

Chapter 8

Nonnegative Matrices and Graph Theory

The interconnections and coupling patterns of dynamical systems are best described in terms of graph theory. This chapter serves the purpose of summarizing the main results and tools from matrix analysis and graph theory that will be important for the analysis of interconnected systems in subsequent chapters. This includes a proof of the Perron–Frobenius theorem for irreducible nonnegative matrices using a contraction mapping principle on convex cones due to Birkhoff (1957). We introduce adjacency matrices and Laplacians associated to a weighted directed graph and study their spectral properties. The analysis of eigenvalues and eigenvectors for graph adjacency matrices and Laplacians is the subject of spectral graph theory, which is briefly summarized in this chapter; see the book by Godsil and Royle (2001) for a comprehensive presentation. Explicit formulas for the eigenvalues and eigenvectors of Laplacians are derived for special types of graphs such as cycles and paths. These formulas will be used later on, in Chapter 9, in an examination of homogeneous networks. The technique of graph compression is briefly discussed owing to its relevance for the model reduction of networks. Properties of graphs are increasingly important for applications to, for example, formation control and molecular geometry. Therefore, a brief section is included on graph rigidity and the characterization of Euclidean distance matrices.

We begin by establishing some notation to be used subsequently and presenting some basic facts on Kronecker products of matrices over a field \mathbb{F} . For rectangular matrices $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{k \times l}$, the **Kronecker product** is defined as the $mk \times nl$ matrix

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}.$$

By this definition, the Kronecker product of an upper triangular matrix A with a rectangular matrix B is block-upper triangular. In particular, the Kronecker product

$B \otimes I_N$ is of the form

$$B \otimes I_N = \begin{pmatrix} b_{11}I_N & \dots & b_{1l}I_N \\ \vdots & \ddots & \vdots \\ b_{k1}I_N & \dots & b_{kl}I_N \end{pmatrix},$$

while

$$I_N \otimes A = \text{diag}(A, \dots, A) = \begin{pmatrix} A & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A \end{pmatrix}.$$

If A and B are invertible $n \times n$ and $m \times m$ matrices, respectively, then the Kronecker product $A \otimes B$ is invertible and

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

The eigenvalues of $A \otimes B$ are the products $\lambda_i(A)\lambda_j(B)$ of the eigenvalues $\lambda_i(A)$ and $\lambda_j(B)$ of A and B , respectively. Therefore, the trace and determinant of $A \otimes B$ of matrices A and B are $\text{tr}(A \otimes B) = \text{tr}(A)\text{tr}(B)$ and $\det(A \otimes B) = \det(A)^m \det(B)^n$. Similarly, the eigenvalues of $A \otimes I_m + I_n \otimes B$ are the sums $\lambda_i(A) + \lambda_j(B)$. The following rules for the Kronecker product are easily verified:

$$(A \otimes B)(C \otimes D) = AC \otimes BD,$$

$$(A \otimes B)^T = A^T \otimes B^T.$$

Let $\text{vec}(A) \in \mathbb{F}^{mn}$ denote a column vector obtained by stacking the second column of A under the first, then the third under the second, and so on. This leads to the following important identity:

$$\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B).$$

8.1 Nonnegative Matrices and Contractions

Let \mathbb{R}_+ denote the subset of all nonnegative real numbers. A matrix $A \in \mathbb{R}^{n \times n}$ is called **nonnegative (positive)** if all entries a_{ij} of A are nonnegative (positive) real numbers. The notation for nonnegative and positive matrices

$$A \geq 0 \iff a_{ij} \geq 0, \quad i, j = 1, \dots, n,$$

$$A > 0 \iff a_{ij} > 0, \quad i, j = 1, \dots, n$$

should not be confused with the notion of positive definiteness for symmetric matrices. To distinguish these notions from each other, the **positive (semi-) definiteness** of real symmetric matrices $A = A^\top \in \mathbb{R}^{n \times n}$ is denoted by

$$A \succeq 0 \iff x^\top A x \geq 0 \quad \text{for all } x \in \mathbb{R}^n,$$

$$A \succ 0 \iff x^\top A x > 0 \quad \text{for all } x \in \mathbb{R}^n \setminus \{0\}.$$

The sum and product of finitely many nonnegative matrices is nonnegative. Moreover, the scalar product λA of a nonnegative matrix A with a nonnegative real number λ is a nonnegative matrix. Thus the set of nonnegative matrices forms a closed convex cone $\mathbb{R}_+^{n \times n}$ in the matrix space $\mathbb{R}^{n \times n}$, which is multiplicatively closed.

The set of nonnegative matrices forms the largest class of matrices that leave the cone \mathbb{R}_+^n invariant. The Hilbert metric allows for a natural generalization of this situation. We refer to the papers by Birkhoff (1957), Bushell (1986), and Kohlberg and Pratt (1982) for additional background. Let $C \subset \mathbb{R}^n$ denote a closed convex cone that is **pointed**, i.e., C has nonempty interior and satisfies $C \cap (-C) = \{0\}$. We use the notation $x \geq 0$ whenever $x \in C$ and $x > 0$ whenever x is an interior point of C . Recall that the dual cone of C is defined as

$$C^* = \{\lambda \in \mathbb{R}^{1 \times n} \mid \lambda(x) \geq 0\}.$$

Clearly, the dual cone $(\mathbb{R}_+^n)^*$ of \mathbb{R}_+^n is equal to $\mathbb{R}_+^{1 \times n}$. We mention, without proof, the well-known fact that for pointed closed convex sets the interior of the dual cone C^* is nonempty. Note further that for C a closed pointed convex cone, every linear functional λ in the interior of C^* satisfies $\lambda(x) > 0$ for all nonzero $x \in C$. This implies the following lemma.

Lemma 8.1. *Let C be a closed convex and pointed cone in \mathbb{R}^n . Then the subset $C_1 = \{x \in C \mid \lambda(x) = 1\}$ is compact for all interior points λ of the dual cone C^* .*

Proof. Clearly, $C_1 \subset \mathbb{R}^n$. Suppose C_1 is unbounded. Then there exists a sequence $x_k \in C_1$, with $\|x_k\| \rightarrow \infty$ and $\lambda(x_k) = 1$ for all k . Thus $\lambda(\frac{x_k}{\|x_k\|})$ converges to 0. By the compactness of the unit sphere in \mathbb{R}^n , there exists an infinite subsequence $y_m, m \in \mathbb{N}$ of $\frac{x_k}{\|x_k\|}$ that converges to a unit vector $y \in C$. Thus $\lambda(y) = \lim_{m \rightarrow \infty} \lambda(y_m) = 0$. But λ is in the interior of C^* , and therefore $\lambda(x) > 0$ for all $x \in C \setminus \{0\}$. Thus $y = 0$, which is a contradiction. ■

A **projective metric** on C is a map $d : C \times C \rightarrow \mathbb{R} \cup \{\infty\}$ such that for all $x, y, z \in C$ (and $r \leq \infty, r + \infty = \infty = \infty + \infty$ for all real r):

1. $d(x, y) = d(y, x)$;
2. $d(x, y) \geq 0, d(x, y) = 0$ if and only if $x = \lambda y$ for some real $\lambda > 0$;
3. $d(x, z) \leq d(x, y) + d(y, z)$.

Conditions 1–3 imply the identity

$$4. \quad d(\lambda x, \mu y) = d(x, y) \text{ for all } \lambda > 0, \mu > 0.$$

Let

$$M(x, y) = \inf\{\lambda \geq 0 \mid x \leq \lambda y\},$$

$$m(x, y) = \sup\{\lambda \geq 0 \mid x \geq \lambda y\} = \frac{1}{M(y, x)}.$$

The following properties are easily seen:

$$0 \leq m(x, y) \leq M(x, y) \leq \infty,$$

$$m(x, y)y \leq x \leq M(x, y)y.$$

Definition 8.2. The **Hilbert metric** on C is a projective metric defined by

$$d(x, y) = \log \frac{M(x, y)}{m(x, y)} = \log M(x, y) + \log M(y, x).$$

Here, $d(0, 0) = 0$; $d(x, 0) = d(0, y) = \infty$ for $x, y \in C$.

The preceding definitions are illustrated by the following examples.

Example 8.3. (a) Let $C = \mathbb{R}_+^n$. Then for all $x > 0, y > 0$

$$m(x, y) = \min_{i=1, \dots, n} \frac{x_i}{y_i}, \quad M(x, y) = \max_{i=1, \dots, n} \frac{x_i}{y_i}.$$

Thus the Hilbert metric on \mathbb{R}_+^n is

$$d(x, y) = \max_{1 \leq i, j \leq n} \log \frac{x_i y_j}{x_j y_i}, \quad \text{for } x > 0, y > 0.$$

(b) Let $C = \{X \in \mathbb{R}^{n \times n} \mid X = X^T \succeq 0\}$ denote the closed convex cone of positive semidefinite real symmetric matrices. Let $\lambda_{\min}(X)$ and $\lambda_{\max}(X)$ denote the smallest and largest eigenvalues of a symmetric matrix X , respectively. For positive definite matrices $X \succ 0, Y \succ 0$ then

$$m(X, Y) = \lambda_{\min}(XY^{-1}), \quad M(X, Y) = \lambda_{\max}(XY^{-1}).$$

Thus the Hilbert metric of two positive definite matrices $X \succ 0, Y \succ 0$ is

$$d(X, Y) = \log \frac{\lambda_{\max}(XY^{-1})}{\lambda_{\min}(XY^{-1})}.$$

For the proof of the following result see Kohlberg and Pratt (1982).

Theorem 8.4. Let C be a closed convex cone in \mathbb{R}^n that is pointed. Let $\lambda \in C^*$ be an interior point of C^* . Then the following properties are satisfied:

1. *The Hilbert metric is a projective metric on C . The distance satisfies $d(x, y) < \infty$ if and only if x and y are interior points of $L(x, y) \cap C$. In particular, $d(x, y) < \infty$ for all interior points x, y of C .*
2. *Let $C_1 := \{x \in C \mid \lambda(x) = 1\}$. Then (C_1, d) is a compact metric space, and there exists a constant $\gamma > 0$ with*

$$\|x - y\| \leq \gamma d(x, y) \quad \forall x, y \in C_1.$$

3. *Let $x, y \in C$ be such that $d(x, y) < \infty$. Then there exist $a, b \in C$ such that*

$$d(x, y) = \log \frac{\|a - y\| \|x - b\|}{\|a - x\| \|y - b\|}.$$

A linear map $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is called C -monotonic whenever $AC \subset C$. Let d denote the Hilbert metric on C . Then

$$k(A) = \inf\{k \geq 0 \mid d(Ax, Ay) \leq kd(x, y) \quad \forall x, y \in C\}$$

denotes the contraction constant of A . The operator A is called a **contraction** if $k(A) < 1$. There is a beautiful formula for the contraction constant of a C -monotonic linear map A .

Theorem 8.5 (Birkhoff (1957)). *A linear C -monotonic map $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a contraction if and only if*

$$\delta = \sup\{d(Ax, Ay) \mid x, y \in C\} < \infty.$$

Whenever this is satisfied, the contraction constant is equal to

$$k(A) = \frac{e^{\delta/2} - 1}{e^{\delta/2} + 1}.$$

For a linear contraction on a closed convex and pointed cone C the Banach fixed-point theorem applies. The following result extends the classical Perron–Frobenius theorem to monotonic maps on a closed convex pointed cone.

Theorem 8.6 (Contraction Mapping Theorem). *Let $C \subset \mathbb{R}^n$ denote a closed convex and pointed cone, $\lambda \in C^*$ an interior point of the dual cone C^* , and $C_1 := \{x \in C \mid \lambda(x) = 1\}$. Let $\mu \geq 0$ be a nonnegative real number. Let $N \in \mathbb{N}$ and $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear map such that $(\mu I + A)^N$ maps C_1 into the interior of C .*

1. *Then $(\mu I + A)^N$ is a contraction on C and A has a unique eigenvector $x_* \in C_1$. The vector x_* is contained in the interior of C_1 with a positive eigenvalue $r_* > 0$.*
2. *The discrete dynamical system*

$$x_{t+1} = \frac{(\mu I + A)_t^{Nx}}{\lambda((\mu I + A)_t^{Nx})}, \quad x_t \in C_1 \quad (8.1)$$

converges to x_* from each initial condition $x_0 \in C_1$.

Proof. Let $B = (\mu I + A)^N$. By Lemma 8.1, C_1 is compact, and therefore the image $K := B(C_1)$ is a compact subset of the interior of C . Thus

$$\begin{aligned} \sup\{d(Bx, By) \mid x, y \in C\} &= \sup\{d(Bx, By) \mid x, y \in C_1\} \\ &\leq \delta(K) < \infty, \end{aligned}$$

where $\delta(K)$ denotes the diameter of the compact set K . The Birkhoff Theorem 8.5 therefore implies that B is a contraction on the complete metric space C_1 . Consider the discrete-dynamical system $x_{t+1} = f(x_t)$ on C_1 , defined by iterating the map

$$f : C_1 \longrightarrow C_1, \quad f(x) = \frac{Bx}{\lambda(Bx)}.$$

By our assumption on A , the map f is well defined and satisfies $d(f(x), f(y)) = d(Bx, By)$. Since B is a contraction on C_1 , so is f . Thus there exists $0 \leq k < 1$ with $d(f(x), f(y)) \leq kd(x, y)$ for all $x, y \in C_1$. Therefore, one can apply the Banach fixed-point theorem to f and conclude that there exists a unique fixed point $x_* \in C_1$ of f . Moreover, the dynamical system (8.1) converges to x_* from every initial point $x_0 \in C_1$. This shows that $x_* \in C_1$ is an eigenvector of B , with $Bx_* = \sigma x_*$. Since B maps C_1 into the interior of C , $\sigma x_* = Bx_*$ must be an interior point of C . But this implies that $\sigma > 0$ as well as that x_* is an interior point of C_1 . By projective invariance of the Hilbert metric,

$$\begin{aligned} d(Ax_*, x_*) &= d(A(Bx_*), Bx_*) \\ &= d(B(Ax_*), Bx_*) \leq kd(Ax_*, x_*), \end{aligned}$$

and therefore $d(Ax_*, x_*) = 0$. But this implies $Ax_* = r_* x_*$ for some $r > 0$. The result follows. \blacksquare

One can give a steepest-descent interpretation of the preceding arguments that will be useful for the proof of the Perron–Frobenius theorem. Consider the continuous function

$$R_A : K \longrightarrow \mathbb{R}_+, \quad R_A(x) = m(Ax, x). \quad (8.2)$$

It is instructive to compute this function in the special case $C = \mathbb{R}_+^n$, where for $x > 0$ one has

$$R_A(x) = \min_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}.$$

This form is reminiscent of the Rayleigh quotient function.

Proposition 8.7 (Generalized Rayleigh Quotient). *The same notation is used here as in Theorem 8.6. The function (8.2) has x_* as its unique maximum with $R_A(x_*) = r_*$. Moreover, $R_A(x_t)$, $t \geq 0$, is monotonically increasing in t and converges to $R_A(x_*) = r_*$ for each sequence of points $(x_t)_{t \geq 0}$ that are generated by the power iterations (8.1).*

Proof. We use the notation from the proof of Theorem 8.6. Note that R_A is well defined and continuous, because $K = B(C_1)$ is a compact subset of the interior of C . Since $m(cx, cy) = m(x, y)$ for $c > 0$, one obtains for all $x \in K$

$$\begin{aligned} R_A(f(x)) &= m(ABx, Bx) = \sup\{\lambda \mid ABx \geq \lambda Bx\} \\ &= \sup\{\lambda \mid B(Ax - \lambda x) \geq 0\} \\ &\geq \sup\{\lambda \mid Ax - \lambda x \geq 0\} = R_A(x). \end{aligned}$$

Since $Bv > 0$ for all nonzero vectors $v \in C$, this shows $B(Ax - \lambda x) > 0$ for all $x \in K$ that are not eigenvectors of A . This implies that $\sup\{\lambda \mid B(Ax - \lambda x) \geq 0\} > \sup\{\lambda \mid Ax - \lambda x \geq 0\}$ is a strict inequality, unless $x \in K$ is an eigenvector of A . Thus $R_A(x_{t+1}) = R_A(f(x_t)) < R_A(x_t)$, unless $x_t = f(x_t)$ is a fixed point of f . By Theorem 8.6, the eigenvector x_* is the only fixed point of f and satisfies $R_A(x_*) = m(Ax_*, x_*) = m(r_*x_*, x_*) = r_*$. This completes the proof. ■

The next result yields an explicit form for the contraction constant for positive matrices $A \in \mathbb{R}^{n \times n}$, $A > 0$. Here $C = \mathbb{R}_+^n$.

Corollary 8.8. *Every positive matrix $A \in \mathbb{R}^{n \times n}$ is a contraction with respect to the Hilbert metric on \mathbb{R}_+^n . The contraction constant is*

$$k(A) = \frac{\sqrt{\gamma} - 1}{\sqrt{\gamma} + 1},$$

where

$$\gamma = \max_{i,j,k,l} \frac{a_{ki}a_{lj}}{a_{kj}a_{li}}.$$

8.2 Perron–Frobenius Theorem

The Perron–Frobenius theorem establishes a surprising and deep connection between the spectral properties of nonnegative matrices and the properties of the associated graph. Let us begin by deriving the theorem using the contraction mapping theorem on closed convex cones. For other approaches and further details see, for example, the beautiful books by Fiedler (2008) and Sternberg (2010).

The notion of the irreducibility of nonnegative matrices plays a central role in the subsequent analysis. In Section 8.6, a simple graph-theoretic characterization of irreducibility is derived.

Definition 8.9. A matrix $A \in \mathbb{R}^{n \times n}$ is called **reducible** if either $(n = 1, A = 0)$ or $n \geq 2$ and there exists a permutation matrix $P \in \mathbb{R}^{n \times n}$, $1 \leq r \leq n - 1$, and the matrices $B \in \mathbb{R}^{r \times r}$, $C \in \mathbb{R}^{r \times (n-r)}$, $D \in \mathbb{R}^{(n-r) \times (n-r)}$, with

$$P^T A P = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}.$$

Otherwise, A is called **irreducible**.

Irreducible nonnegative matrices have useful properties.

Lemma 8.10. *Let $A \in \mathbb{R}^{n \times n}$ be a nonnegative irreducible matrix and $x \in \mathbb{R}^n$ a vector with nonnegative components. Then $Ax = 0$ implies $x = 0$.*

Proof. After a suitable permutation of the entries of x (and an induced similarity transformation on A), one may assume that $x = (\xi^T, 0)^T$, with $\xi = (x_1, \dots, x_r)^T$ and $x_1 > 0, \dots, x_r > 0$. Suppose $r \geq 1$. Partition the matrix A accordingly as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \tag{8.3}$$

with $A_{11} \in \mathbb{R}^{r \times r}$, and so forth. Thus $Ax = 0$ is equivalent to $A_{11}\xi = 0, A_{12}\xi = 0$, which implies $A_{11} = 0, A_{21} = 0$. This is a contradiction to A being irreducible. Therefore, $Ax = 0$ implies $x = 0$. ■

Using this lemma, we next prove a basic existence and uniqueness result for positive eigenvectors of nonnegative irreducible matrices. Let

$$\rho(A) = \max\{|\lambda| \mid \det(\lambda I - A) = 0\} \tag{8.4}$$

denote the **spectral radius** of matrix A . Suppose that A has exactly h eigenvalues with absolute value $\rho(A)$. Then h is called the **index** of A .

Theorem 8.11. *Let $A \in \mathbb{R}^{n \times n}$ be a nonnegative irreducible matrix. Let*

$$e = (1, \dots, 1)^T \in \mathbb{R}^n, \quad C_1 = \{x \in \mathbb{R}_+^n \mid e^T x = 1\}.$$

1. *Then A has a unique nonnegative eigenvector $x_* \in \mathbb{R}_+^n$, with $e^T x_* = 1$, called the **Perron vector**. Both the Perron vector and the associated eigenvalue are positive.*
2. *If A is positive, then the sequence of power iterates*

$$x_{t+1} = \frac{Ax_t}{e^T Ax_t}, \quad x_t \in C_1 \tag{8.5}$$

converges to x_ from each initial condition $x_0 \in C_1, x_0 > 0$.*

Proof. C_1 is a compact convex subset of the closed convex and pointed cone $C = \mathbb{R}_+^n$. By Lemma 8.10, matrix A maps $C \setminus \{0\}$ into itself. Moreover, $e^T Ax > 0$ for all $x \in C \setminus \{0\}$. Since A is irreducible, the matrix $(I + A)^{n-1}$ is positive (see the subsequent Theorem 8.26) and therefore maps C_1 into the interior of C . Thus one can apply Theorem 8.6, with $N = n - 1$ and $\mu = 1$, to conclude that A possesses a unique eigenvector x_* in C_1 . Moreover, x_* is contained in the interior of C_1 with positive eigenvalue $r > 0$. In particular, $x_* > 0$. The second part follows again from Theorem 8.6. ■

After this first step we can state and prove the Perron–Frobenius theorem.

Theorem 8.12 (Perron–Frobenius). *Let $A \in \mathbb{R}^{n \times n}$ denote an irreducible nonnegative matrix. Then the spectral radius $\rho(A)$ is a simple positive eigenvalue of A and there exists a positive eigenvector $x > 0$ for $\rho(A)$. No eigenvector corresponding to other eigenvalues of A is positive.*

Proof. Suppose, $x \in \mathbb{R}_+^n$ is a nonnegative eigenvector of A . By Theorem 8.11, then $x = x_*$, where $x_* > 0$ denotes the unique Perron vector of A ; thus $Ax_* = r_*x_*$ and $r_* > 0$. Since $\mathbb{R}x_*$ is the only eigenspace that intersects \mathbb{R}_+^n , the eigenvalue r_* has an algebraic multiplicity of one. It suffices, therefore, to show that $\rho(A)$ coincides with the eigenvalue r_* for x_* . In fact, let $\lambda \in \mathbb{C}$ be an eigenvalue of A so that $Az = \lambda z$ for some complex vector $z = (z_1, \dots, z_n)^T$. Let $|z| = (|z_1|, \dots, |z_n|) \in \mathbb{R}_+^n$ be the associated nonnegative vector of absolute values. From the triangle inequality one obtains $|\lambda||z| \leq A|z|$. Let

$$m(x, y) = \sup\{\lambda \geq 0 \mid x \geq \lambda y\}$$

be the order function for \mathbb{R}_+^n . Applying Proposition 8.7, then

$$|\lambda| = m(\lambda|z|, |z|) \leq m(A|z|, |z|) \leq r_*.$$

Thus r_* is an eigenvalue with eigenvector x_* and is equal to the spectral radius. This completes the proof. ■

The following perturbation result is of independent interest.

Proposition 8.13 (Wielandt). *Let $A \in \mathbb{R}^{n \times n}$ be an irreducible nonnegative matrix. Let $B \in \mathbb{C}^{n \times n}$ be a complex matrix with*

$$|b_{ij}| \leq a_{ij} \quad \text{for all } i, j = 1, \dots, n.$$

Then $\rho(B) \leq \rho(A)$. If $\rho(B) = \rho(A)$ and $\rho(B)e^{\phi\sqrt{-1}}$ is an eigenvalue of B , then there exists a diagonal matrix $D = \text{diag}(z_1, \dots, z_n) \in \mathbb{C}^{n \times n}$ with $|z_1| = \dots = |z_n| = 1$ such that

$$B = e^{\phi\sqrt{-1}}DAD^{-1}.$$

In particular, $|b_{ij}| = a_{ij}$ for all $i, j = 1, \dots, n$.

Proof. In the preceding proof of the Perron–Frobenius theorem it was shown that the maximum $r_* = R_A(x_*)$ of the function $R_A : \mathbb{R}_+^n \rightarrow \mathbb{R} \cup \{\infty\}$, $R_A(x) = m(Ax, x)$ exists and $r_* = \rho(A)$. Let $z \in \mathbb{C}^n$, $z \neq 0$, and $\lambda \in \mathbb{C}$, with $Bz = \lambda z$. Then, using the triangle inequality, one obtains $|\lambda||z| = |Bz| \leq A|z|$, and therefore

$$|\lambda| = m(|Bz|, |z|) \leq m(A|z|, |z|) \leq r_* = \rho(A). \quad (8.6)$$

This shows that $\rho(B) \leq \rho(A)$. Assume that $\rho(B) = \rho(A) = r_*$ and λ is an eigenvalue of B , with $|\lambda| = r_*$. Then (8.6) implies that $r_* = |\lambda| = m(A|z|, |z|)$. The Perron–Frobenius theorem therefore implies $A|z| = r_*|z|$ and $|z| > 0$. Similarly, for $|B| = (|b_{ij}|)$ one obtains

$$r_*|z| = |\lambda z| = |Bz| \leq |B| \cdot |z| \leq A|z| = r_*|z|,$$

and therefore $|B| \cdot |z| = A \cdot |z|$. Since $A - |B|$ is a nonnegative matrix and $|z| > 0$, this implies $A = |B|$.

Define

$$D = \text{diag} \left(\frac{z_1}{|z_1|}, \dots, \frac{z_n}{|z_n|} \right).$$

Then $D|z| = z$ and $BD|z| = Bz = \lambda D|z|$. Thus $C := e^{\phi\sqrt{-1}}D^{-1}AD$ satisfies $C|z| = A|z|$ and $|C| = |B| = A$. Split the complex matrix $C = \text{Re}C + \sqrt{-1}\text{Im}C$ into real and imaginary parts $\text{Re}C$ and $\text{Im}C$, respectively. Since A is real, $C|z| = A|z|$ implies $\text{Re}C|z| = A|z|$. From $\text{Re}C \leq |C| = A$ it follows that $A - \text{Re}C$ is nonnegative, with $(A - \text{Re}C)|z| = 0$. Since $|z| > 0$, this implies $|C| = A = \text{Re}C$, and therefore $C = A$. This completes the proof. ■

Let $A \in \mathbb{R}^{n \times n}$ be nonnegative and irreducible. For $i = 1, \dots, n$, the **period** $p(i)$ is defined as the greatest common divisor of all $m \in \mathbb{N}$ satisfying $(A^m)_{ii} > 0$. By a theorem of Romanovsky, $p(1) = \dots = p(n)$ for all irreducible nonnegative matrices A . The common value $p(A) := p(1) = \dots = p(n)$ is called the **period** of A . A nonnegative matrix A of period 1 is called **aperiodic**. We now state, without providing full proof details, a full characterization of the structure of irreducible nonnegative matrices. A stronger form of the subsequent result and its proof appeared as Theorem 4.3.1 in the book by Fiedler (2008).

Theorem 8.14. *Let $A \in \mathbb{R}^{n \times n}$ be irreducible and nonnegative, and let $\lambda_0, \dots, \lambda_{k-1}$ denote the eigenvalues of A with absolute value equal to the spectral radius $\rho(A)$. The following statements are true:*

1. $\lambda_0, \dots, \lambda_{k-1}$ are simple eigenvalues of A and satisfy

$$\lambda_j = e^{\frac{2\pi\sqrt{-1}j}{k}} \rho(A), \quad j = 0, \dots, k-1.$$

2. The spectrum of A is invariant under rotations with angle $\frac{2\pi}{k}$.

3. If $k > 1$, then there exists a permutation matrix P such that

$$PAP^T = \begin{pmatrix} 0 & B_{12} & 0 & \dots & 0 \\ 0 & 0 & B_{23} & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 0 & B_{k-1,k} \\ B_{k1} & 0 & \dots & 0 & 0 \end{pmatrix},$$

with block matrices B_{ij} of suitable sizes.

4. The index k of A in (1) coincides with the period of A .

5. A is primitive if and only if the spectral radius $\rho(A)$ is the only eigenvalue λ of A with $|\lambda| = \rho(A)$.

Proof. Only the first two statements are shown; we refer the reader to Fiedler (2008) for a proof of the remaining claims. Let $\lambda_j = \rho e^{2\pi\phi_j\sqrt{-1}}$, $\phi_0 = 0$, denote the eigenvalues of absolute value $\rho = \rho(A)$. Applying Proposition 8.13 with $B = A$ one obtains

$$A = e^{2\pi\phi_j\sqrt{-1}} D_j A D_j^{-1}, \quad j = 0, \dots, k-1$$

for suitable unitary diagonal matrices D_0, \dots, D_{k-1} . Thus the spectrum of A is invariant under multiplications by $e^{2\pi\phi_j\sqrt{-1}}$, $j = 0, \dots, k-1$. Since the spectral radius $\lambda_0 = \rho$ is a simple eigenvalue of A , $\lambda_0, \dots, \lambda_{k-1}$ are simple eigenvalues of $e^{2\pi\phi_j\sqrt{-1}} D_j A D_j^{-1} = A$. For all $0 \leq r, s \leq k-1$ and $D_{rs} = D_r D_s$,

$$\begin{aligned} A &= e^{2\pi\phi_r\sqrt{-1}} D_r A D_r^{-1} = e^{2\pi\phi_r\sqrt{-1}} e^{2\pi\phi_s\sqrt{-1}} D_r D_s A D_r^{-1} D_s^{-1} \\ &= e^{2\pi(\phi_r+\phi_s)\sqrt{-1}} D_{rs} A D_{rs}^{-1}. \end{aligned}$$

Thus $e^{2\pi(\phi_r+\phi_s)\sqrt{-1}} \rho$ are eigenvalues of A for all $0 \leq r, s \leq k-1$. This implies that $\{1, e^{2\pi\phi_1\sqrt{-1}}, \dots, e^{2\pi\phi_{k-1}\sqrt{-1}}\}$ is a multiplicative subgroup of $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ of order k . Thus

$$\lambda_j = e^{\frac{2\pi j\sqrt{-1}}{k}}, \quad j = 0, \dots, k-1.$$

In particular, the spectrum of A is invariant under rotations by $\frac{2\pi}{k}$. This completes the proof for the first two items. ■

The preceding result allows for an interesting dynamical interpretation of irreducible nonnegative matrices A in terms of discrete-time periodic linear systems. In fact, if A is not primitive, then the dynamical system $x_{t+1} = Ax_t$ is permutation equivalent to a **periodic** time-varying system $x_{t+1} = A_{[t]}x_t$, with local states $x_t \in \mathbb{R}^{n_t}$ and a periodic sequence of matrices $A_{[0]} = B_{12}, A_{[1]} = B_{23}, \dots, A_{[k-1]} = B_{k1}, A_{[k]} = A_{[0]}$, and so forth.

8.3 Stochastic Matrices and Markov Chains

We present a simple version of the ergodic theorem in the context of finite-dimensional matrix algebras. Recall that a norm $\|\cdot\|$ on the matrix space $\mathbb{C}^{n \times n}$ is called submultiplicative if $\|AB\| \leq \|A\|\|B\|$ for all matrices $A, B \in \mathbb{C}^{n \times n}$. Standard examples of submultiplicative matrix norms include the 1-norm $\|A\|_1 = \sum_{i,j=1}^n |a_{ij}|$, the Frobenius norm $\|A\|_F = \sqrt{\sum_{i,j=1}^n |a_{ij}|^2}$, and the operator norm $\|A\| = \sup_{\|x\|=1} \|Ax\|$. Let $\rho(A)$ denote the spectral radius of A .

Proposition 8.15. *Let $\rho(A)$ denote the spectral radius of a matrix $A \in \mathbb{C}^{n \times n}$. Then the following assertions are true:*

1. *If $\rho(A) < 1$, then*

$$\lim_{k \rightarrow \infty} A^k = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} A^i = 0.$$

2. *If $\rho(A) > 1$, then both sequences (A^k) and $\frac{1}{k} \sum_{i=0}^{k-1} A^i$ diverge.*
 3. *Let $\rho(A) = 1$. The limit $\lim_{k \rightarrow \infty} A^k$ exists if and only if 1 is the only eigenvalue of A with absolute value 1 and all Jordan blocks for 1 are 1×1 .*
 4. *Let $\rho(A) = 1$. The limit $\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} A^i$ exists if and only if all eigenvalues λ of A with absolute value $|\lambda| = 1$ have a geometric multiplicity of one.*

Proof. The simple proofs of assertions 1–3 are omitted. To prove assertion 4, assume that the limit $\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} A^i$ exists. From assertion it follows that each eigenvalue λ of A must satisfy $|\lambda| \leq 1$. Suppose $|\lambda| = 1$. Without loss of generality, one can assume that $A = \lambda I + N$ is a Jordan block. Then

$$\frac{1}{k} \sum_{i=0}^{k-1} A^i = \frac{1}{k} \sum_{i=0}^{k-1} \lambda^i I + \frac{1}{k} \sum_{i=0}^{k-1} i \lambda^{i-1} N + \dots + \frac{1}{k} \sum_{i=0}^{k-1} N^i$$

diverges whenever

$$\frac{1}{k} \sum_{i=0}^{k-1} i \lambda^{i-1} = \begin{cases} \frac{k\lambda^{k-1}(\lambda-1) - (\lambda^{k-1})}{k(\lambda-1)^2} & \text{if } \lambda \neq 1 \\ \frac{k-1}{2} & \text{if } \lambda = 1 \end{cases}$$

diverges. This completes the proof. ■

Theorem 8.16 (Ergodic Theorem). *Let $\|\cdot\|$ be a submultiplicative matrix norm on $\mathbb{C}^{n \times n}$ and $A \in \mathbb{C}^{n \times n}$ with $\|A\| \leq 1$. Then:*

1. *The limit*

$$P = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} A^i$$

exists and satisfies $P^2 = P = PA = AP$. Moreover,

$$\begin{aligned} \text{Im } P &= \text{Ker } (A - I) \\ \text{Ker } P &= \bigoplus_{\lambda \neq 1} \text{Ker } (A - \lambda I)^n; \end{aligned}$$

2. If 1 is the only eigenvalue of A with an absolute value of one, then

$$P = \lim_{k \rightarrow \infty} A^k$$

exists.

Proof. Assume that the limit

$$P = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} A^i \tag{8.7}$$

exists. Proposition 8.15 then implies $\rho(A) \leq 1$. For complex numbers λ with $|\lambda| \leq 1$,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} \lambda^i = \lim_{k \rightarrow \infty} \frac{1}{k} \frac{\lambda^k - 1}{\lambda - 1} = \begin{cases} 0 & \text{if } \lambda \neq 1, \\ 1 & \text{if } \lambda = 1. \end{cases}$$

Thus, by decomposing the Jordan canonical form $J = \text{diag}(J_1, J_2)$ of A , where J_1 and J_2 have eigenvalues $\lambda = 1$ and $|\lambda| < 1$, respectively, one obtains $\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} J_1^i = I$ and $\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} J_2^i = 0$. Thus $P = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} A^i$ is a projection operator that commutes with A . Therefore, $P^2 = P = AP = PA$. The preceding argument also implies the formulas for the image space $\text{Im } P$ and kernel of P . The second claim follows from Proposition 8.15. Thus it remains to show that the limit (8.7) exists.

For all complex matrices and submultiplicative norms the inequality $\rho(A) \leq \|A\|$ is valid. By Proposition 8.15, it is enough to show that the Jordan blocks for the eigenvalues λ with $|\lambda| = 1$ are 1×1 . Assume that $A = SJS^{-1}$, where $J = \text{diag}(J_1, \dots, J_r)$ are in Jordan canonical form. Let J_1, \dots, J_ν be the Jordan blocks for the eigenvalues λ with an absolute value of one. For all $m \in \mathbb{N}$, therefore, $\|J^m\| \leq \|S\| \|S^{-1}\| \|A^m\| \leq \|S\| \|S^{-1}\|$. Since all norms on $\mathbb{C}^{n \times n}$ are equivalent, one obtains, for $i = 1, \dots, \nu$,

$$\|J_i^m\|_1 \leq \|J^m\|_1 \leq \gamma \|J^m\| \leq \gamma \|S\| \|S^{-1}\|.$$

For every Jordan block $J_i = \lambda I + N, i = 1, \dots, \nu$, of size s and $m \geq 1$, one has the estimate

$$\|J_i^m\|_1 = \|(\lambda I + N)^m\|_1 \geq \begin{cases} 1 & s = 1, \\ m + 1 & s > 1. \end{cases}$$

Thus the sequence $(\|J^m\|_1)_m$ grows unbounded if there exists an eigenvalue λ of geometric multiplicity greater than one and $|\lambda| = 1$. This completes the proof. ■

We establish a concrete form of the ergodic theorem for doubly stochastic matrices.

Definition 8.17. A nonnegative matrix $A \in \mathbb{R}^{n \times n}$ is called **stochastic** if

$$\sum_{j=1}^n a_{ij} = 1, \quad i = 1, \dots, n.$$

A is called **doubly stochastic** if it is nonnegative and satisfies

$$\sum_{j=1}^n a_{ij} = 1, \quad \sum_{l=1}^n a_{lj} = 1, \quad i, l = 1, \dots, n. \quad (8.8)$$

Let

$$\mathbf{e}_n = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

denote a vector in \mathbb{R}^n with all components equal to one. Then a nonnegative matrix A is stochastic (or doubly stochastic) if and only if $A\mathbf{e}_n = \mathbf{e}_n$ (or $A\mathbf{e}_n = \mathbf{e}_n, \mathbf{e}_n^\top A = \mathbf{e}_n^\top$).

Theorem 8.18. Let $A \in \mathbb{R}^{n \times n}$ be a stochastic matrix. Then

1. 1 is an eigenvalue of A and A has a spectral radius equal to 1;
2. The limit

$$P = \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} A^i$$

exists and is a stochastic matrix that satisfies $P^2 = P = PA = AP$;

3. If 1 is the only eigenvalue of A with an absolute value of one, then

$$P = \lim_{k \rightarrow \infty} A^k$$

exists. In particular, this is case if A is primitive;

4. If A is irreducible, then

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} A^i = \mathbf{e}_n \mathbf{y}^\top.$$

Here \mathbf{y} is a uniquely determined positive vector with

$$\mathbf{y}^\top A = \mathbf{y}^\top, \quad y_1 + \dots + y_n = 1.$$

Proof. The matrix norm $\|A\| = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$ is submultiplicative, and every stochastic matrix A satisfies $\|A\| = 1$. The vector \mathbf{e}_n is an eigenvector of A with eigenvalue 1. Therefore, the spectral radius satisfies $1 \leq \rho(A) \leq \|A\| = 1$, i.e.,

$$\rho(A) = \|A\| = 1.$$

This proves claim 1. Moreover, claim 2 follows from the Ergodic Theorem 8.16, together with the observation that A being stochastic implies that for each k the matrix $\frac{1}{k} \sum_{i=0}^{k-1} A^i$ is nonnegative, with $\frac{1}{k} \sum_{i=0}^{k-1} A^i \mathbf{e}_n = \frac{1}{k} \sum_{i=0}^{k-1} \mathbf{e}_n = \mathbf{e}_n$. The first claim in statement 3 follows from Theorem 8.16, while the second claim follows from the Perron–Frobenius theorem. To prove the last claim, one applies the Perron–Frobenius theorem to the irreducible matrix A^\top . Thus, there exists a unique vector $y \in \mathbb{R}^n, y > 0$, with $y^\top A = y^\top$ and $\mathbf{e}^\top y = 1$. Moreover, 1 is a simple eigenvalue of A . Thus Theorem 8.16 implies that $\text{rk} P = \dim \text{Ker}(A - I) = 1$, and therefore P is of the form $P = bc^\top$ for unique nonzero vectors $b, c \in \mathbb{R}^n$, with $c^\top \mathbf{1} = 1$. Since $P\mathbf{e} = \mathbf{e}$, $b = \mathbf{e}$. Moreover, $y^\top A = y^\top$ implies $y^\top P = y^\top$, and therefore $y^\top = y^\top P = y^\top \mathbf{e} c^\top = c^\top$, since $y^\top \mathbf{e} = 1$. This completes the proof. ■

Corollary 8.19. For every irreducible, doubly stochastic matrix $A \in \mathbb{R}^{n \times n}$,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} A^i = \frac{1}{n} \mathbf{e} \mathbf{e}^\top.$$

It is straightforward to apply the preceding results to Markov chains. Consider random variables with values in the finite alphabet $\{1, \dots, n\}$ and associated probabilities $\pi_i = P(X = i), i = 1, \dots, n$. Thus $\pi = (\pi(1), \dots, \pi(n))^\top \in \mathbb{R}^n$ satisfies $\pi \geq 0$ and $\pi(1) + \dots + \pi(n) = 1$. One can easily generalize this simple static model to a dynamic one by considering a stochastic process (X_t) defined by a sequence of random variables $X_t, t \in \mathbb{N}_0$, with values in $\{1, \dots, n\}$. Let $(\pi_t) \in \mathbb{R}^n$ denote the associated vector of probabilities. Markov chains are special stochastic processes where for each time t the vector of probabilities π_t depends only on π_{t-1} . More precisely, it is assumed that for each $i, j \in \{1, \dots, n\}$ the conditioned probabilities

$$p_{ij} := P(X_{t+1} = j | X_t = i), \quad t \in \mathbb{N}_0,$$

are independent of time t . Therefore, one can describe the transition probabilities between states i and j by a matrix

$$A = (p_{ij}) \in \mathbb{R}^{n \times n}$$

of real numbers $p_{ij} \geq 0$, with $\sum_{j=1}^n p_{ij} = 1$ for $i = 1, \dots, n$. Thus $A \in \mathbb{R}^{n \times n}$ is a stochastic matrix of transition probabilities, with $A\mathbf{e} = \mathbf{e}$.

Definition 8.20. A **Markov chain** on the finite state space $\{1, \dots, n\}$ is a discrete dynamical system

$$\pi_{t+1}^\top = \pi_t^\top A, \quad \pi_0 = \pi, \quad t \in \mathbb{N}_0 \quad (8.9)$$

defined by a stochastic matrix A . Here the initial probability distribution π is allowed to be an arbitrary vector of nonnegative numbers p_1, \dots, p_n , $p_1 + \dots + p_n = 1$.

The preceding results on stochastic matrices can be reformulated as follows.

Theorem 8.21. *Let (A, π) be a Markov chain on $\{1, \dots, n\}$ with initial probability distribution π . Let $\pi_t^\top = (\pi_t(1), \dots, \pi_t(n))$ denote the probability distributions that evolve according to the Markov chain (8.9).*

1. *If A is irreducible, then there exists a unique stationary probability distribution $\pi_\infty^\top = (\pi_\infty(1), \dots, \pi_\infty(n)) \in \mathbb{R}^{1 \times n}$ satisfying*

$$\pi_\infty > 0, \quad \pi_\infty^\top A = \pi_\infty^\top, \quad \mathbf{e}^\top \pi_\infty = 1.$$

Moreover,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^{k-1} A^i = \mathbf{e} \pi_\infty^\top,$$

which implies that

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{t=0}^{k-1} \pi_t^\top = \pi_\infty^\top.$$

2. *Assume A is primitive, i.e., $A^m > 0$ for some $m \in \mathbb{N}$. Then the following limits exist:*

$$\begin{aligned} \lim_{k \rightarrow \infty} A^k &= \mathbf{e} \pi_\infty^\top, \\ \lim_{t \rightarrow \infty} E(X_t) &= \sum_{i=1}^n i \pi_\infty(i). \end{aligned}$$

Here the expectation value of X_t is defined as $E(X_t) := \sum_{i=1}^n i \pi_t(i)$.

Example 8.22. We discuss the Ehrenfest diffusion model from statistical mechanics. Assume a domain Ω is partitioned into two regions, Ω_1 and Ω_2 . Assume further that Ω contains exactly n particles that may move around in Ω , passing from one region to the other. Let $X_t \in \{0, \dots, n\}$ denote the number of particles that are in region Ω_1 at time t . Assume that the probability for a change of a particle from region Ω_1 to region Ω_2 , or vice versa, is exactly $\frac{1}{n}$. The transition probability matrix, then, is the $(n+1) \times (n+1)$ tridiagonal matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ \frac{1}{n} & 0 & \frac{n-1}{n} & & & \vdots \\ 0 & \frac{2}{n} & 0 & \frac{n-2}{n} & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & \frac{1}{n} \\ 0 & \cdots & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Note that A is an irreducible stochastic matrix. Therefore, a unique stationary probability distribution π_∞ of X_t exists and satisfies $\pi_\infty^\top A = \pi_\infty^\top$. Define $y = (y_0, \dots, y_n)^\top$, with $y_j = 2^{-j} \binom{n}{j}$. A straightforward computation shows that $y^\top A = y^\top, y^\top \mathbf{e} = 1$. Therefore, the stationary probabilities are

$$\pi_\infty(j) = \frac{1}{2^j} \binom{n}{j}, \quad j = 0, \dots, n.$$

In particular, the expectation value of the number of particles in region Ω_1 is equal to

$$E(X_t) = \sum_{j=0}^n j \pi_\infty(j) = \sum_{j=0}^n j \frac{1}{2^j} \binom{n}{j} = \frac{n}{2},$$

as expected. One can show that the eigenvalues of A are the real numbers $1 - \frac{2k}{n}$, $k = 0, \dots, n$. The convergence rate of the Markov chain is dependent on the second largest eigenvalue of A , i.e., it is equal to $1 - \frac{2}{n}$. Thus, for large numbers n of particles, the Markov chain will converge quite slowly to the equilibrium distribution.

8.4 Graphs and Matrices

We now introduce some of the basic notions from graph theory and relate the graph concepts to the structure of nonnegative matrices. A **directed graph (digraph)** $\Gamma = (V, E)$ consists of a finite set $V = \{v_1, \dots, v_N\}$ of **vertices**, together with a finite subset $E \subset V \times V$ of pairs of vertices called **edges**. Thus each edge of a graph is a pair (v, w) of vertices v and w , which are called the initial and terminal vertices of e , respectively. This leads to well-defined maps $\iota, \tau : E \rightarrow V$ that assign to each edge the initial vertex $\iota(v, w) = v$ and terminal vertex $\tau(v, w) = w$, respectively. We refer to the pair ι, τ as the **canonical orientation** on a digraph (Figures 8.1 and 8.2).

Each vertex element v in a digraph has two kinds of neighborhoods, the **in-neighborhood**,

$$\mathcal{N}^i(v) = \{u \in V \mid (u, v) \in E\},$$

Fig. 8.1 Directed graph

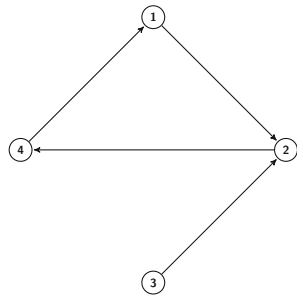
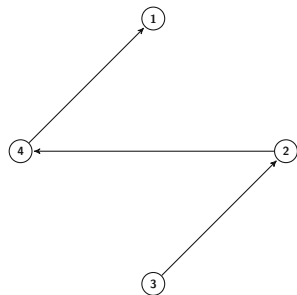


Fig. 8.2 Spanning tree



and the **out-neighborhood**,

$$\mathcal{N}^o(v) = \{w \in V \mid (v, w) \in E\}.$$

The cardinalities $d_i(v) = |\mathcal{N}^i(v)|$ and $d_o(v) = |\mathcal{N}^o(v)|$ are called the **in-degree** and **out-degree** of v , respectively. A **subgraph** of a digraph $\Gamma = (V, E)$ is a digraph $\Gamma' = (V', E')$, with $V' \subset V$ and $E' \subset E$. It is called a **spanning subgraph** if $V' = V$ and $E' \subset E$. An **induced subgraph** of $\Gamma = (V, E)$ is a subgraph $\Gamma' = (V', E')$ that contains all edges in E between pairs of vertices in V' . A **walk** in a directed graph Γ of length $r - 1$ is a finite sequence of vertices (v_1, \dots, v_r) such that (v_i, v_{i+1}) are edges for $i = 1, \dots, r - 1$. A walk is **cyclic** if $v_1 = v_r$. A path is a walk where all vertices v_1, \dots, v_r are distinct. Thus a path cannot be cyclic. A directed graph is called **acyclic** if it does not contain a cycle.

An important topological concept in graph theory is that of connectivity. A digraph Γ is called **strongly connected** if there exists a directed path between all pairs $(u, v) \in V \times V$ of distinct vertices. Γ is called **connected** if, for each $(u, v) \in V \times V$, there exists a directed path from u to v or from v to u . A **strong component** of a digraph is a maximal, strongly connected induced subgraph (Figure 8.3).

For a proof of the following characterization of strong components we refer the reader to Fiedler (2008).

Fig. 8.3 Strongly connected graph

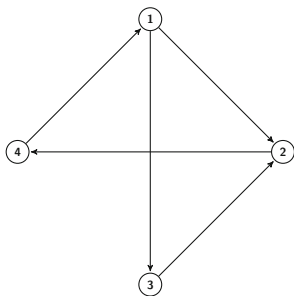
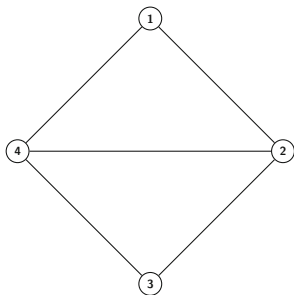


Fig. 8.4 Undirected graph



Proposition 8.23. Let $\Gamma = (V, E)$ be a digraph.

1. Every vertex v is contained in a unique strong component.
2. Two distinct strong components have disjoint sets of vertices.
3. Two distinct vertices $u, v \in V$ belong to the same strong component if and only if there are directed paths from u to v and from v to u .

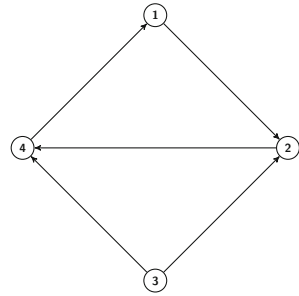
An **undirected graph** $\Gamma = (V, E)$, often simply called a **graph**, consists of a finite set $V = \{v_1, \dots, v_N\}$ of **vertices**, together with a finite set $E = \{\{v_i, v_j\} \mid (i, j) \in I\}$ of **edges**. Here I denotes a finite subset of $\{1, \dots, N\} \times \{1, \dots, N\}$. Thus the edges of an undirected graph are unordered pairs of vertices. Frequently, self-loops are excluded from the definition. A graph Γ is **oriented** if there exist maps $\iota, \tau : E \rightarrow V$ such that for each edge, $e = \{\iota(e), \tau(e)\}$. Thus, $\hat{\Gamma} = (V, E)$ is a directed graph with a set of edges $E = \{\{\iota(e), \tau(e)\} \mid e \in E\}$. In many situations concerning graphs one often assumes that an underlying orientation of the graph is specified. A directed graph $\Gamma = (V, E)$ carries a natural orientation by defining $\iota(v, w) = v, \tau(v, w) = w$ for all vertices (Figures 8.4 and 8.5).

We briefly mention a number of elementary operations one can perform with graphs. Let $\Gamma = (V, E)$ and $\Gamma' = (V', E')$ denote graphs with a disjoint set of vertices, i.e., $V \cap V' = \emptyset$. Then the following operations yield new graphs:

1. **Union:** $\Gamma \cup \Gamma' := (V \cup V', E \cup E')$.
2. **Join:** $\Gamma + \Gamma' := (V \cup V', \langle E \cup E' \rangle)$, where

$$\langle E \cup E' \rangle := E \cup E' \cup \{\{v, v'\} \mid v \in V, v' \in V'\}.$$

Fig. 8.5 Orientation of a graph



3. **Product:** $\Gamma \times \Gamma' := (V \times V', \hat{E})$, where the set of edges \hat{E} is defined as

$$\{(v, v'), (w, w')\} \in \hat{E} \iff \begin{cases} \{v, w\} \in E \text{ and } v' = w' \\ \text{or} \\ \{v', w'\} \in E' \text{ and } v = w. \end{cases}$$

4. **Insertion of a vertex into an edge:** Geometrically this means one places a new vertex w in the middle of an edge uv and replaces the old edge uv with the two new ones uw, wv . This operation does not change the topology of the graph. Thus, for every edge $\{u, v\} \in E$, a new graph $\Gamma' = (V', E')$ is defined as

$$\begin{aligned} V' &= V \cup \{u, v\} \\ E' &= E \setminus \{\{u, v\}\} \cup \{\{u, w\}, \{w, v\}\}. \end{aligned}$$

5. **Contraction of an edge:** Here one replaces an edge $\{u, v\}$ with a new vertex w and adds new edges to all neighboring vertices of u, v .

The preceding definitions and constructions for digraphs carry over to undirected graphs in an obvious way. For a vertex $v \in V$, let

$$N(v) = \{w \in V \mid \{v, w\} \in E\}$$

denote the neighborhood of v in the graph. The degree of v is the number of all neighboring vertices, i.e., it is equal to $|N(v)|$. A graph is called k -regular if all vertices have the same degree k . A subgraph of Γ is defined by a pair $\Gamma' = (V', E')$ such that $V' \subset V$ and $E' \subset E$. A **spanning subgraph** of Γ is a subgraph $\Gamma' = (V', E')$ with the same set of vertices $V' = V$. A **path** in Γ of length $r - 1$ is a finite sequence of vertices (v_0, \dots, v_r) such that $e_i = \{v_{i-1}, v_i\}$ are edges of Γ for $i = 1, \dots, r$. One says that the path connects the vertices v_0 and v_r . If $v_r = v_0$, then the path is called closed or a **cycle**. A graph is called **connected** if two distinct vertices $v \neq w$ are always connected through a suitable path in Γ . A maximal, connected, induced subgraph of a graph is called a **connected component**. The counterpart to Proposition 8.23 is true, too, i.e., each vertex is contained in a unique connected

component, and the connected components of a graph form a disjoint decomposition of the set of vertices. A **tree** is a connected graph Γ without cycles. This is easily seen to be equivalent to $|V| = |E| + 1$. A **forest** is a graph whose connected components are trees. A **spanning tree** in a graph $\Gamma = (V, E)$ is a spanning subgraph $\Gamma' = (V', E')$, which is a tree. The number of spanning trees in a graph can be counted by the so-called Matrix-Tree Theorem 8.43.

Weighted Digraphs and Matrices. Nonnegative matrices are associated with digraphs in various ways. A digraph $\Gamma = (V, E)$ is called **weighted** if for each edge $(v_i, v_j) \in E$ one specifies a nonzero real number $a_{ij} \in \mathbb{R}$. For $(v_i, v_j) \notin E$ set $a_{ij} = 0$. Thus, using a labeling $\{1, \dots, N\} \rightarrow V$, one associates with the graph a real $N \times N$ matrix $A(\Gamma) = (a_{ij}) \in \mathbb{R}^{N \times N}$. We refer to $A(\Gamma)$ as the **weighted adjacency matrix**. The labelings of the set of vertices differ from each other by a permutation π on $\{1, \dots, N\}$. Thus the associated adjacency matrix changes by a similarity transformation $\pi A(\Gamma) \pi^{-1}$. Conversely, if A denotes a real $N \times N$ matrix, then let $\Gamma_A = (V_A, E_A)$ denote the associated finite directed graph with vertex set $V_A = \{1, \dots, N\}$. A pair $(i, j) \in V_A \times V_A$ is an edge of Γ_A if and only if $a_{ij} \neq 0$. Then A is the weighted adjacency matrix of Γ_A . Similarly, weighted undirected graphs are defined by specifying for each edge $\{v_i, v_j\}$ a real number a_{ij} and $a_{ij} = 0$ for $\{v_i, v_j\} \notin E$. Thus the weight matrix $A = (a_{ij})$ of an undirected graph is always a real symmetric matrix and therefore has only real eigenvalues.

Every digraph can be considered in a canonical way as a weighted digraph by defining the weight matrix with 0, 1 entries as

$$\mathfrak{A} = (a_{ij}) \in \{0, 1\}^{N \times N}, \quad \text{with} \quad a_{ij} = \begin{cases} 1 & (v_i, v_j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

We refer to the digraph $\Gamma_{\mathfrak{A}}$ as a **canonically weighted** digraph.

Example 8.24. A simple example of digraphs with nonnegative weights arises in Euclidean distance geometry and shape analysis. Thus, consider an arbitrary directed graph $\Gamma = (V, E)$ with vertex set $V = \{v_1, \dots, v_N\} \subset \mathbb{R}^m$. Using the Euclidean distance $\|v - w\|$ between two vertices, define the weights as $a_{ij} = \|v_i - v_j\|$ if and only if $(v_i, v_j) \in E$, and $a_{ij} = 0$ otherwise. Then the weighted adjacency matrix contains all the mutual distances between ordered pairs of points (v_i, v_j) that are specified by the edges of the graph. Thus this matrix contains very interesting information on the geometric configuration of the vertex points.

One can express the classical adjacency matrices of a graph in terms of basic graph operations. Let Γ and Γ' be graphs on m and n vertices, respectively. The classical adjacency matrices for unions, sums, and products are

$$\begin{aligned} \mathfrak{A}_{\Gamma \cup \Gamma'} &= \text{diag}(\mathfrak{A}_{\Gamma}, \mathfrak{A}_{\Gamma'}), & \mathfrak{A}_{\Gamma \times \Gamma'} &= \mathfrak{A}_{\Gamma} \otimes I_n + I_m \otimes \mathfrak{A}_{\Gamma'}, \\ \mathfrak{A}_{\Gamma + \Gamma'} &= \begin{pmatrix} \mathfrak{A}_{\Gamma} & J \\ J^{\top} & \mathfrak{A}_{\Gamma'} \end{pmatrix}. \end{aligned}$$

Here J denotes the $m \times n$ matrix with all entries equal to 1.

The analysis of the structural properties of matrices is closely related to graph theory. The basic connectivity properties of digraphs are reflected in the associated properties of the weighted adjacency matrix. There is a simple graph-theoretic characterization of irreducible matrices.

Proposition 8.25. *The following conditions are equivalent for a matrix $A \in \mathbb{R}^{N \times N}$:*

1. A is irreducible.
2. The digraph Γ_A is strongly connected.

Proof. Assume that Γ_A is not strongly connected. Then there exist vertices $w \neq j \in V_A$ such that there exists no directed path from w to j . Let $V' \subset V_A$ denote the set of vertices v such that there exists a directed path from v to j . Define $V_1 = V' \cup \{j\}$ and $V_2 = V_A \setminus V_1$. By construction of these sets, there does not exist a path from a vertex w in V_2 to some vertex v in V_1 ; otherwise, one could concatenate the path from w to v with the path from v to j to obtain a path from w to j . This is a contradiction, since $V_2 \cap V_1 = \emptyset$. By assumption on w, j , one has $j \in V_1 \neq \emptyset, w \in V_2$. After a suitable renumbering of vertices one can assume, without loss of generality, that $V_1 = \{1, \dots, r\}, V_2 = \{r+1, \dots, N\}, 1 \leq r \leq N-1$. Thus there exists a permutation matrix P such that

$$P^T A P = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}.$$

Therefore, A is reducible. Conversely, if A is reducible, then there exists no path from the set $V_2 = \{r+1, \dots, N\}$ to $V_1 = \{1, \dots, r\}$. Thus the graph is not strongly connected. This completes the proof. ■

For nonnegative matrices a stronger form of Proposition 8.25 is valid.

Theorem 8.26. *The following conditions are equivalent for a nonnegative $A \in \mathbb{R}^{N \times N}$.*

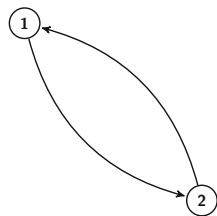
1. A is irreducible.
2. For every pair of indices $i, j \in \{1, \dots, N\}$ there exists $m \in \mathbb{N}$, with $(A^m)_{ij} \neq 0$.
3. $(I_N + A)^{N-1} > 0$.
4. The digraph Γ_A is strongly connected.

Proof. The equivalence of conditions 1 and 4 has already been shown. For $m \in \mathbb{N}$, the sum

$$(A^m)_{ij} = \sum_{k_1, \dots, k_{m-1}} a_{ik_1} a_{k_1 k_2} \dots a_{k_{m-1} j} = 0$$

is zero if and only if each summand is zero. But this is equivalent to the property that there exists no walk from i to j of length m . Thus condition 2 is equivalent to condition 4.

Fig. 8.6 Cyclic graph



Assume condition 3. Since A is nonnegative, the entries of the matrix

$$(I + A)^{N-1} = \sum_{j=0}^{N-1} \binom{N-1}{j} A^j$$

are all positive. Thus, for each i, j , the ij entry of A^m is positive, for $m = 1, \dots, N-1$. Thus condition 2 is satisfied and condition 3 implies condition 2 and, therefore, condition 4. Conversely, assume that Γ_A is strongly connected. Between every two distinct vertices there exists, then, a path of length $\leq N-1$ that connects them. Thus the off-diagonal entries of $(I + A)^{N-1}$ are positive. Since the diagonal entries of $(I + A)^{N-1}$ are positive, too, this implies condition 3. This completes the proof. ■

The adjacency matrix of the cyclic graph in Figure 8.6 is

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and satisfies, for each m ,

$$A^{2m+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A^{2m} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus A is an example of a nonnegative matrix that satisfies assertion 2 in Theorem 8.26 but does not satisfy $A^m > 0$ for some m . Such phenomena are explained by the Perron–Frobenius theorem.

Definition 8.27. A nonnegative matrix $A \in \mathbb{R}^{N \times N}$ is called **primitive** if $A^m > 0$ for some $m \in \mathbb{N}$. The smallest such integer m is called the **primitivity index** $\gamma(A)$.

Thus a primitive nonnegative matrix A is irreducible, but the converse does not hold. In general, the primitivity index does not satisfy $\gamma(A) \leq N$. One can show that the primitivity index always satisfies the sharp bound

$$\gamma(A) \leq N^2 - 2N + 2.$$

8.5 Graph Rigidity and Euclidean Distance Matrices

Graph rigidity is an important notion from Euclidean distance geometry that plays a central role in diverse areas such as civil engineering Henneberg (1911), the characterization of tensegrity structures, molecular geometry and 2D-NMR spectroscopy Havel and Wüthrich (1985), and formation shape control of multiagent systems Anderson et. al. (2007). Formations of specified shape are useful in control for sensing and localizing objects, and formations of fixed shape can be contemplated for moving massive objects placed upon them. To steer formations of points from one location to another, steepest-descent methods are used to optimize a suitable cost function. Typically, the smooth cost function

$$V(X) = \frac{1}{4} \sum_{ij \in E} (\|x_i - x_j\|^2 - d_{ij}^2)^2$$

on the space of all formations $X = (x_1, \dots, x_N)$ of N points x_i in \mathbb{R}^m is used. The gradient flow of V is

$$\dot{x}_i = \sum_{j: ij \in E} (\|x_i - x_j\|^2 - d_{ij}^2)(x_i - x_j), \quad i = 1, \dots, N,$$

and can be shown to converge from every initial condition pointwise to a single equilibrium point. It thus provides a simple computational approach to find a formation that realizes a specified set of distances $d_{ij} > 0, ij \in E$, indexed by the edges of a graph. The characterization of such critical formations and the analysis of their local stability properties in terms of the properties of the graph are among the open research problems in this field. Such research depends crucially on a deeper understanding of Euclidean distance geometry and associated graph-theoretic concepts, such as rigidity. We next turn to a brief description of such methods.

In Euclidean distance geometry one considers a finite tuple of points x_1, \dots, x_N in Euclidean space \mathbb{R}^m , called a formation, together with an undirected, connected graph $\Gamma = (V, E)$ on the set of vertices $V = \{1, \dots, N\}$, with prescribed distances $d_{ij} = \|x_i - x_j\|$ for each edge $ij \in E$ of Γ . Conversely, by assigning positive real numbers d_{ij}^* to the edges ij of a graph Γ , one is asked to find a formation $x_1, \dots, x_N \in \mathbb{R}^m$ that realizes the d_{ij}^* as distances $\|x_i - x_j\|$. In heuristic terms (see below for a more formal definition), a formation (x_1, \dots, x_N) is then called rigid whenever there does not exist a nontrivial perturbed formation (x'_1, \dots, x'_N) near (x_1, \dots, x_N) that realizes the same prescribed distances.

Let us associate with a vertex element $i \in V$ a point x_i in Euclidean space \mathbb{R}^m . The $m \times N$ matrix $X = (x_1, \dots, x_N)$ then describes a **formation** X of N points labeled by the set of vertices of Γ . With this notation at hand, consider the smooth distance map

$$\mathcal{D}: \mathbb{R}^{m \times N} \longrightarrow \mathbb{R}^M, \quad \mathcal{D}(X) = (\|x_i - x_j\|^2)_{(i,j) \in E}.$$

The image set

$$CM_m(\Gamma) = \{\mathcal{D}(X) \mid X \in \mathbb{R}^{m \times N}\}$$

is called the **Cayley–Menger variety**. Being the image of a real polynomial map, the Cayley–Menger variety defines a semialgebraic subset of \mathbb{R}^M , which is in fact closed. It is a fundamental geometric object that is attached to the set of all realizations of a graph in \mathbb{R}^m . For simplicity, let us focus on the **complete graph** K_N with a set of vertices $V = \{1, \dots, N\}$ and a set of edges $E = V \times V$. Then the elements of the Cayley–Menger variety $CM_m(K_N)$ are in bijective correspondence with the set of $N \times N$ **Euclidean distance matrices**

$$D(x_1, \dots, x_N) = \begin{pmatrix} 0 & \|x_1 - x_2\|^2 & \dots & \|x_1 - x_N\|^2 \\ \|x_1 - x_2\|^2 & 0 & \dots & \|x_2 - x_N\|^2 \\ \vdots & \ddots & \ddots & \vdots \\ \|x_1 - x_{N-1}\|^2 & \dots & 0 & \|x_{N-1} - x_N\|^2 \\ \|x_1 - x_N\|^2 & \dots & \|x_{N-1} - x_N\|^2 & 0 \end{pmatrix}$$

defined by $x_1, \dots, x_N \in \mathbb{R}^m$. Thus,

$$D(x_1, \dots, x_N) = -2X^T X + \mathbf{x}\mathbf{e}^T + \mathbf{e}\mathbf{x}^T.$$

Here $\mathbf{x} = \text{col}(\|x_1\|^2, \dots, \|x_N\|^2) \in \mathbb{R}^N$ and $X^T X = (x_i^T x_j) \in \mathbb{R}^{N \times N}$ denote the Gramian matrix associated with x_1, \dots, x_N . In particular, $D(x_1, \dots, x_N)$ is a rank two perturbation of the rank $\leq m$ Gramian matrix $X^T X$. This observation implies that Euclidean distance matrices of N points in \mathbb{R}^m have rank $\leq m + 2$, while for generic choices of x_1, \dots, x_N the rank is equal to $m + 2$. To characterize the set of Euclidean distance matrices, one needs a simple lemma from linear algebra.

Lemma 8.28. *Let $A = A^T \in \mathbb{R}^{n \times n}$, and assume that $B \in \mathbb{R}^{r \times n}$ has full row rank r . Let Q_A denote the quadratic form $x^T A x$ defined on the kernel $\text{Ker} B$ of B . Then the symmetric matrix*

$$C = \begin{pmatrix} 0 & B^T \\ B & A \end{pmatrix}$$

satisfies the equations for rank and signature

$$\text{rk } C = \text{rk } Q_A + 2r,$$

$$\text{sign } C = \text{sign } Q_A.$$

Proof. Let R and L be invertible $r \times r$ and $n \times n$ matrices, with $LBR = (0, I_r)$, respectively. Then

$$\begin{pmatrix} R^\top & 0 \\ 0 & L \end{pmatrix} \begin{pmatrix} 0 & B^\top \\ B & A \end{pmatrix} \begin{pmatrix} R & 0 \\ 0 & L^\top \end{pmatrix} = \begin{pmatrix} 0 & (LBR)^\top \\ LBR & LAL^\top \end{pmatrix},$$

and after such a suitable transformation one can assume without loss of generality that $B = (I_r, 0)$. Partition the matrix A

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^\top & A_{22} \end{pmatrix},$$

where A_{11} , A_{12} , and A_{22} have sizes $r \times r$, $r \times (n - r)$, and $(n - r) \times (n - r)$. Applying elementary row and column operations we obtain

$$C = \begin{pmatrix} 0 & I_r & 0 \\ I_r & A_{11} & A_{12} \\ 0 & A_{12}^\top & A_{22} \end{pmatrix} = \begin{pmatrix} I & 0 & 0 \\ \frac{1}{2}A_{11} & I & 0 \\ A_{12}^\top & 0 & I \end{pmatrix} \begin{pmatrix} 0 & I_r & 0 \\ I_r & 0 & 0 \\ 0 & 0 & A_{22} \end{pmatrix} \begin{pmatrix} I & \frac{1}{2}A_{11} & A_{12} \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix}.$$

Thus the inertia theorem of Sylvester implies that

$$\begin{aligned} \text{rk } C &= \text{rk } A_{22} + 2r = \text{rk } Q_A + 2r, \\ \text{sign } C &= \text{sign } A_{11} = \text{sign } Q_A. \end{aligned}$$

■

A classical result by Menger (1928), see also Blumenthal (1953), asserts that the Euclidean distance matrices $D(x_1, \dots, x_N)$ of N points in \mathbb{R}^m have nonpositive **Cayley–Menger determinants**

$$\det \begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & -\frac{1}{2}\|x_1 - x_2\|^2 & \dots & -\frac{1}{2}\|x_1 - x_k\|^2 \\ 1 & -\frac{1}{2}\|x_1 - x_2\|^2 & 0 & \dots & -\frac{1}{2}\|x_2 - x_k\|^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & -\frac{1}{2}\|x_1 - x_k\|^2 & -\frac{1}{2}\|x_2 - x_k\|^2 & \dots & 0 \end{pmatrix} \leq 0 \tag{8.10}$$

for each $k \leq N$ (and are equal to zero for $k > m + 1$). One can easily deduce this condition from the following more general characterization of Euclidean distance matrices.

Theorem 8.29. *Let $A = (a_{ij}) \in \mathbb{R}_+^{N \times N}$ be a nonnegative symmetric matrix, with $a_{11} = \dots = a_{NN} = 0$. The following assertions are equivalent:*

- (a) *A is a Euclidean distance matrix of N points in \mathbb{R}^m .*
- (b) *There exists a nonnegative vector $\mathbf{a} \in \mathbb{R}_+^N$ that satisfies the linear matrix inequality with rank constraint*

$$-A + \mathbf{ae}^\top + \mathbf{ea}^\top \succeq 0, \quad \text{rk}(-A + \mathbf{ae}^\top + \mathbf{ea}^\top) \leq m.$$

(c) *There exists a positive semidefinite matrix S of rank $\leq m$, with*

$$-\frac{1}{2}A = S - \frac{1}{2} \left(\text{diag}(S)\mathbf{ee}^\top + \mathbf{ee}^\top \text{diag}(S) \right). \tag{8.11}$$

(d) *The matrix*

$$S_A := -\frac{1}{2} \left(I_N - \frac{1}{N}\mathbf{ee}^\top \right) A \left(I_N - \frac{1}{N}\mathbf{ee}^\top \right)$$

is positive semidefinite of rank $\leq m$.

(e) *The restriction of the quadratic form $x^\top Ax$ on $(\mathbb{R}e)^\perp$ is negative semidefinite and has rank $\leq m$.*

(f) *The **Cayley–Menger matrix***

$$CM(A) := \begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & -\frac{1}{2}a_{12} & \dots & -\frac{1}{2}a_{1N} \\ 1 & -\frac{1}{2}a_{12} & 0 & \dots & -\frac{1}{2}a_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & -\frac{1}{2}a_{1N} & -\frac{1}{2}a_{2N} & \dots & 0 \end{pmatrix}$$

has exactly one negative eigenvalue and at most $m + 1$ positive eigenvalues.

Proof. Using the identity $\|x_i - x_j\|^2 = \|x_i\|^2 + \|x_j\|^2 - 2x_i^\top x_j$, we see that $A = D(x_1, \dots, x_N)$ for some points $x_1, \dots, x_N \in \mathbb{R}^m$ if and only if

$$-\frac{1}{2}A = X^\top X - \frac{1}{2}\mathbf{ae}^\top - \frac{1}{2}\mathbf{ea}^\top, \tag{8.12}$$

where $\mathbf{a} = \text{col}(\|x_1\|^2, \dots, \|x_N\|^2)$. Equivalently, A is a Euclidean distance matrix in \mathbb{R}^m if and only if there exists a positive semidefinite matrix S of rank $\leq m$ with

$$-\frac{1}{2}A = S - \frac{1}{2} \left(\text{diag}(S)\mathbf{ee}^\top + \mathbf{ee}^\top \text{diag}(S) \right).$$

Here $\text{diag } S$ is a diagonal matrix with the same diagonal entries as S . This completes the proof of the equivalence of the first three conditions. It is easily seen that

$$S := -\frac{1}{2} \left(I_N - \frac{1}{N}\mathbf{ee}^\top \right) A \left(I_N - \frac{1}{N}\mathbf{ee}^\top \right)$$

satisfies (8.11). Thus (d) implies (c) and hence also (a). Conversely, (a) implies (8.12), and therefore

$$S_A = -\frac{1}{2} \left(I_N - \frac{1}{N} \mathbf{e} \mathbf{e}^\top \right) X^\top X \left(I_N - \frac{1}{N} \mathbf{e} \mathbf{e}^\top \right)$$

is positive semidefinite of rank $\leq m$. This shows the equivalence of (a) and (d). The equivalence of (a) with (e) and (f) follows from Lemma 8.28 by noting that S_A in (d) satisfies $x^\top S_A x = -\frac{1}{2} x^\top A x$ for all $x \in (\mathbb{R} \mathbf{e})^\perp$. This completes the proof. ■

There are two simple facts about Cayley–Menger determinants that are worth mentioning. First, for $N = 3$ points in \mathbb{R}^2 , one has the expression for the determinant of the distance matrix (with $a_{ij} := \|x_i - x_j\|$)

$$\begin{aligned} \det D(x_1, x_2, x_3) &= -\frac{1}{4} (a_{12} + a_{13} + a_{23})(a_{12} + a_{13} - a_{23})(a_{12} - a_{13} + a_{23})(-a_{12} + a_{13} + a_{23}). \end{aligned}$$

This relates to the familiar triangle inequalities that characterize a triple of nonnegative real numbers d_1, d_2, d_3 as the side lengths of a triangle. Second, a well-known formula for the volume $\text{Vol}(\Sigma)$ of the simplex Σ defined by $N + 1$ points x_0, \dots, x_N in \mathbb{R}^N asserts that

$$\text{Vol}(\Sigma) = \frac{1}{N!} |\det(x_1 - x_0, \dots, x_N - x_0)|.$$

From the translational invariance of the norm, the distance matrix $A := D(x_0, \dots, x_N) = D(0, p_1, \dots, p_N)$, with $p_i := x_i - x_0$ for $i = 1, \dots, N$. Applying (8.12) to $d := \text{col}(0, \|p_1\|^2, \dots, \|p_N\|^2)$, $P := (p_1, \dots, p_N) \in \mathbb{R}^{d \times N}$, we obtain

$$\begin{aligned} \text{Vol}^2(\Sigma) &= \left(\frac{1}{N!} \right)^2 |\det P|^2 = \det \begin{pmatrix} I_2 & 0 \\ 0 & P^\top P \end{pmatrix} \\ &= -\det \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & p_1^\top p_1 & \cdots & p_1^\top p_N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & p_N^\top p_1 & \cdots & p_N^\top p_N \end{pmatrix} \\ &= -\det \begin{pmatrix} 0 & \mathbf{e}^\top \\ \mathbf{e} & -\frac{1}{2} D \end{pmatrix}. \end{aligned}$$

Thus we obtain the formula for the squared volume of the simplex in terms of the mutual distances as

$$\text{Vol}^2(\Sigma) = -\left(\frac{1}{N!} \right)^2 \det CM,$$

where CM is the Cayley–Menger matrix of the distance matrix $D(x_0, \dots, x_N)$.

Returning to the situation of formations defined over a graph, the **rigidity matrix** of a formation is defined as the $M \times mN$ Jacobi matrix $R(X) = Jac_{\mathcal{D}}(X)$ whose ij th row ($ij \in E$) is

$$R(X)_{ij} = (e_i - e_j)^\top \otimes (x_i - x_j)^\top.$$

A formation X is called **regular** whenever

$$\text{rk}R(X) = \max_{Z \in \mathbb{R}^{m \times N}} \text{rk}R(Z).$$

The regular formations form an open and dense subset in the space of all formations. The geometry of formations is closely connected with the standard action of the Euclidean group on \mathbb{R}^m . Let $O(m)$ denote the compact matrix Lie group of all real orthogonal $m \times m$ matrices. The Euclidean group $E(m)$ then parameterizes all Euclidean group transformations of the form $p \mapsto gp + v$, where $g \in O(m)$ and $v \in \mathbb{R}^m$ denotes a translation vector. Thus $E(m)$ is a Lie group of dimension $\frac{m(m+1)}{2}$, which is in fact a semidirect product of $O(m)$ and \mathbb{R}^m . Since \mathcal{D} is invariant under orthogonal rotations, i.e., $\mathcal{D}(SX) = \mathcal{D}(X)$ for $S \in O(m)$, the tangent space to such a group orbit is contained in the kernel of the rigidity matrix $R(X)$.

Lemma 8.30. *The kernel of the rigidity map $R(X)$ contains the tangent space $T_X(O(m) \cdot X)$. Suppose the linear span of the columns x_1, \dots, x_N of X has dimension r . Then the kernel of $R(X)$ has at least dimension $\frac{1}{2}r(2m - r - 1)$.*

Proof. The first statement is a simple consequence of the invariance of \mathcal{D} under the group of orthogonal transformations $X \mapsto SX$. Note that the stabilizer group $O(m)_X$ of X coincides with the subgroup of $O(m)$ that leaves the elements of the linear span $\langle x_1, \dots, x_N \rangle$ pointwise invariant. Thus a straightforward computation reveals that the dimension of $O(m)_X$ is equal to $\frac{1}{2}(m - r)(m - r - 1)$. Therefore, the dimension of the group orbit $O(m) \cdot X$ is equal to $\frac{1}{2}m(m - 1) - \frac{1}{2}(m - r)(m - r - 1) = \frac{1}{2}r(2m - r - 1)$. This completes the proof. ■

A formation is called **infinitesimally rigid** if the kernel of the rigidity matrix coincides with the tangent space $T_X(O(m) \cdot X)$. Equivalently, infinitesimal rigidity holds if and only if the following rank condition is satisfied:

$$\text{rk}R(X) = m(N - 1) - \frac{1}{2} \text{rk}X(2m - \text{rk}X - 1).$$

Note that from the structure of $R(X)$ one can easily check that $\text{rk}R(X) \leq N - 1$ for $r = 1$. Likewise one can check that $\text{rk}R(X) \leq 2N - 3$ for $r = 2$ and all d with $M \geq 2N - 3$. With the aid of these bounds, one can then verify using the rank condition that a formation of N points in the plane \mathbb{R}^2 is infinitesimally rigid if and only if $r = 2$ and the rank of $R(X)$ is equal to $2N - 3$. Similarly, a formation of $N \geq 4$ points in \mathbb{R}^3 is infinitesimally rigid if and only if $r = 3$ and the rank of $R(X)$ is equal

Fig. 8.7 Complete graphs are rigid

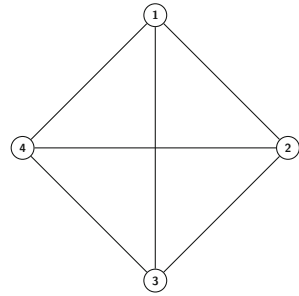


Fig. 8.8 Nonrigid graph

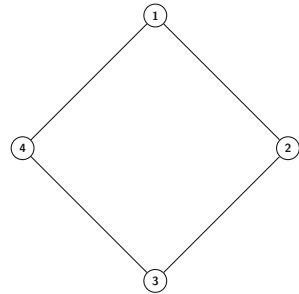
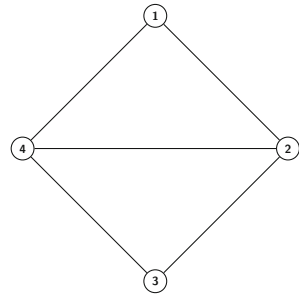


Fig. 8.9 Minimally rigid graph



to $3N - 6$. A formation X is called **rigid** whenever the orbit $SO(m) \cdot X$ is isolated in the fiber $\mathcal{D}^{-1}(\mathcal{D}(X))$. Every infinitesimally rigid formation is rigid, but the converse does not hold. In fact, regular formations are infinitesimally rigid if and only if they are rigid (Figures 8.7 and 8.8).

A **rigid graph** in \mathbb{R}^m is one for which almost every $X \in \mathbb{R}^m$ is infinitesimally rigid. Thus Γ is rigid in \mathbb{R}^m if and only if the rigidity matrix $R(X)$ has generic rank equal to $mN - \frac{m(m+1)}{2}$. A rigid graph is called **minimally rigid** if it has exactly $mN - \frac{m(m+1)}{2}$ edges (Figure 8.9). An example of a rigid graph is the complete graph K_N that has an edge between each pair of the N vertices. K_N is minimally rigid if and only if $N = 2, 3$, but not if $N \geq 4$. In contrast, a graph with 4 vertices and 5 edges realized in \mathbb{R}^2 is minimally rigid.

Rigid graphs in \mathbb{R}^2 are characterized in a combinatorial manner by the so-called Laman condition stated in the next theorem.

Theorem 8.31 (Laman (1970)). *A graph Γ with M edges and N vertices is minimally rigid in \mathbb{R}^2 if and only if the following two conditions are satisfied:*

- (a) $M = 2N - 3$.
- (b) Every subgraph Γ' of Γ with N' vertices and M' edges satisfies $M' \leq 2N' - 3$.

An explicit combinatorial characterization of rigid graphs in \mathbb{R}^3 is unknown.

8.6 Spectral Graph Theory

Definition 8.32. The **spectrum of a weighted graph** is defined as the set of eigenvalues of the adjacency matrix $A(\Gamma)$, counted with their multiplicities. The characteristic polynomial of Γ is defined as the characteristic polynomial

$$\det(zI - A(\Gamma)) = z^N + c_1 z^{N-1} + \dots + c_N.$$

The field of spectral graph theory is concerned with attempting to characterize the properties of graphs using information on the spectrum. Typically, the graphs are not weighted, and thus the adjacency matrix considered is the classical adjacency matrix of a graph. The first three coefficients of the characteristic polynomial of an unweighted graph (without self-loops) are easily characterized as follows:

1. $c_1 = 0$,
2. $c_2 = -|E|$,
3. $c_3 = -2\delta$, where δ denotes the number of triangles in Γ .

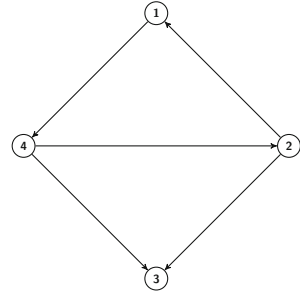
We refer the reader to Cvetkovic, Rowlinson and Simic (2010) for a detailed study of graph spectra.

Laplacian matrices are constructed from the adjacency matrix of a weighted graph through the notion of oriented incidence matrices. Their spectral properties dominate the present theory of consensus and synchronization, as will be further explained in Chapter 11. Let $\Gamma = (V, E)$ be an oriented (directed or undirected) weighted graph with associated maps $\iota : E \rightarrow V$ and $\tau : E \rightarrow V$ on the initial and terminal points, respectively. Here and in the sequel we will always use the canonical orientation on a digraph. Assume that Γ has n vertices $\{v_1, \dots, v_N\}$ and M edges $\{e_1, \dots, e_M\}$. Thus, for an edge $e \in E$ with initial and terminal points $v_i = \iota(e), v_j = \tau(e)$, respectively, there are associated weights $a_e = a_{ij} \geq 0$. Let $A \in \mathbb{R}_+^{N \times N}$ denote the weight adjacency matrix of $\Gamma = (V, E)$. Equivalently, one can present the weights as the diagonal matrix

$$W = \text{diag}(a_e)_{e \in E} \in \mathbb{R}^{M \times M}.$$

The **oriented incidence matrix** is defined as

Fig. 8.10 Orientation labeling of a graph



$$B = (b_{ij}) \in \mathbb{R}^{N \times M}, \quad b_{ij} = \begin{cases} 1 & \text{if } v_i = \tau(e_j) \neq \iota(e_j), \\ -1 & \text{if } v_i = \iota(e_j) \neq \tau(e_j), \\ 0 & \text{otherwise.} \end{cases} \quad (8.13)$$

Thus every incidence matrix has in each of its columns a single entry 1 and a single entry -1 . All other entries in the column are zero. If B, B' are incidence matrices of two identical graphs but with different orientations, then $B' = BS$ for a unique diagonal matrix $S = \text{diag}(s_1, \dots, s_M)$ and $s_i = \pm 1$. Thus the product $B'W(B')^\top = BWSB^\top$ is independent of the orientation. If the graph Γ is strongly connected, then the incidence matrix B is known to have full row rank $N - 1$; see, for example, Fiedler (2008). This shows that B has full row rank $N - 1$ and the kernel of B has dimension $M - N + 1$. Each vector in $\text{Ker } B$ describes a cycle in the graph Γ . Thus there are exactly $M - N + 1$ linearly independent cycles for a (directed or undirected) graph defined by a basis of the kernel of B with integer coefficients (Figure 8.10).

Let $\overleftarrow{\Gamma} = (V, \overleftarrow{E})$ denote the so-called **reverse graph**, i.e., Γ and $\overleftarrow{\Gamma}$ have the same set V of vertices and for each pair of vertices u, v :

$$(u, v) \in E \iff (v, u) \in \overleftarrow{E}.$$

Then the adjacency matrix of $\overleftarrow{\Gamma}$ is the transpose of that of Γ , i.e.,

$$A(\overleftarrow{\Gamma}) = A(\Gamma)^\top.$$

Definition 8.33 (Laplacian). Let $\Gamma = (V, E)$ be a weighted (directed or undirected) graph with nonnegative weight matrix $A(\Gamma) \in \mathbb{R}_+^{N \times N}$. Let $D_\Gamma = \text{diag}(d_1, \dots, d_N)$ denote the real diagonal matrix with entries $d_i := \sum_{j=1}^N a_{ij}$. The **Laplacian** of the weighted digraph is defined as

$$L(\Gamma) = D(\Gamma) - A(\Gamma) \in \mathbb{R}^{N \times N}.$$

The **canonical Laplacian** $\mathcal{L}(\Gamma) \in \mathbb{Z}^{N \times N}$ of a (directed or undirected) graph is defined as a Laplacian with the canonical 0, 1 adjacency weight matrix $\mathfrak{A}(\Gamma)$.

Proposition 8.34. *Let $\Gamma = (V, E)$ be an oriented directed graph with nonnegative weight matrix A_Γ . Then*

$$BWB^\top = L(\Gamma) + L(\overleftarrow{\Gamma}).$$

If Γ is undirected, then

$$BWB^\top = L(\Gamma).$$

Proof. Let e_1, \dots, e_M denote the edges of Γ . The ij entry of BWB^\top is equal to $\sum_{k=1}^M a_{e_k} b_{ik} b_{jk}$. Note that $b_{ik} b_{jk} = 0$ if and only if either i or j is not a vertex of the edge e_k . This shows that

$$(BWB^\top)_{ij} = \begin{cases} -(a_{ij} + a_{ji}) & \text{for } i \neq j \\ \sum_{r=1}^N (a_{ir} + a_{ri}) & \text{for } i = j. \end{cases}$$

This shows the formula for directed graphs. For undirected graphs, the edges ij and ji appear only once. This accounts for the factor of $\frac{1}{2}$. This completes the proof. ■

The Laplacian of an undirected, weighted graph is always a real symmetric matrix and therefore has only real eigenvalues. Of course, for a directed graph this does not need to be true. However, there are important constraints on the spectrum of a Laplacian that are specified by the next theorem.

For $\xi = \text{col}(\xi_1, \dots, \xi_N) \in \mathbb{R}^N$, introduce the diagonal matrix $\text{diag}(\xi_i - \xi_j)_{(i,j) \in E} \in \mathbb{R}^{M \times M}$. The oriented incidence matrix then satisfies the useful identity

$$B^\top \xi = \text{diag}(\xi_i - \xi_j) \mathbf{e}.$$

The preceding lemma then implies the explicit representation of the associated quadratic form as

$$\begin{aligned} Q_\Gamma(\xi) &= \xi^\top (L(\Gamma) + L(\overleftarrow{\Gamma})) \xi = \xi^\top BWB^\top \xi \\ &= \sum_{i,j=1}^n a_{ij} (\xi_i - \xi_j)^2. \end{aligned} \tag{8.14}$$

Most spectral properties of Laplacian matrices can be derived via the **canonical quadratic form** $Q_\Gamma : \mathbb{R}^{N \times N} \rightarrow \mathbb{R}$ associated with the graph.

Lemma 8.35. *The quadratic form Q_Γ vanishes on $\mathbb{R} \mathbf{e}$. It is exactly zero on $\mathbb{R} \mathbf{e}$ if the graph Γ is strongly connected.*

Proof. The first claim is obvious. By (8.14), it follows that $Q_\Gamma(\xi) = 0$ if and only if $a_{ij}(\xi_i - \xi_j)^2 = 0$ for all i, j . This implies $\xi_i = \xi_j$ for all edges $(i, j) \in E$. Thus, if Γ is strongly connected, this implies $\xi_1 = \dots = \xi_n$. ■

Theorem 8.36. *Let $L(\Gamma)$ denote the Laplacian of a weighted graph Γ on N vertices with nonnegative weights. Let $\mathbf{e} = (1, \dots, 1)^\top \in \mathbb{R}^N$. Then $L(\Gamma)\mathbf{e} = L(\overleftarrow{\Gamma})\mathbf{e} = 0$. Moreover:*

1. *The eigenvalues of $L(\Gamma)$ have nonnegative real part;*
2. *If Γ is strongly connected, then $L(\Gamma)$ has rank $N - 1$, i.e.,*

$$\text{Ker}L(\Gamma) = \mathbb{R}\mathbf{e}.$$

Thus, 0 is a simple eigenvalue of $L(\Gamma)$, and all other eigenvalues of L_Γ have positive real part.

3. *The quadratic form Q_Γ is positive semidefinite if and only if $\mathbf{e}^\top L(\Gamma) = 0$.*

Proof. The claim $L(\Gamma)\mathbf{e} = L(\overleftarrow{\Gamma})\mathbf{e} = 0$ is obvious from the definition of Laplacians. Note that $\mathbf{e}^\top L(\Gamma) = 0$ if and only if $\sum_{j=1}^N a_{ij} = \sum_{j=1}^N a_{ji}$ for all $i = 1, \dots, N$. This in turn is equivalent to $L(\overleftarrow{\Gamma}) = L(\Gamma)^\top$. For simplicity, we only prove claims 1 and 2 for symmetric weights, i.e., for $L(\overleftarrow{\Gamma}) = L(\Gamma)^\top$. See Bullo, Cortés and Martínez (2009), Theorem 1.32, for the proof in the general case.

Claims 1 and 2. Assume $Lx = \lambda x$ for a nonzero complex vector x and $\lambda \in \mathbb{C}$. By Proposition 8.34, then $L + L^\top = BWB^\top$, and therefore, using $x^* = \bar{x}^\top$,

$$2\text{Re}\lambda \|x\|^2 = x^*Lx + x^*L^\top x = x^*BWB^\top x \geq 0.$$

Thus $\text{Re}\lambda \geq 0$. Now suppose that Γ is strongly connected. By Lemma 8.35, the symmetric matrix $L(\Gamma) + L(\Gamma)^\top$ is positive semidefinite and degenerates exactly on $\mathbb{R}\mathbf{e}$. This proves the claim.

Claim 3. If $\mathbf{e}^\top L(\Gamma) = 0$, then $L(\Gamma) + L(\Gamma)^\top = L(\Gamma) + L(\overleftarrow{\Gamma}) = BWB^\top$ is positive semidefinite. Conversely, assume that $L(\Gamma) + L(\Gamma)^\top$ is positive semidefinite. Then

$$0 = \mathbf{e}^\top (L(\Gamma) + L(\Gamma)^\top) \mathbf{e}.$$

By the positive semidefiniteness of $L(\Gamma) + L(\Gamma)^\top$, thus $(L(\Gamma)^\top \mathbf{e} = L(\Gamma) + L(\Gamma)^\top) \mathbf{e} = 0$. This proves (3). ■

A classical result on the eigenvalues of Hermitian matrices is the Courant–Fischer minimax principle, which characterizes the eigenvalues of Hermitian matrices.

Theorem 8.37 (Courant–Fischer). *Let $A = A^* \in \mathbb{C}^{n \times n}$ be Hermitian with eigenvalues $\lambda_1 \geq \dots \geq \lambda_n$. Let E_i denote the direct sum of the eigenspaces corresponding to the eigenvalues $\lambda_1, \dots, \lambda_i$. Then*

$$\lambda_i = \min_{\dim S = n-i+1} \max_{0 \neq x \in S} \frac{x^* A x}{x^* x} = \min_{0 \neq x \in E_i} \frac{x^* A x}{x^* x}.$$

The minimax principle can be extended to establish bounds for the eigenvalues of sums of Hermitian matrices, such as the Weyl inequalities. The **Weyl inequality** asserts that the eigenvalues of the sum $A + B$ of Hermitian matrices $A, B \in \mathbb{C}^{n \times n}$, ordered decreasingly, satisfy, for $1 \leq i + j - 1 \leq n$, the inequality

$$\lambda_{i+j-1}(A+B) \leq \lambda_i(A) + \lambda_j(B).$$

More refined eigenvalue estimates are obtained from generalizations of the Weyl inequalities such as the Freede–Thompson inequality. We next state a straightforward consequence of the minimax principle to Laplacian matrices.

Corollary 8.38. *Let $L(\Gamma)$ denote the Laplacian of a weighted, directed graph on N vertices with nonnegative weights. Assume that Γ is strongly connected and satisfies $\mathbf{e}^\top L = 0$. Then the eigenvalue λ_{N-1} of $L(\Gamma)$ with smallest positive real part satisfies*

$$\operatorname{Re}(\lambda_{N-1}) \geq \min_{0 \neq x \in (\mathbb{R}\mathbf{e})^\perp} \frac{x^* L x}{x^* x} > 0. \quad (8.15)$$

Moreover, if $L(\Gamma)$ is symmetric, then λ_{N-1} is real, and equality in (8.15) holds.

Proof. By Theorem 8.36 the quadratic form $Q_\Gamma(x) = 2x^* L x = x^*(L + L^*)x$ is positive semidefinite and degenerates exactly on $\mathbb{R}\mathbf{e}$. Moreover, all nonzero eigenvalues σ of L have positive real part. Thus the eigenvectors v of L with $Lv = \sigma v$ satisfy $v^* \mathbf{e} = 0$ and $v^*(L + L^*)v = 2\operatorname{Re}(\sigma)\|v\|^2$. Thus the result follows from the minimax principle applied to $A := L + L^*$. ■

Let us briefly mention a coordinate-free approach to Laplacian matrices. Let $S(N)$ denote the vector space of real symmetric $N \times N$ matrices. For an $N \times N$ matrix S , let $\delta(S) = (s_{11}, s_{22}, \dots, s_{NN})^\top$ denote the column vector defined by the diagonal entries of S . Moreover, let $\operatorname{diag}(S)$ denote the diagonal matrix obtained from S by setting all off-diagonal entries equal to zero. Define a linear map $\mathcal{D} : S(N) \rightarrow S(N)$ as

$$\mathcal{D}(S) = S - \frac{1}{2}(\delta(S)\mathbf{e}^\top + \mathbf{e}\delta(S)^\top),$$

where $\mathbf{e} = (1, \dots, 1)^\top \in \mathbb{R}^N$.

Lemma 8.39.

$$\operatorname{Ker} \mathcal{D} = \{a\mathbf{e}^\top + \mathbf{e}a^\top\}.$$

Proof. The inclusion \subset follows from the definition of $\mathcal{D}(S)$ with $a = \frac{1}{2}\delta(S)$. For the other direction let $S = a\mathbf{e}^\top + \mathbf{e}a^\top$. Then $\delta(S) = 2a$, and thus $\mathcal{D}(S) = 0$. ■

It is easily seen that the adjoint operator of \mathcal{D} with respect to the Frobenius inner product $\langle S_1, S_2 \rangle := \text{tr}(S_1 S_2)$ is $\mathcal{D}^* : S(N) \rightarrow S(N)$, with

$$\mathcal{D}^*(S) = S - \frac{1}{2} \text{diag}(\delta(S) \mathbf{e} \mathbf{e}^\top + \mathbf{e} \mathbf{e}^\top \delta(S)^\top).$$

The **Laplacian operator** is the linear map $L = \mathcal{D}^* \circ \mathcal{D}$. Obviously, a symmetric matrix S satisfies

$$\text{tr}(L(S)S) = \text{tr}(\mathcal{D}(S)\mathcal{D}(S)) = \|\mathcal{D}(S)\|^2 \geq 0,$$

and therefore $\text{tr}(L(S)S) = 0$ if and only if $\mathcal{D}(S) = 0$. Note further that $\mathcal{D} \circ \mathcal{D}^* = \mathcal{D}^*$ and the operators \mathcal{D} and \mathcal{D}^* commute on the space of symmetric matrices with zero diagonal entries. A brute force calculation shows

$$L(S) = S - \frac{1}{2}(\delta(S)\mathbf{e}^\top + \mathbf{e}\delta(S)^\top) - \frac{1}{2} \text{diag}(\mathbf{S} \mathbf{e} \mathbf{e}^\top + \mathbf{e} \mathbf{e}^\top S) + \frac{N}{2} \text{diag}(S) + \frac{1}{2} \text{tr}(S) I_N.$$

Using the preceding formula for the Laplacian operator one concludes for each symmetric matrix S that

$$L_{ij}(S) = \begin{cases} s_{ij} - \frac{s_{ii} + s_{jj}}{2} & \text{if } i \neq j \\ -\sum_{k=1}^N (s_{ik} - \frac{s_{ii} + s_{kk}}{2}) & \text{if } i = j. \end{cases}$$

In particular, $L(S)\mathbf{e} = 0$ for all S . This explicit formula implies the following corollary.

Corollary 8.40. For $(x_1, \dots, x_N) \in \mathbb{R}^{m \times N}$, define the distance matrix $D(x_1, \dots, x_N) = (\|x_i - x_j\|^2)$. Then the Laplacian of $S = X^\top X$ is

$$L(X^\top X) = -\frac{1}{2}(D(x_1, \dots, x_N) - \text{diag}(D(x_1, \dots, x_N)\mathbf{e})).$$

In particular, if the matrix $D(x_1, \dots, x_N)$ is irreducible, then

$$\text{Ker} L(X^\top X) = \mathbb{R} \mathbf{e}.$$

Laplacian operators share an important monotonicity property.

Proposition 8.41. L is a monotonic operator, i.e., if $S_1 - S_2$ is positive semidefinite, then so is $L(S_1) - L(S_2)$. The kernel of L is equal to the kernel of \mathcal{D} .

Proof. The second claim is obvious from the definition of $L = \mathcal{D}^* \circ \mathcal{D}$. For the first claim it suffices to show that L maps positive semidefinite matrices into positive semidefinite matrices. If S is positive semidefinite, then there exists a full column

row matrix $X = (x_1, \dots, x_N)$, with $S = X^\top X$. Therefore, $L_{ij}(X^\top X) = -\|x_i - x_j\|^2$ for $i \neq j$ and $L_{ii}(X^\top X) = \sum_{j=1}^N \|x_i - x_j\|^2$. Thus, for a vector ξ ,

$$\xi^\top L(X^\top X)\xi = \frac{1}{2} \sum_{i < j} \|x_i - x_j\|^2 (\xi_i - \xi_j)^2 \geq 0.$$

This completes the proof. ■

A different version of the Laplacian matrix of a graph that is frequently of interest in applications is the **normalized Laplacian**, or the **flocking matrix**

$$\mathcal{L} = D^{-1}A.$$

Here A denotes the weighted adjacency matrix and $D = \text{diag}(Ae)$. We list a few spectral properties of the normalized Laplacian for undirected graphs.

Theorem 8.42. *The normalized Laplacian \mathcal{L} of an undirected, weighted, connected graph Γ has the following properties:*

1. \mathcal{L} is a stochastic matrix with only real eigenvalues $-1 \leq \lambda \leq 1$;
2. 1 is a simple eigenvalue of \mathcal{L} with eigenspace $\mathbb{R}e$. Moreover, -1 is not an eigenvalue of Γ if and only if Γ is not bipartite;
3. If A has at least one positive entry on the diagonal, then -1 is not an eigenvalue of \mathcal{L} .

Proof. \mathcal{L} is similar to the real symmetric matrix

$$D^{\frac{1}{2}} \mathcal{L} D^{-\frac{1}{2}} = D^{-\frac{1}{2}} A D^{-\frac{1}{2}} = I - D^{-\frac{1}{2}} L D^{-\frac{1}{2}}$$

and therefore has only real eigenvalues. Moreover, $D^{-1}Ax = x$ if and only if $Lx = (D - A)x = 0$. Theorem 8.36 implies that 1 is a simple eigenvalue of \mathcal{L} with eigenspace equal to the kernel of L , i.e., it coincides with $\mathbb{R}e$. \mathcal{L} is nonnegative, with $\mathcal{L}e = e$, and therefore a stochastic matrix. Thus Theorem 8.18 implies that \mathcal{L} has spectral radius 1. Moreover, the irreducibility of the adjacency matrix A implies that \mathcal{L} is irreducible. Suppose that -1 is an eigenvalue of \mathcal{L} . Applying Theorem 8.23 we conclude that \mathcal{L} and, hence, A are permutation equivalent to a matrix of the form

$$\begin{pmatrix} 0 & B_1 \\ B_1^\top & 0 \end{pmatrix}. \tag{8.16}$$

But this is equivalent to the graph being bipartite. Conversely, assume that A is the adjacency matrix of a bipartite graph. Then \mathcal{L} is permutation equivalent to (8.16). Thus the characteristic polynomial of \mathcal{L} is even, and therefore -1 is an eigenvalue. This proves the first two claims. Now suppose that A and, hence, \mathcal{L} have at least one positive diagonal entry. Then \mathcal{L} cannot be permutation equivalent to a matrix of the

form (8.16) because diagonal entries remain on the diagonal under permutations. Thus Theorem 8.23 implies that -1 cannot be an eigenvalue of \mathcal{L} . This completes the proof. ■

Finally, we prove the classical matrix-tree theorem for weighted graphs, which plays an important role in algebraic combinatorics. We use the following notation. For each edge $e \in E$ let $a_e > 0$ denote the associated weight. For a subset $E' \subset E$ define

$$a_{E'} = \prod_{e \in E'} a_e.$$

Note that the classical adjoint of a matrix A is the transposed matrix $\text{adj } A$ of cofactors, i.e., $\text{adj } A_{ij} = (-1)^{i+j} \det A_{ji}$.

Theorem 8.43 (Matrix-Tree Theorem). *Let Γ be an undirected weighted graph and $L(\Gamma)$ the associated Laplacian with real eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$. Then:*

1. *The $(N - 1) \times (N - 1)$ leading principal minor of $L(\Gamma)$ is equal to*

$$\kappa(\Gamma) = \sum_{|E'|=N-1} a_{E'}, \tag{8.17}$$

where the sum is over all spanning subtrees (V, E') of Γ ;

2. *The adjoint matrix of $L(\Gamma)$ is*

$$\text{adj } L(\Gamma) = \kappa(\Gamma) \mathbf{e} \mathbf{e}^\top; \tag{8.18}$$

- 3.

$$\kappa(\Gamma) = \frac{\lambda_2 \cdots \lambda_N}{N}. \tag{8.19}$$

Proof. By Lemma 8.34, the $(N-1)$ -minor of the Laplacian is equal to $\det(B_1 W B_1^\top)$, where $B_1 \in \mathbb{R}^{(N-1) \times M}$ denotes the $(N - 1) \times M$ submatrix of the oriented incidence matrix B formed by the first $N - 1$ rows. By the Cauchy–Binet formula,

$$\det(B_1 W B_1^\top) = \sum_{|E'|=N-1} \det^2(B_{E'}) \det W_{E'}, \tag{8.20}$$

where the summation is over all subsets of edges $E' \subset E \cap (\{1, \dots, N - 1\} \times \{1, \dots, N - 1\})$ of cardinality $N - 1$. One can always assume without loss of generality that the subgraphs $(\{1, \dots, N - 1\}, E')$ are connected, because otherwise $B_{E'}$ would contain a zero row (or zero column), and thus $\det^2(B_{E'}) = 0$. Assume that $(\{1, \dots, N - 1\}, E')$ contains a cycle of length $r \leq N - 1$. Then, after a suitable permutation of rows and columns, $B_{E'}$ would be of the form

$$\begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix},$$

with

$$B_{11} = \begin{pmatrix} 1 & 0 & \cdots & 0 & -1 \\ -1 & 1 & 0 & & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & & -1 & 1 & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}$$

a circulant matrix whose last column is the negative of the sum of the previous ones. Thus $\det B_{11} = 0$, and therefore the corresponding summand in (8.20) vanishes. Thus only the cycle-free subgraphs $\Gamma' = (\{1, \dots, N-1\}, E')$ contribute. These are exactly the spanning subtrees of $(\{1, \dots, N-1\}, E')$ with determinant $\det B_{E'} = \pm 1$. This proves (8.20).

For the second claim note that Γ is connected if and only if the rank of the Laplacian is $n - 1$. Thus, if Γ is not connected, then both sides of (8.19) are zero. Hence, one can assume that Γ is connected. From $L(\Gamma)\text{adj } L(\Gamma) = \det L(\Gamma)I = 0$ we conclude that every column of $\text{adj } L(\Gamma)$ is in the kernel of $L(\Gamma)$, i.e., is a scalar multiple of \mathbf{e} . Since $L(\Gamma)$ is symmetric, so is $\text{adj } L(\Gamma)$. Thus $\text{adj } L(\Gamma)$ is a multiple of $\mathbf{e}\mathbf{e}^\top$.

The last claim follows by taking traces in (8.19). Thus $N\kappa(\Gamma) = \text{tr } \text{adj } L(\Gamma)$ coincides with the sum of eigenvalues of $\text{adj } L(\Gamma)$. If $\lambda_1, \dots, \lambda_N$ denote the eigenvalues of a matrix M , then the eigenvalues of the adjoint $\text{adj } M$ are $\prod_{j \neq i} \lambda_j$, $j = 1, \dots, N$. Thus we obtain $\text{tr } \text{adj } L(\Gamma) = \lambda_2 \cdots \lambda_N$. This completes the proof. ■

As a consequence of the matrix-tree theorem one can derive an explicit formula for the number of spanning trees in a graph.

Corollary 8.44 (Temperley (1964)). *Let $L := L(\Gamma)$ be the Laplacian of an undirected weighted graph on N vertices and $J = \mathbf{e}\mathbf{e}^\top$. Then*

$$\kappa(\Gamma) = \frac{\det(J + L)}{N^2}.$$

Proof. The identities $NJ = J^2, JL = 0$ imply $(NI - J)(J + L) = NL$. By taking adjoints, therefore, $\text{adj } (J + L)\text{adj } (NI - J) = \text{adj } (NL) = N^{N-1}\text{adj } L$. Thus, using $\text{adj } (NI - J) = N^{N-2}J$, the matrix-tree theorem implies

$$N^{N-1}\kappa(\Gamma)J = N^{N-1}\text{adj } L = N^{N-2}\text{adj } (J + L)J.$$

Thus $N\kappa(\Gamma)J = \text{adj } (J + L)J$, and therefore

$$\begin{aligned} \det (J+L)J &= (J+L)\text{adj} (J+L)J = \text{adj} (J+L)(J+L)J = N\text{adj} (J+L)J \\ &= N^2\kappa(\Gamma). \end{aligned}$$

■

For the complete graph K_N on N vertices, the classical graph Laplacian is $\mathcal{L} = NI - J$. This implies the well-known formula

$$\kappa(K_N) = \frac{\det (J+\mathcal{L})}{N^2} = \frac{N^N}{N^2} = N^{N-2}$$

for the number of spanning trees in K_N .

8.7 Laplacians of Simple Graphs

We determine spectral information for important classes of classical Laplacians and adjacency matrices for directed and undirected graphs Γ . For simplicity we focus on unweighted graphs, i.e., on classical Laplacians and adjacency matrices. Note that both Laplacians and adjacency matrices of undirected, weighted graphs satisfy the following properties.

$$\begin{aligned} L(\Gamma \cup \Gamma') &= \text{diag} (L(\Gamma), L(\Gamma')), \\ L(\Gamma \times \Gamma') &= L(\Gamma) \otimes I_n + I_m \otimes L(\Gamma'). \end{aligned}$$

In particular, the eigenvalues of the Laplacian of the direct product graph $\Gamma \times \Gamma'$ are the sums $\lambda_i + \mu_j$ of the eigenvalues λ_i and μ_j of Γ and Γ' , respectively.

1. Simple Path Graph. Our first example is the simple directed path graph Γ_n on the vertex set $V = \{1, \dots, N\}$ and edges $E = \{(1, 2), (2, 3), \dots, (N - 1, N)\}$ (Figure 8.11).

The adjacency matrix and Laplacian of Γ_N are respectively

$$\mathfrak{A} = \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}, \quad L = \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ -1 & 1 & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -1 & 1 \end{pmatrix}.$$

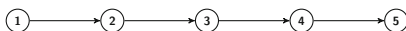


Fig. 8.11 Directed simple path

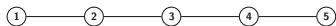


Fig. 8.12 Undirected simple path

Thus \mathfrak{A} is the standard cyclic nilpotent matrix, while L has 0 as a simple eigenvalue and the eigenvalue 1 has a geometric multiplicity of one (and an algebraic multiplicity of $N - 1$) (Figure 8.12).

More interesting is the associated undirected graph with a set of edges $E = \{\{1, 2\}, \{2, 3\}, \dots, \{N - 1, N\}\}$ and graph adjacency and Laplacian matrices, respectively,

$$\mathfrak{A}_N = \begin{pmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ & 1 & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & & 1 & 0 \end{pmatrix}, \quad L_N = \begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{pmatrix}.$$

We begin with a spectral analysis of \mathfrak{A}_N . This is a classical exercise from analysis.

- Theorem 8.45.** 1. The eigenvalues of \mathfrak{A}_N are the distinct real numbers $\lambda_k(\mathfrak{A}_N) = 2 \cos \frac{k\pi}{N+1}$, $k = 1, \dots, N$.
2. The unique eigenvector $x^{(k)} = (\xi_0^{(k)}, \dots, \xi_{N-1}^{(k)})^\top$, normalized as $\xi_0^{(k)} := 1$, for the eigenvalue $2 \cos \frac{k\pi}{N+1}$ is

$$\xi_v^{(k)} = \frac{\sin k \left(\frac{v+1}{N+1}\right) \pi}{\sin \left(\frac{k\pi}{N+1}\right)}, \quad v = 0, \dots, N - 1.$$

In particular, the coordinates of the eigenvector $x^{(k)}$ are reflection symmetric, that is, they satisfy $\xi_v^{(k)} = \xi_{n-1-v}^{(k)}$.

Proof. Let

$$e_N(z) = \det(zI - \mathfrak{A}_N)$$

be the characteristic polynomial of \mathfrak{A}_N . Expanding by the first row leads to the three-term recursion

$$e_N(z) = ze_{N-1}(z) - e_{N-2}(z). \tag{8.21}$$

For the 1-norm $\|\mathfrak{A}\| \leq 2$, and hence it follows that the eigenvalues satisfy $|\lambda| \leq 2$ and, as eigenvalues of a real symmetric matrix, they are all real. So one can set

$\lambda = 2 \cos x$. From (8.21) it follows that λ is an eigenvalue if and only if $e_N(\lambda) = 0$. The difference equation (8.21) can be written now as

$$e_N(2 \cos x) = 2 \cos x \cdot e_{N-1}(2 \cos x) - e_{N-2}(2 \cos x). \tag{8.22}$$

We try an exponential solution to this difference equation, i.e., we put $e_N = A\zeta_1^N + B\zeta_2^N$, where ζ_1 and ζ_2 are the two roots of the characteristic polynomial $\zeta^2 - 2\zeta \cos x + 1 = 0$. This leads to $\zeta = e^{\pm\sqrt{-1}x}$. The initial conditions for the difference equation (8.21) are $e_0(z) = 1$ and $e_1(z) = z$. Setting $e_N(2 \cos x) = Ae^{\sqrt{-1}Nx} + Be^{-\sqrt{-1}Nx}$ leads to the pair of equations

$$A + B = 1, \quad Ae^{\sqrt{-1}x} + Be^{-\sqrt{-1}x} = 2 \cos x.$$

Solving and substituting back in (8.22) one obtains

$$e_N(2 \cos x) = \frac{\sin(N+1)x}{\sin x}.$$

The right-hand side vanishes for $x = \frac{k\pi}{N+1}, k = 1, \dots, N$, and therefore the eigenvalues of \mathfrak{A}_N are $2 \cos \frac{k\pi}{N+1}$.

We proceed now to the computation of the eigenvectors of \mathfrak{A}_N . Let $x^{(k)} = (\xi_0^{(k)}, \dots, \xi_{N-1}^{(k)})$ be the eigenvector corresponding to the eigenvalue $\lambda_k = 2 \cos(\frac{k\pi}{N+1})$. The characteristic equation $Tx^{(k)} = \lambda_k x^{(k)}$ is equivalent to the system

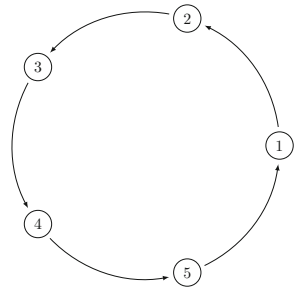
$$\begin{aligned} \xi_1^{(k)} &= \lambda_k \xi_0^{(k)} \\ \xi_0^{(k)} + \xi_2^{(k)} &= \lambda_k \xi_1^{(k)} \\ &\vdots \\ \xi_{N-3}^{(k)} + \xi_{N-1}^{(k)} &= \lambda_k \xi_{N-2}^{(k)} \\ \xi_{N-2}^{(k)} &= \lambda_k \xi_{N-1}^{(k)}. \end{aligned} \tag{8.23}$$

The coordinates of the eigenvector $x^{(k)}$ satisfy the recursion

$$\xi_v^{(k)} = 2 \cos\left(\frac{k\pi}{N+1}\right) \xi_{v-1}^{(k)} - \xi_{v-2}^{(k)}.$$

Normalize the eigenvector by requiring $\xi_0^{(k)} = 1$. The second coordinate is determined by the first equation in (8.23) and $\xi_1^{(k)} = 2 \cos(\frac{k\pi}{N+1})$. As is the case with eigenvalues, one solves the difference equation using a linear combination of two exponential solutions. Thus

Fig. 8.13 Directed cycle graph



The unique eigenvector $x^{(k)} = (\xi_0^{(k)}, \dots, \xi_{N-1}^{(k)})^\top$, normalized as $\xi_v^{(1)} := 1$, for the eigenvalue $\lambda_k(L_N)$ is

$$\xi_v^{(k)} = \cos\left(\frac{(k-1)(2v+1)\pi}{2N}\right), \quad v = 0, \dots, N-1.$$

2. Simple Cycle Graph. In this case the set of edges of the digraph is $E = \{(1, 2), \dots, (N-1, N), (N, 1)\}$ and in the undirected case

$$E = \{\{1, 2\}, \dots, \{N-1, N\}, \{1, N\}\}.$$

Consider first the directed graph case (Figure 8.13). Then the adjacency matrix is the circulant matrix

$$C_N = \begin{pmatrix} 0 & & & & 1 \\ 1 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}. \tag{8.24}$$

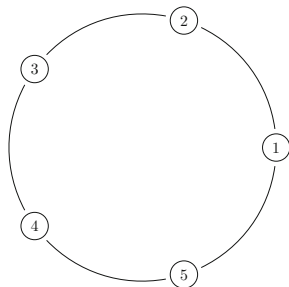
Being a circulant matrix, C is diagonalized by the Fourier matrix. Explicitly, let

$$\Phi = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2N-2} \\ \vdots & & & \ddots & \\ 1 & \omega^{N-1} & \omega^{2N-2} & \dots & \omega^{(N-1)^2} \end{pmatrix} \tag{8.25}$$

denote the Fourier matrix, where $\omega = e^{2\pi\sqrt{-1}/N}$ denotes a primitive N th root of unity. Notice that Φ is both a unitary and a symmetric matrix:

$$C_N = \Phi \text{diag}(1, \omega, \dots, \omega^{N-1}) \Phi^*.$$

Fig. 8.14 Undirected cycle graph



This proves the following theorem.

Theorem 8.47. *The eigenvalues of C_N are distinct and are the N th roots of unity:*

$$\lambda_k(C_N) = \omega^k = e^{\frac{2k\sqrt{-1}\pi}{N}}, \quad k = 1, \dots, N.$$

An eigenvector $x^{(k)} = (\xi_0^{(k)}, \dots, \xi_{N-1}^{(k)})^\top$ for the eigenvalue $\lambda_k(C_N)$ is

$$x^{(k)} = \Phi e_k = \sum_{j=0}^{N-1} \omega^{(k-1)j} e_{j+1}, \quad k = 1, \dots, N.$$

The associated Laplacian matrix is equal to $L_N = I_N - C_N$. Thus the eigenvalues and eigenvector are trivially related to those of C_N (Figure 8.14).

The undirected case is more interesting. The adjacency matrix and Laplacian matrices are

$$\mathfrak{A}_N = \begin{pmatrix} 0 & 1 & & & 1 \\ 1 & 0 & 1 & & \\ & 1 & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ 1 & & & 1 & 0 \end{pmatrix}, \quad L_N = 2I_N - \mathfrak{A}_N = \begin{pmatrix} 2 & 1 & & & -1 \\ -1 & 2 & -1 & & \\ & -1 & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ -1 & & & -1 & 2 \end{pmatrix}. \quad (8.26)$$

Theorem 8.48. *1. The eigenvalues of \mathfrak{A}_N and L_N defined in (8.26) are*

$$\lambda_k(\mathfrak{A}_N) = 2 \cos\left(\frac{2k\pi}{N}\right), \quad k = 1, \dots, N,$$

$$\lambda_k(L_N) = 2 - 2 \cos\left(\frac{2k\pi}{N}\right), \quad k = 1, \dots, N.$$

In either case, $\lambda_k = \lambda_l$ for $1 \leq k, l \leq N$ if and only if $l = N - k$. For $N = 2m$ even, λ_m and λ_N are simple and λ_k has a multiplicity of two for all other k . For $N = 2m + 1$ odd, λ_N is simple and all other eigenvalues have a multiplicity of two.

2. An orthonormal basis for the eigenspaces of \mathfrak{A}_N and L_N for the eigenvalue $\lambda_k(\mathfrak{A}_N)$ and $\lambda_k(L_N)$, respectively, is as follows:

(a) A single generator

$$\frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ -1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

for $k = N$ or $k = m, N = 2m$, respectively.

(b) Otherwise, a basis of two orthonormal vectors

$$x^{(k)} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ \cos\left(\frac{2k\pi}{N}\right) \\ \cos\left(\frac{4k\pi}{N}\right) \\ \vdots \\ \cos\left(\frac{2(N-1)k\pi}{N}\right) \end{pmatrix}, \quad y^{(k)} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ \sin\left(\frac{2k\pi}{N}\right) \\ \sin\left(\frac{4k\pi}{N}\right) \\ \vdots \\ \sin\left(\frac{2(N-1)k\pi}{N}\right) \end{pmatrix}.$$

Proof. Since

$$\mathfrak{A}_N = C_N + C_N^\top = \Phi \text{diag}(1, \omega, \dots, \omega^{N-1}) \Phi^* + \Phi \text{diag}(1, \bar{\omega}, \dots, \bar{\omega}^{N-1}) \Phi^*,$$

the eigenvalues of $C + C^\top$ are equal to $\text{Re}(\omega^k + \bar{\omega}^k) = 2 \cos\left(\frac{2k\pi}{N}\right)$. Moreover, the complex eigenvectors of \mathfrak{A}_N are simply the columns

$$\phi_k = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ \omega^k \\ \omega^{2k} \\ \vdots \\ \omega^{(N-1)k} \end{pmatrix}$$

of the Fourier matrix Φ_N . Thus the real and imaginary parts

$$x^{(k)} = \frac{1}{2}(\phi_k + \bar{\phi}_k), \quad y^{(k)} = \frac{1}{2i}(\phi_k - \bar{\phi}_k)$$

form a real basis of the corresponding eigenspaces. Writing $x^{(k)} = \frac{1}{\sqrt{N}}(\xi_0^{(k)}, \dots, \xi_{N-1}^{(k)})^\top$ and $y^{(k)} = \frac{1}{\sqrt{N}}(\eta_0^{(k)}, \dots, \eta_{N-1}^{(k)})^\top$ one obtains for each $k = 1, \dots, N$

$$x_v^{(k)} = \cos \frac{2kv\pi}{N}, \quad y_v^{(k)} = \sin \frac{2kv\pi}{N}, \quad k = 1, \dots, N, v = 0, \dots, N-1.$$

This completes the proof for \mathfrak{A}_N . The result on the Laplacian follows trivially as $L_N = 2I_N - \mathfrak{A}_N$. ■

8.8 Compressions and Extensions of Laplacians

We begin by recalling the definition of the Schur complement. Let M be an $N \times N$ matrix and $I, J \subset \{1, \dots, N\}$. Then M_{IJ} denotes the submatrix of M with row indices in I and column indices in J , respectively.

Definition 8.49. Let M be an $N \times N$ matrix and $I \subset \{1, \dots, N\}$ such that M_{II} is invertible. Let $J = \{1, \dots, N\} \setminus I \neq \emptyset$. Then

$$M/M_{II} := M_{JJ} - M_{JI}M_{II}^{-1}M_{IJ}$$

is called the **Schur complement**.

The Schur complement has some basic properties that are easily established, as follows.

Proposition 8.50. Let M be an $N \times N$ matrix and $I \subset \{1, \dots, N\}$ such that M_{II} is invertible. The Schur complement M/M_{II} has the following properties:

1. $\text{rk}M = \text{rk}M_{II} + \text{rk}M/M_{II}$;
2. Let M be Hermitian; then M/M_{II} is Hermitian with signature

$$\text{sign}(M) = \text{sign}(M_{II}) + \text{sign}(M/M_{II}).$$

Proof. Without loss of generality, assume that $I = \{1, \dots, r\}$, $1 \leq r < n$ and M is partitioned as

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}.$$

The result follows easily from the identity

$$\begin{pmatrix} I & 0 \\ -M_{21}M_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} I - M_{11}^{-1}M_{12} \\ 0 & I \end{pmatrix} = \begin{pmatrix} M_{11} & 0 \\ 0 & M_{22} - M_{21}M_{11}^{-1}M_{12} \end{pmatrix}.$$

■

The 2×2 matrix

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

shows that the Schur complement of a Hurwitz matrix need not be a Hurwitz matrix. Information about the spectral properties of the Schur complement is provided by the next result. For the proof we refer the reader to Fiedler (2008).

Theorem 8.51. *Let M be a real $N \times N$ matrix with nonpositive off-diagonal entries $m_{ij} \leq 0, i \neq j$. Let $I \subset \{1, \dots, N\}$, with M_{II} invertible. Suppose there exists a vector $x \geq 0$ with $Mx > 0$. Then:*

1. *All eigenvalues of M have positive real part;*
2. *The eigenvalues of M_{II} and M/M_{II} have positive real parts, respectively;*
3. *The off-diagonal entries of M_{II} and the Schur complement M/M_{II} are both non-positive. The inverses M_{II}^{-1} and $(M/M_{II})^{-1}$ exist and are nonnegative matrices.*

Let L denote the Laplacian matrix of an undirected, weighted graph Γ . Assume that $\Delta = \text{diag}(\delta_1, \dots, \delta_N)$ denotes a diagonal matrix with nonnegative entries $\delta_1 \geq 0, \dots, \delta_N \geq 0$. Then the matrix

$$\mathcal{L} = L + \Delta$$

is called a **generalized Laplacian** for Γ . Thus the generalized Laplacians \mathcal{L} are characterized as those matrices with nonpositive off-diagonal entries that satisfy $\mathcal{L}\mathbf{e} \geq 0$. Let A denote the weighted adjacency matrix of Γ . The **loopy Laplacian** is then the generalized Laplacian $Q = L + \Delta$ defined by setting $\delta_i := a_{ii}$ for $i = 1, \dots, N$.

We now prove a central result on the submatrices of generalized Laplacians.

Theorem 8.52. *Let Γ be an undirected, weighted graph with generalized Laplacian matrix \mathcal{L} .*

1. *Γ is connected if and only if \mathcal{L} is irreducible.*
2. *Every eigenvalue and every principal minor of \mathcal{L} are nonnegative.*
3. *For each $I \neq \{1, \dots, N\}$, \mathcal{L}_{II} is a positive definite matrix and its inverse $(\mathcal{L}_{II})^{-1}$ is a nonnegative matrix.*
4. *Let Γ be connected. Then, for each $I \neq \{1, \dots, N\}$, both \mathcal{L}_{II} and the Schur complement $\mathcal{L} / \mathcal{L}_{II}$ are generalized Laplacians.*

Proof. Since the off-diagonal entries of $-\mathcal{L}$ coincide with those of the graph adjacency matrix A of Γ , it follows that \mathcal{L} is irreducible if and only if A is irreducible. But this is equivalent to Γ being connected.

By Proposition 8.34, every principal submatrix of \mathcal{L} is of the form

$$\mathcal{L}_{II} = \Delta_{II} + B_I W B_I^\top,$$

where B_I denotes the submatrix of the incidence matrix B formed by the rows that are indexed by I . The matrix Δ_{II} is a diagonal matrix with nonnegative entries. Thus

\mathcal{L}_I has nonnegative off-diagonal terms and is positive semidefinite. This proves claim 2. Assume that Γ is connected. Then for each proper subset $I \subset \{1, \dots, N\}$, the matrix $B_I W B_I^\top$ is positive definite. In fact, $x^\top B_I W B_I^\top x = 0$ implies $B_I^\top x = 0$. Extend x to $z \in \mathbb{R}^N$ by adding zeros, so that $B^\top z = B_I^\top x$. Since $\text{Ker } B^\top = \mathbb{R}e$, therefore $z = \lambda e$. Since at least one entry of z is zero, we obtain $\lambda = 0$. Thus $x = 0$, which proves positive definiteness of $B_I W B_I^\top$. In particular, $\mathcal{L}_I = \Delta_I + B_I W B_I^\top$ is positive definite for all proper index sets I . Moreover, \mathcal{L}_I is a generalized Laplacian matrix because the off-diagonal entries are all nonpositive and $\mathcal{L}_I e \geq 0$. Let A denote a real matrix with nonpositive off-diagonal entries such that all eigenvalues of A have positive real part. By Theorem 5.2.1 in Fiedler (2008), one obtains that A^{-1} is a nonnegative matrix. Applying this result to $A = \mathcal{L}_I$ we conclude that \mathcal{L}_I^{-1} is nonnegative. This completes the proof of claim 3. Since \mathcal{L}_I is invertible, the Schur complement $\mathcal{L} / \mathcal{L}_I = \mathcal{L}_{JJ} - \mathcal{L}_{JI} \mathcal{L}_I^{-1} \mathcal{L}_{IJ}$ exists. Moreover, \mathcal{L}_I^{-1} is nonnegative and the entries of $\mathcal{L}_{JI}, \mathcal{L}_{IJ}$ are nonpositive. Thus all entries of $-\mathcal{L}_{JI} \mathcal{L}_I^{-1} \mathcal{L}_{IJ}$ are ≤ 0 . Since the off-diagonal entries of \mathcal{L}_{JJ} are all ≤ 0 , this shows that the off-diagonal entries of $\mathcal{L} / \mathcal{L}_I$ are nonpositive. Thus it remains to show that the diagonal entries of the Schur complement are nonnegative. To this end, we simplify the notation by assuming that $I = \{1, \dots, r\}$. Then diagonal entries of $\mathcal{L} / \mathcal{L}_I$ are of the form

$$v^\top \left(-\mathcal{L}_{21} \mathcal{L}_{11}^{-1} I \right) \begin{pmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} \\ \mathcal{L}_{21} & \mathcal{L}_{22} \end{pmatrix} \begin{pmatrix} -\mathcal{L}_{11}^{-1} \mathcal{L}_{12} \\ I \end{pmatrix} v = w^\top \mathcal{L} w,$$

for suitable choices of v, w . By claim 2, then $w^\top \mathcal{L} w \geq 0$, and the result follows. ■

We now explore in more detail the underlying graphs that are associated with forming submatrices and Schur complements. Let $\Gamma = (V, E)$ be an undirected weighted graph and $V' \subset V$ a nonempty subset of r vertices in V . Let $\Gamma_{V'} = (V', E \cap (V' \times V'))$ denote an induced graph with the induced weight adjacency matrix A' . The relation between the Laplacians of $\Gamma_{V'}$ and Γ is established by the following result, whose easy proof is omitted.

Proposition 8.53. *Let $I = \{i_1 < \dots < i_r\} \subset \{1, \dots, N\}$ and $V' = \{v_{i_1}, \dots, v_{i_r}\}$ denote the corresponding set of vertices in V . Let*

$$L(\Gamma)_I = (L_{ab})_{a,b \in I}$$

denote the $r \times r$ principal submatrix of the Laplacian $L(\Gamma) = (L_{ij})$, with row and column indices in I . Then

$$L(\Gamma)_I = L(\Gamma_{V'}) + D_{V'},$$

where $D_{V'} = \text{diag}(\delta_1, \dots, \delta_r)$ is a diagonal matrix with nonnegative entries $\delta_i = \sum_{j \notin I} (a_{ij} + a_{ji})$. In particular, the submatrix $L(\Gamma)_I$ of the Laplacian L is a generalized Laplacian of the induced graph $\Gamma_{V'} = (V', E \cap (V' \times V'))$.

For the Schur complement we introduce the following notion; see Fiedler (2008) and Horn and Johnson (1990).

Definition 8.54. Let $\Gamma = (V, E)$ be an undirected weighted graph on a vertex set $V = \{1, \dots, N\}$ and $I \subset V$. The **Schur complement**, or the **Kron reduction**, on J is the graph $\Gamma_J = (J, E_J)$ with the set of vertices $J := V \setminus I$. Between vertices $i, j \in J$ an edge $(i, j) \in E'$ is defined if and only if there exists a path from i to j such that all its interior vertices (i.e., those of the path that differ from i and j) belong to W .

The Kron reduction of an undirected graph is an undirected graph on a subset of vertices; however, it may contain self-loops even if $\Gamma = (V, E)$ does not have self-loops. The Kron reduction graph has some appealing properties that are partially stated in the next result.

Theorem 8.55. *Let $\Gamma = (V, E)$ be an undirected weighted graph that is connected. The Kron reduction of Γ is connected. The Schur complement $\mathcal{L}/\mathcal{L}_I$ of a generalized Laplacian of Γ is a generalized Laplacian of the Kron reduction graph $\Gamma_J = (V_J, E_J)$.*

Proof. For a proof that the Kron reduction is connected, we refer to Doerfler and Bullo (2013). By the preceding reasoning, $L(J) := \mathcal{L}/\mathcal{L}_I$ is a generalized Laplacian on the vertex set J . Thus it remains to show that the off-diagonal entries $L_{ij}(J)$ are nonzero if and only if ij is an edge of the Kron reduction $\Gamma_J = (J, E_J)$. This is shown in Theorem 14.1.2 by Fiedler (2008) in the case where \mathcal{L} possesses a vector $x \geq 0$ with $\mathcal{L}x > 0$. In Theorem 3.4 by Doerfler and Bullo (2013), this is shown for the so-called loopy Laplacian matrix of a graph. ■

The Courant–Fischer minimax principle has important implications for the characterization of the eigenvalues of submatrices of Hermitian matrices via interlacing conditions. We state one of the simplest known results here, which is often attributed to Cauchy and Weyl.

Theorem 8.56 (Eigenvalue Interlacing Theorem). *Let M be a Hermitian $n \times n$ matrix and $I \subset \{1, \dots, n\}$ a subset of cardinality r . Assume that the eigenvalues of Hermitian matrices A are ordered increasingly as $\lambda_1(A) \leq \dots \leq \lambda_n(A)$. Then, for $1 \leq k \leq r$,*

$$\lambda_k(M) \leq \lambda_k(M_{II}) \leq \lambda_{k+n-r}(M). \tag{8.27}$$

If in addition M_{II} is positive definite, then for $1 \leq k \leq n - r$,

$$\lambda_k(M) \leq \lambda_k(M/M_{II}) \leq \lambda_k(M_{JJ}) \leq \lambda_{k+r}(M).$$

Proof. For the proof of the first inequality we refer to Theorem 4.3.15 in Horn and Johnson (1990). For the second claim note that the positive definiteness of M_{II} implies that of $M_{II}M_{II}^{-1}M_{IJ}$. Therefore, $M_{JJ} \succeq M_{JJ} - M_{JI}M_{II}^{-1}M_{IJ}$, and thus $\lambda_k(M/M_{II}) \leq \lambda_k(M_{JJ})$ for all $1 \leq k \leq n$. Applying (8.27) to the submatrix M_{JJ} gives the result. ■

We next describe inequalities between the eigenvalues of the Laplacians of a graph and induced subgraph.

Theorem 8.57. *Let $\Gamma = (V, E)$ be an undirected weighted graph and $\Gamma_{V'} = (V', E \cap (V' \times V'))$ an induced subgraph on $V' \subset V$. Let $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, $\lambda'_1 \leq \lambda'_2 \leq \dots \leq \lambda'_r$, and $0 = \mu_1 \leq \mu_2 \leq \dots \leq \mu_r$ denote the eigenvalues of $L(\Gamma)$, $L(\Gamma)_{V'}$, and $L(\Gamma_{V'})$, respectively. Then for all $1 \leq k \leq r$,*

$$\lambda_k \leq \lambda'_k \leq \lambda_{N-r+k},$$

$$\sum_{j=1}^k \lambda'_j \geq \sum_{j=1}^k \mu_j + \sum_{j=1}^k \delta_j.$$

In particular,

$$\lambda_N + \dots + \lambda_{N-r+1} \geq \sum_{i=1}^r \sum_{j \neq i} (a_{ij} + a_{ji}).$$

Proof. The first inequality follows from the interlacing theorem for eigenvalues of nested Hermitian matrices; see, for example, Horn, Rhee and So (1998). The second estimate follows from a standard eigenvalue inequality for sums of Hermitian matrices. The last inequality follows from the other two. In fact, by the first inequality, $\lambda_n + \dots + \lambda_{n-r+1} \geq \text{tr}(L(\Gamma)_{V'}) = \text{tr}L(\Gamma_{V'}) + D_{V'}$. This completes the proof. ■

Similar eigenvalue inequalities exist for the Schur complement of generalized Laplacians. The straightforward proof of the next theorem is omitted.

Theorem 8.58. *Let \mathcal{L} denote a generalized Laplacian of an undirected, connected graph, and let $\mathcal{L}_{\text{red}} = \mathcal{L} / \mathcal{L}_I$ denote the Schur complement, $|I| = r$. Then the following interlacing conditions for eigenvalues are satisfied:*

$$\lambda_k(\mathcal{L}) \leq \lambda_k(\mathcal{L}_{\text{red}}) \leq \lambda_k(\mathcal{L}_{JJ}) \leq \lambda_{k+r}(\mathcal{L}) \quad \text{for } 1 \leq k \leq N - r.$$

8.9 Exercises

1. Let $\lambda_1, \dots, \lambda_n$ and μ_1, \dots, μ_m be the eigenvalues of the matrices $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{m \times m}$, respectively. Prove that the eigenvalues of the Kronecker product $A \otimes B$ and of $A \otimes I_m + I_n \otimes B$ are $\lambda_i \mu_j$ and $\lambda_i + \mu_j$, respectively, for $i = 1, \dots, n; j = 1, \dots, m$. Deduce that the Sylvester operator $A \otimes I_m - I_n \otimes B$ is invertible if and only if A and B have disjoint spectra.
2. Let $\lambda_1, \dots, \lambda_n$ and μ_1, \dots, μ_m be the eigenvalues of the matrices $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{m \times m}$, respectively. Let $p(x, y) = \sum_{i,j} c_{ij} x^i y^j$ denote a real polynomial in

two commuting variables. Generalize the preceding exercise by showing that the eigenvalues of

$$\sum_{ij} c_{ij} A^i \otimes B^j$$

are equal to $p(\lambda_k, \mu_l)$.

3. The *Hadamard product* of two matrices $A, B \in \mathbb{R}^{n \times n}$ is defined as the $n \times n$ matrix $A * B = (a_{ij} b_{ij})$. Prove that $A * B$ is a principal submatrix of $A \otimes B$. Deduce that the Hadamard product $A * B$ of two positive definite symmetric matrices A and B is again positive definite.
4. Prove that the set of matrices $A \in \mathbb{R}^{n \times n}$, with $\mathbf{e}^T A = \mathbf{e}^T$ and $A \mathbf{e} = \mathbf{e}$, forms an affine space of dimension $(n-1)^2$.
5. Prove Birkhoff's theorem, stating that the set of $n \times n$ doubly stochastic matrices form a convex polyhedron whose $n!$ vertices are permutation matrices.
6. Let $A \in \mathbb{C}^{n \times n}$ be unitary. Prove that the $n \times n$ matrix $(|a_{ij}|^2)$ is doubly stochastic.
7. Let $A \in \mathbb{R}^{n \times n}$ be irreducible and $D \in \mathbb{R}^{n \times n}$ be diagonal with $AD = DA$. Prove that $D = \lambda I_n$ is suitable for $\lambda \in \mathbb{R}$.
8. Let $A \in \mathbb{R}^{n \times n}$ be irreducible and $D_1, \dots, D_{N-1} \in \mathbb{R}^{n \times n}$ diagonal, with

$$A = e^{2\pi\sqrt{-1}k/N} D_k A D_k^{-1}, \quad k = 1, \dots, N-1.$$

Then there exists $\lambda_k \in \mathbb{C}$, with $D_k = \lambda_k D_1^k$ for $k = 1, \dots, N-1$.

9. A connected graph with N vertices, without loops and multiple edges, has at least $N-1$ edges. If the graph has more than $N-1$ edges, then it contains a polygon as a subgraph.
10. Prove that a graph is connected if and only if it has a spanning tree.
11. Consider the directed graph Γ on the vertex set $\mathcal{V} = \{1, 2, 3, 4, 5, 6\}$ with adjacency matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}.$$

- (a) Prove that Γ is strongly connected.
 - (b) Prove that there exists a cycle of period two through vertex 1 and 1 has no cycle of odd period.
 - (c) Prove that the period of A is 2.
 - (d) Compute the eigenvalues of A .
12. Let $A \in \mathbb{R}^{n \times n}$ be nonnegative and irreducible. Show that $(A + \varepsilon I)^{n-1} > 0$ for all $\varepsilon > 0$.

13. Consider the matrices

$$A_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

Check for the primitivity of the matrices and, if possible, determine the smallest $m \in \mathbb{N}$ such that $A_i^m > 0$.

14. Show that the contraction constant for the Hilbert metric of

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

is equal to

$$k(A) = \frac{2 - \sqrt{3}}{2 + \sqrt{3}},$$

while the eigenvalues of A are $\lambda_{\pm} = \frac{4 \pm \sqrt{13}}{6}$. Deduce that $k(A)$ is strictly smaller than the convergence rate for the power iteration defined by A .

15. The primitivity index $\gamma(A)$ of a nonnegative matrix A is defined as the smallest $m \in \mathbb{N}$, with $A^m > 0$. Prove that the $n \times n$ Wielandt matrix

$$A = \begin{pmatrix} 0 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & & \ddots & 1 \\ 1 & 1 & \dots & 0 \end{pmatrix}$$

is primitive with primitivity index $\gamma(A) = n^2 - 2n + 2$.

16. Prove that every nonnegative irreducible matrix $A \in \mathbb{R}^{n \times n}$ with at least one positive diagonal element is primitive.
17. Consider a real $n \times n$ tridiagonal matrix

$$A = \begin{pmatrix} a_1 & b_1 & \dots & 0 \\ c_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & b_{n-1} \\ 0 & \dots & c_{n-1} & a_n \end{pmatrix}$$

with spectral radius $r(A)$. Prove:

- (a) If $b_j c_j \geq 0$ for all j , then A has only real eigenvalues.

- (b) If $b_j c_j > 0$ for all j , then A has only real simple eigenvalues.
- (c) Assume $b_j > 0$, $c_j > 0$, and $a_j \geq 0$ for all j . Then A is irreducible. Matrix A is primitive if at least one $a_j > 0$. If $a_1 = \dots = a_n = 0$, then $-r(A)$ is an eigenvalue of A .
18. Let $\Gamma = (V = \{1, \dots, N\}, E)$ be a finite directed graph and $d_o(j) = |\mathcal{N}^o(j)| = |\{i \in V \mid (j, i) \in \Gamma\}|$ the out-degree of vertex j . For a real number $0 \leq \alpha < 1$ define the $N \times N$ **Google matrix** $\mathcal{G} = (g_{ij})$ of the digraph Γ as

$$g_{ij} := \begin{cases} \frac{\alpha}{d_o(j)} + \frac{1-\alpha}{N} & i \in \mathcal{N}^o(j) \neq \emptyset \\ \frac{1-\alpha}{N} & i \notin \mathcal{N}^o(j) \neq \emptyset \\ \frac{1}{N} & \mathcal{N}^o(j) = \emptyset. \end{cases}$$

- (a) Prove that \mathcal{G} is column stochastic and primitive,
- (b) Prove that the largest eigenvalue of \mathcal{G} is $\lambda_1 = 1$ and the second largest eigenvalue of \mathcal{G} is $\lambda_2 = \alpha$.
19. The **Leslie matrix** is a nonnegative matrix of the form

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & b_{n-1} & 0 \end{pmatrix}.$$

We assume $a_n > 0$ and $b_1 > 0, \dots, b_n > 0$.

- (a) Show that A is irreducible.
- (b) Show that A is primitive whenever there exists i with $a_i > 0$ and $a_{i+1} > 0$.
- (c) Show that A is not primitive if $n = 3$ and $a_1 = a_2 = 0$.
20. Prove that the Cayley–Menger determinants of a formation $x_1, \dots, x_N \in \mathbb{R}^m$ are nonpositive for $k \leq N$ and are zero for $k > m + 1$.

8.10 Notes and References

Classical references for nonnegative matrices, Markov chains, and the Perron–Frobenius theorem include Gantmacher (1959), Horn and Johnson (1990), and Seneta (1981). We also mention the excellent book by Fiedler (2008), which provides a useful collection of results on special matrices and connections to graph theory. Part of the material on stochastic matrices and the ergodic theorem in Section 8.3 was inspired by the book of Huppert (1990). Infinite-dimensional generalizations of the Perron–Frobenius theory can be found in the work of Jentzsch (1912), Krein and Rutman (1950), and Krasnoselskii (1964). A special case of the

Contraction Mapping Theorem 8.6 was applied by Pollicot and Yuri (1998) to prove the existence of a unique Perron vector for aperiodic $\{0, 1\}$ matrices. For positive matrices A , the existence of the Perron vector in Theorem 8.11 is well known and easily deduced from purely topological arguments. In fact, the standard simplex C_1 is homeomorphic to the closed unit ball, on which

$$x \mapsto \frac{Ax}{\mathbf{e}^\top Ax}$$

defines a continuous map. Thus the Brouwer fixed-point theorem implies the existence of an eigenvector $x \in C_1$ with positive eigenvalue. The papers by Bushell (1986) and Kohlberg and Pratt (1982) provide further background on the Hilbert metric and the analysis of positive operators. The sequence of power iterates (8.5) to compute the Perron vector is reminiscent of the well-known power method from numerical linear algebra,

$$x_{t+1} = \frac{Ax_t}{\|Ax_t\|},$$

for computing dominant eigenvectors of a matrix A . The convergence speed of the general power method depends on the ratio $\frac{|\lambda_1|}{|\lambda_2|}$ of the largest and second largest eigenvalues. See Parlett and Poole (1973) and Helmke and Moore (1994) for convergence proofs of the power method on projective spaces and Grassmann manifolds.

Graph-theoretic methods have long been used for studying the convergence properties of random walks on a graph, for analyzing synchronization and clustering phenomena in physical systems, and for algorithms in distributed computing, formation control, and networked control systems; we refer the reader to the monographies by Bullo, Cortés and Martínez (2009) and Meshbahi and Egerstedt (2010) for extensive background material and further references. Boyd, Diaconis and Xiao (2004) developed linear matrix inequalities characterizing the fastest Markov chain on a graph, while Xiao and Boyd (2004) studied linear iterations for distributed averaging and consensus in networks. Ottaviani and Sturmfels (2013) studied the problem of finding weights in a complete graph such that the associated Markov chain has a stationary probability distribution that is contained in a specified linear subspace. This problem is equivalent to characterizing the Laplacian matrices A of a graph such that the pair (C, A) is not observable. This problem seems widely open, but in a special situation (complete graph; weights are complex numbers), Ottaviani and Sturmfels (2013) successfully computed the degree of the variety of unobservable pairs.

The literature on formation shape control via distances and graph concepts includes early work by Olfati-Saber, Fax and Murray (2007) and Doerfler and Francis (2010). For characterizations of rigid graphs see Asimov and Roth (1978) and Connelly (1993). References on Euclidean distance geometry and applications include Crippen and Havel (1988), Dress and Havel (1993), and Blumenthal (1953).

The nonpositivity condition (8.10) for the Cayley–Menger determinants yields a simple determinantal condition that is necessary for a nonnegative symmetric matrix A with zero diagonal entries to be a Euclidean distance matrix. Blumenthal (1953), Chapter IV, p. 105, has shown that every such matrix A is a Euclidean distance matrix if and only if the Cayley–Menger determinants for all $k \times k$ principal submatrices of A are nonpositive, $k = 1, \dots, N$. The conditions of parts (d) and (e) in Theorem 8.29 are due to Gower (1985) and Schoenberg (1935), respectively.

A reference for the proof of Theorem 8.45 and related material is Grenander and Szegö (1958). The spectral properties of circulant matrices are well studied. An important fact is that the set of all circulant matrices is simultaneously diagonalized by the Fourier matrix (8.25). Further information on circulant matrices can be found in Davis (1979) and the recent survey by Kra and Simanca (2012). For a statement and proof of the Courant–Fischer minimax principle, see Horn and Johnson (1990). A generalization is Wielandt’s minimax theorem on partial sums of eigenvalues. The eigenvalue inequalities appearing in Theorem 8.57 are the simplest of a whole series of eigenvalue inequalities, which can be derived from eigenvalue inequalities on sums of Hermitian matrices. For a derivation of such eigenvalue inequalities via Schubert calculus on Grassmann manifolds, see Helmke and Rosenthal (1995). The full set of eigenvalue inequalities for sums of Hermitian matrices has been characterized by Klyachko; his work is nicely summarized by Fulton (2000). Such results should be useful in deriving sharp eigenvalue bounds for the Schur complement of Laplacian matrices.