mathrm{I}) + \mathrm{E})\,\mathbf{x} = \mathbf{0}$, that is, passing to the $\|\cdot\|_\infty$ norm,

$$\mathbf{x} = -(\mathrm{D} - \lambda\mathrm{I})^{-1}\mathrm{E}\mathbf{x}, \qquad \|\mathbf{x}\|_\infty \le \|(\mathrm{D} - \lambda\mathrm{I})^{-1}\mathrm{E}\|_\infty \|\mathbf{x}\|_\infty,$$

and thus

$$1 \le \|(\mathrm{D} - \lambda\mathrm{I})^{-1}\mathrm{E}\|_\infty = \sum_{j=1}^{n} \frac{|e_{kj}|}{|a_{kk} - \lambda|} = \sum_{\substack{j=1 \\ j\neq k}}^{n} \frac{|a_{kj}|}{|a_{kk} - \lambda|}, \tag{5.5}$$

for a certain $k$, $1 \le k \le n$. Inequality (5.5) implies $\lambda \in \mathcal{R}_k$ and thus (5.4). $\diamond$

The bounds (5.4) ensure that any eigenvalue of A lies within the union of the circles $\mathcal{R}_i$. Moreover, since A and $A^T$ share the same spectrum, Theorem 5.2 also holds in the form

$$\sigma(A) \subseteq \mathcal{S}_{\mathcal{C}} = \bigcup_{j=1}^{n} \mathcal{C}_j, \qquad \mathcal{C}_j = \{z \in \mathbb{C} : |z - a_{jj}| \leq \sum_{\substack{i=1\\ i\neq j}}^{n} |a_{ij}|\}. \tag{5.6}$$

The circles $\mathcal{R}_i$ in the complex plane are called row circles, and $\mathcal{C}_j$ column circles. The immediate consequence of (5.4) and (5.6) is the following.

**Property 5.1 (First Gershgorin theorem)** *For a given matrix* $A \in \mathbb{C}^{n\times n}$,

$$\forall \lambda \in \sigma(A), \qquad \lambda \in \mathcal{S}_{\mathcal{R}} \bigcap \mathcal{S}_{\mathcal{C}}. \tag{5.7}$$

The following two location theorems can also be proved (see [Atk89], pp. 588-590 and [Hou75], pp. 66-67).

**Property 5.2 (Second Gershgorin theorem)** *Let*

$$\mathcal{S}_1 = \bigcup_{i=1}^{m} \mathcal{R}_i, \quad \mathcal{S}_2 = \bigcup_{i=m+1}^{n} \mathcal{R}_i.$$

*If* $\mathcal{S}_1 \cap \mathcal{S}_2 = \emptyset$, *then* $\mathcal{S}_1$ *contains exactly* $m$ *eigenvalues of* A, *each one being accounted for with its algebraic multiplicity, while the remaining eigenvalues are contained in* $\mathcal{S}_2$.

**Remark 5.1** Properties 5.1 and 5.2 do not exclude the possibility that there exist circles containing no eigenvalues, as happens for the matrix in Exercise 1. ■

**Definition 5.1** A matrix $A \in \mathbb{C}^{n\times n}$ is called *reducible* if there exists a permutation matrix P such that

$$PAP^T = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix},$$

where $B_{11}$ and $B_{22}$ are square matrices; A is *irreducible* if it is not reducible. ■

To check if a matrix is reducible, the *oriented graph* of the matrix can be conveniently employed. Recall from Section 3.9 that the oriented graph of a real matrix A is obtained by joining $n$ points (called vertices of the graph) $P_1, \ldots, P_n$ through a line oriented from $P_i$ to $P_j$ if the corresponding matrix entry $a_{ij} \neq 0$. An oriented graph is *strongly connected* if for any pair of distinct vertices $P_i$ and $P_j$ there exists an oriented path from $P_i$ to $P_j$. The following result holds (see [Var62] for the proof).

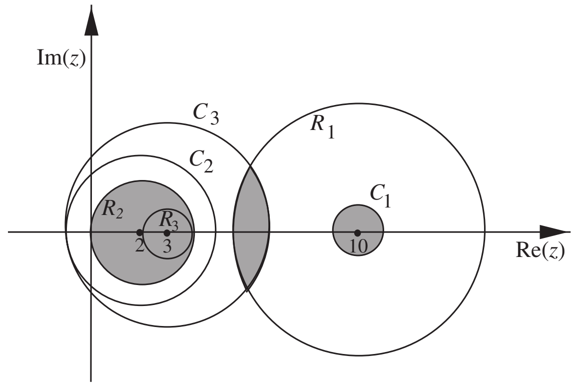


**Fig. 5.1.** Row and column circles for matrix A in Example 5.1

**Property 5.3** *A matrix* $\mathrm{A} \in \mathbb{C}^{n\times n}$ *is irreducible iff its oriented graph is strongly connected.*

**Property 5.4 (Third Gershgorin theorem)** *Let* $\mathrm{A} \in \mathbb{C}^{n\times n}$ *be an irreducible matrix. An eigenvalue* $\lambda \in \sigma(\mathrm{A})$ *cannot lie on the boundary of* $\mathcal{S}_\mathcal{R}$ *unless it belongs to the boundary of every circle* $\mathcal{R}_i$*, for* $i = 1, \ldots, n$*.*

**Example 5.1** Let us consider the matrix

$$\mathrm{A} = \begin{bmatrix} 10 & 2 & 3 \\ -1 & 2 & -1 \\ 0 & 1 & 3 \end{bmatrix},$$

whose spectrum is (to four significant figures) $\sigma(\mathrm{A}) = \{9.687, 2.656 \pm i0.693\}$. The following values of the norm of A: $\|\mathrm{A}\|_1 = 11$, $\|\mathrm{A}\|_2 = 10.72$, $\|\mathrm{A}\|_\infty = 15$ and $\|\mathrm{A}\|_F = 11.36$ can be used in the estimate (5.1). Estimate (5.2) provides instead $1.96 \leq \mathrm{Re}(\lambda(\mathrm{A})) \leq 10.34$, $-2.34 \leq \mathrm{Im}(\lambda(\mathrm{A})) \leq 2.34$, while the row and column circles are given respectively by $\mathcal{R}_1 = \{|z| : |z-10| \leq 5\}$, $\mathcal{R}_2 = \{|z| : |z-2| \leq 2\}$, $\mathcal{R}_3 = \{|z| : |z-3| \leq 1\}$ and $\mathcal{C}_1 = \{|z| : |z-10| \leq 1\}$, $\mathcal{C}_2 = \{|z| : |z-2| \leq 3\}$, $\mathcal{C}_3 = \{|z| : |z-3| \leq 4\}$.
In Figure 5.1, for $i = 1, 2, 3$ the $\mathcal{R}_i$ and $\mathcal{C}_i$ circles and the intersection $\mathcal{S}_\mathcal{R} \cap \mathcal{S}_\mathcal{C}$ (shaded areas) are drawn. In agreement with Property 5.2, we notice that an eigenvalue is contained in $\mathcal{C}_1$, which is disjoint from $\mathcal{C}_2$ and $\mathcal{C}_3$, while the remaining eigenvalues, thanks to Property 5.1, lie within the set $\mathcal{R}_2 \cup \{\mathcal{C}_3 \cap \mathcal{R}_1\}$. •

## 5.2 Stability and Conditioning Analysis

In this section we introduce some a priori and a posteriori estimates that are relevant in the stability analysis of the matrix eigenvalue and eigenvector problem. The presentation follows the guidelines that have been traced in Chapter 2.