

# Chapter 6

## Eigenvalues and eigenvectors

We review the basic properties of eigenvalues and eigenvectors of a matrix, and we present an algorithm for computational approximation in the case of real symmetric matrices.

### 6.1 Eigenvalues and eigenvectors

#### Complex vector spaces

In this section we change the focus from real vector spaces to complex vector spaces. We let  $z \in \mathbb{C}$  denote a complex scalar, with  $\bar{z}$  the complex conjugate of  $z$ .

The complex vector space  $\mathbb{C}^n$  is defined by the basic operations of componentwise addition and scalar multiplication of complex numbers, and with the transpose of a vector  $x \in \mathbb{C}^n$  replaced by the *adjoint*  $x^*$ , corresponding to the transpose with the imaginary parts negated. Similarly, the adjoint of a complex  $m \times n$  matrix  $A = (a_{ij})$  is the  $n \times m$  matrix  $A^* = (\bar{a}_{ji})$ . If  $A = A^*$  the matrix  $A$  is *Hermitian*, and if  $AA^* = A^*A$  it is *normal*.

The inner product of  $x, y \in \mathbb{C}^n$  is defined by

$$(x, y) = x^*y = \sum_{i=1}^n \bar{x}_i y_i, \quad (6.1)$$

with the associated norm for  $x \in \mathbb{C}^n$ ,

$$\|x\| = (x, x)^{1/2}. \quad (6.2)$$

## Matrix spectrum and eigenspaces

We now consider a square matrix  $A \in \mathbb{C}^{n \times n}$  acting on a complex vector space  $\mathbb{C}^n$ . An *eigenvector* of  $A$  is a nonzero vector  $x \in \mathbb{C}^n$ , such that

$$Ax = \lambda x, \quad (6.3)$$

with  $\lambda \in \mathbb{C}$  the corresponding *eigenvalue*. The subspace of  $\mathbb{C}^n$  spanned by the eigenvectors corresponding to  $\lambda$ , together with the zero vector, is an *eigenspace*  $E_\lambda$ , and the set of all eigenvalues of a matrix  $A$  is the *spectrum* of  $A$ , denoted by  $\Lambda(A)$ .

The eigenspace  $E_\lambda$  is an *invariant subspace* under  $A$ , so that  $AE_\lambda \subseteq E_\lambda$ , and  $\dim(E_\lambda)$  is the number of linearly independent eigenvectors corresponding to the eigenvalue  $\lambda$ , known as the *geometric multiplicity* of  $\lambda$ .

We have that the eigenspace  $E_\lambda = \text{null}(\lambda I - A)$ , since  $(\lambda I - A)x = 0$ , and thus for a nonempty eigenspace  $E_\lambda$ ,  $\lambda I - A$  is a singular matrix, so that

$$\det(\lambda I - A) = 0. \quad (6.4)$$

## Characteristic polynomial

The *characteristic polynomial* of the matrix  $A \in \mathbb{C}^{n \times n}$ , is the degree  $n$  polynomial

$$p_A(z) = \det(zI - A), \quad (6.5)$$

with  $z \in \mathbb{C}$ . For  $\lambda$  an eigenvalue of  $A$ , we thus have that

$$p_A(\lambda) = 0, \quad (6.6)$$

and by the fundamental theorem of algebra we can express  $p_A(\lambda)$  as

$$p_A(\lambda) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n), \quad (6.7)$$

where each  $\lambda_j$  is an eigenvalue of  $A$ , not necessary unique. The multiplicity of each eigenvalue  $\lambda$  as a root to the equation  $p_A(\lambda) = 0$  is the *algebraic multiplicity* of  $\lambda$ , where an eigenvalue is said to be *simple* if its algebraic multiplicity is 1. The algebraic multiplicity of an eigenvalue  $\lambda$  is at least as great as its geometric multiplicity.

## Eigenvalue decompositions

A *defective matrix* is a matrix which has one or more *defective eigenvalues*, where a defective eigenvalue is an eigenvalue for which its algebraic multiplicity exceeds its geometric multiplicity.

**Theorem 5** (Eigenvalue decomposition). *Each nondefective matrix  $A \in \mathbb{C}^{n \times n}$  has an eigenvalue decomposition*

$$A = X\Lambda X^{-1}, \quad (6.8)$$

where  $X \in \mathbb{C}^{n \times n}$  is a nonsingular matrix with the eigenvectors of  $A$  as column vectors, and where  $\Lambda \in \mathbb{C}^{n \times n}$  is diagonal matrix with the eigenvalues of  $A$  on the diagonal.

We also say that a nondefective matrix is *diagonalizable*. Given that the factorization (6.8) exists, we have that

$$AX = X\Lambda, \quad (6.9)$$

which expresses (6.3) as

$$Ax_j = \lambda_j x_j, \quad (6.10)$$

with  $\lambda_j$  the  $j$ th diagonal entry of  $\Lambda$ , and  $x_j$  the  $j$ th column of  $X$ .

For some matrices eigenvectors can be chosen to be pairwise orthogonal, so that a matrix  $A$  is *unitary diagonalizable*, that is

$$A = Q\Lambda Q^*, \quad (6.11)$$

with  $Q \in \mathbb{C}^{n \times n}$  an orthogonal matrix with orthonormal eigenvectors of  $A$  as column vectors, and  $\Lambda \in \mathbb{C}^{n \times n}$  a diagonal matrix with the eigenvalues of  $A$  on the diagonal.

**Theorem 6.** *A matrix is unitary diagonalizable if and only if it is normal.*

Hermitian matrices have real eigenvalues, and thus in the particular case of a real symmetric matrix, no complex vector spaces are needed to characterize the matrix spectrum and eigenspaces.

**Theorem 7.** *An Hermitian matrix is unitary diagonalizable with real eigenvalues.*

Irrespectively if the matrix is nondefective or Hermitian, any square matrix always has a Schur factorization, with the diagonal matrix replaced by an upper triangular matrix.

**Theorem 8** (Schur factorization). *For every square matrix  $A$  there exists a Schur factorization*

$$A = QTQ^*, \quad (6.12)$$

where  $Q$  is an orthogonal matrix, and  $T$  is an upper triangular matrix with the eigenvalues of  $A$  on the diagonal.

More generally, if  $X \in \mathbb{C}^{n \times n}$  is nonsingular, the map  $A \mapsto X^{-1}AX$  is a *similarity transformation* of  $A$ , and we say that two matrices  $A$  and  $B$  are *similar* if there exists a similarity transformation such that  $B = X^{-1}AX$ .

**Theorem 9.** *Two similar matrices have the same eigenvalues with the same algebraic and geometric multiplicity.*

## 6.2 Eigenvalue algorithms

### QR algorithm for real symmetric matrices

To compute the eigenvalues of a matrix  $A$ , one may seek the roots of the characteristic polynomial. Although, it turns out that the most efficient algorithms instead are based on computing eigenvalues and eigenvectors by constructing one of the factorizations (6.8), (6.11) or (6.12).

To simplify the presentation, in this section we restrict attention to matrices that are real and symmetric, for which we now present the basic *QR algorithm*, in which a Schur factorization (6.12) of a real symmetric matrix  $A$  is constructed from successive QR factorizations.

We note that for each iteration  $A^{(k)}$  of the algorithm, we have that

$$A^{(k)} = R^{(k)}Q^{(k)} = (Q^{(k)})^{-1}A^{(k-1)}Q^{(k)}, \quad (6.13)$$

so that  $A^k$  and  $A^{k-1}$  are similar, and thus have the same eigenvalues.

---

#### Algorithm 4: QR algorithm

---

```

 $A^{(0)} = A$  for  $k = 1, 2, \dots$  do
  |  $Q^{(k)}R^{(k)} = A^{(k-1)}$ 
  |  $A^k = R^{(k)}Q^{(k)}$ 
end

```

---

The basic QR algorithm can be accelerated: (i) by Householder reflectors to reduce the initial matrix  $A^{(0)}$  to *Hessenberg form*, that is a matrix with zeros below the first subdiagonal (or in the case of an Hermitian matrix a *tridiagonal form*), and (ii) by introducing a *shift* to instead of  $A^{(k)}$  factorize  $A^{(k)} - \mu^{(k)}I$ , with  $\mu^{(k)}$  an eigenvalue estimate.

### Rayleigh quotient

We recall that for a real symmetric matrix  $A$ , all eigenvalues  $\lambda_j$  are real and the corresponding eigenvectors  $q_j$  are orthonormal. Given a vector  $x \in \mathbb{R}^n$ ,

what is the real number  $\alpha \in \mathbb{R}$  that best approximate an eigenvalue of  $A$  in the sense that  $\|Ax - \alpha x\|$  is minimized?

If  $x = q_j$  is an eigenvector of  $A$ , then  $\alpha = \lambda_j$  is the corresponding eigenvalue. If not,  $\alpha$  is the solution to the  $n \times 1$  least squares problem

$$\min_{\alpha \in \mathbb{R}} \|Ax - \alpha x\|, \quad (6.14)$$

for which the normal equations are given as

$$x^T Ax = x^T \alpha x. \quad (6.15)$$

With  $\alpha = r(x)$ , we define the *Rayleigh quotient* as

$$r(x) = \frac{x^T Ax}{x^T x}, \quad (6.16)$$

where  $r(x)$  is an approximation of an eigenvalue  $\lambda_j$ , if  $x$  is close to the eigenvector  $q_j$ . In fact,  $r(x)$  converges quadratically to  $r(q_j) = \lambda_j$ , that is

$$r(x) - r(q_j) = O(\|x - q_j\|^2), \quad (6.17)$$

as  $x \rightarrow q_j$ .

### Convergence of QR algorithm

The matrix products  $\underline{Q}^{(k)} = Q^{(1)}Q^{(2)} \dots Q^{(k)}$  and  $\underline{R}^{(k)} = R^{(k)}R^{(k-1)} \dots R^{(1)}$ , correspond to a QR factorization of the  $k$ th power of  $A$ ,

$$A^k = \underline{Q}^{(k)} \underline{R}^{(k)}, \quad (6.18)$$

so that the QR algorithm constructs successive orthonormal bases for the powers  $A^k$ . Further, the diagonal elements of the  $k$ th iterate  $A^{(k)}$  are the Rayleigh quotients of  $A$  corresponding to the column vectors of  $\underline{Q}^{(k)}$ ,

$$A^{(k)} = (\underline{Q}^{(k)})^T A \underline{Q}^{(k)}, \quad (6.19)$$

and thus the diagonal elements of  $A^{(k)}$  converges (quadratically) to the eigenvalues of  $A$ .