Abdo Y. Alfakih

Euclidean Distance Matrices and Their Applications in Rigidity Theory

Springer

Euclidean Distance Matrices and Their Applications in Rigidity Theory

Abdo Y. Alfakih

Euclidean Distance Matrices and Their Applications in Rigidity Theory

Springer

Abdo Y. Alfakih
Department of Mathematics and Statistics
University of Windsor
Windsor, ON, Canada

ISBN 978-3-319-97845-1 ISBN 978-3-319-97846-8 (eBook)
https://doi.org/10.1007/978-3-319-97846-8

Library of Congress Control Number: 2018953759

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

In loving memory of my mother.

Preface

This monograph is devoted to a unified, up-to-date, and accessible exposition of Euclidean distance matrices (EDMs) and rigidity theory of bar-and-joint frameworks. EDMs, which comprise the first part of the monograph, are those matrices whose entries can be realized as the squared Euclidean interpoint distances of a point configuration. Such matrices arise in many areas of science and engineering including statistics, computational biochemistry, and computer science, to name a few. The second part of the monograph focuses on rigidity theory of bar-and-joint frameworks. Given a subset E of the interpoint distances of a point configuration, rigidity theory is concerned with the existence of a second point configuration having the same interpoint distances as those of E. Various rigidity notions correspond to various conditions on the second point configuration. Rigidity theory has a long and rich history going at least as far back as Cauchy (1813) and is of great interest to structural engineers and mathematicians. EDMs and rigidity theory fall in the general area of distance geometry. Distance geometry has important applications in statistics (multidimensional scaling [48]), computational biochemistry (molecular conformations [66]), and computer science (sensor networks).

The last four decades have seen a growing body of literature on EDMs and rigidity theory. Much of this literature, unfortunately, is available mainly in scattered form in journals of various disciplines. This, coupled with the lack of a unified notation, makes it difficult to get a firm grasp of the published literature and acts as a barrier to new researchers entering the field. This monograph is an attempt to rectify this situation by presenting a unified account of EDMs and rigidity theory based on the one-to-one correspondence between EDMs and projected Gram matrices. Accordingly, the machinery of semidefinite programming is a common thread that runs throughout the monograph. As a result, two parallel approaches to rigidity theory are presented. The first one is traditional and more intuitive and is based on a vector representation of a point configuration. The second one is novel and less intuitive and is based on a Gram matrix representation of a point configuration. Each of these two approaches, obviously, has its advantages and disadvantages.

The monograph is self-contained and should be accessible to a wide audience including students and researchers in statistics, computational biochemistry, engineering, computer science, operations research, and mathematics. The notation used here is standard in the semidefinite programming literature. Chapters 1 and 2 provide the necessary background for the rest of the chapters. The focus of Chap. 1 is on pertinent results from matrix theory, graph theory, and convexity theory, while Chap. 2 is devoted entirely to positive semidefinite (PSD) matrices due to the key role these matrices play in our approach. Chapters 3–7 are devoted to a detailed study of EDMs, and in particular their various characterizations, classes, eigenvalues, entries, and geometry. Chapters 9 and 10 are devoted to local and universal rigidities of bar-and-joint frameworks. The literature on rigidity theory is vast. We chose to include only those two notions of rigidity because they lend themselves easily to semidefinite programming machinery used throughout the monograph. Moreover, due to space limitation, we discuss only the most significant results and results directly relevant to EDMs. Finally, Chap. 8 is a transitional chapter that links rigidity theory to EDMs by viewing various rigidity problems as EDM completion uniqueness problems.

Finally, I would like to express my sincere thanks and gratitude to Katta G. Murty (thesis advisor) and Henry Wolkowicz (postdoctoral advisor). Murty introduced me to Henry and Henry introduced me to EDMs.

Windsor, ON, Canada Abdo Y. Alfakih

Contents

List of Notation

$\mathbf{0}$	The zero vector or matrix of appropriate dimensions
$\lvert S \rvert$	The cardinality of set S
$\langle .,. \rangle$	Inner product
T^*	The adjoint of linear transformation T
$T^{-1}(S)$	The preimage of set S under linear transformation T
$\mathscr{L}^{\perp}$	The orthogonal complement of subspace $\mathscr{L}$
$\lvert\lvert x \rvert\rvert$	The norm of x
x^i	The ith vector in a set of vectors $x^1,\ldots,x^k$
x_i	The ith component of vector x
e_n^i	The standard ith unit vector in $\mathbb{R}^n$
e_n	The vector of all 1's in $\mathbb{R}^n$
V	The $n \times (n-1)$ matrix whose columns form an orthonormal basis of $e_n^{\perp}$
I_n	The identity matrix of order n
E_n	$= e_n e_n^T$
$E_{n,m}$	$= e_n e_m^T$
J	$= I_n - E_n/n$
$E^{ij}, (i \neq j)$	$= e_n^i (e_n^j)^T + e_n^j (e_n^i)^T$
$M^{ij}, (i \neq j)$	$= -V^T E^{ij} V/2$
$F^{ij}, (i \neq j)$	$= (e^i - e^j)(e^i - e^j)^T$
$\mathscr{S}^n$	The space of $n \times n$ real symmetric matrices
$\mathscr{S}_+^n$	The set of $n \times n$ real symmetric positive semidefinite matrices
$\mathscr{S}_{++}^n$	The set of $n \times n$ real symmetric positive definite matrices
$\mathscr{D}^n$	The set of $n \times n$ Euclidean distance matrices
$A_{.j}$	The jth column of matrix A
$A_{i.}$	The ith row of matrix A
$\mathrm{Diag}(x)$	The diagonal matrix formed by vector x
$\mathrm{diag}(A)$	The vector consisting of the diagonal entries of matrix A
$\chi_A(\lambda)$	The characteristic polynomial of matrix A
$m_A(\lambda)$	The minimal polynomial of matrix A
$\lvert\lvert A \rvert\rvert_F$	The Frobenius norm of matrix A

$\rho(A)$	The spectral radius of a square matrix A
$\text{null}(A)$	The null space of matrix A
$\text{col}(A)$	The column space of matrix A
$\text{gal}(D)$	The Gale space of EDM D
$\bar{r} = n - 1 - r$	The dimension of $\text{gal}(D)$
$A^\dagger$	The Moore–Penrose inverse of matrix A
$A \circ B$	The Hadamard product of matrices A and B
$A \otimes B$	The Kronecker product of matrices A and B
$A \otimes_s A$	The symmetric Kronecker product of matrix A and itself
$\mathscr{L}_1 \oplus \mathscr{L}_2$	The direct sum of subspaces $\mathscr{L}_1$ and $\mathscr{L}_2$
$V(G)$	The vertex set of graph G
$E(G)$	The edge set of simple graph G of cardinality m
$\bar{E}(\bar{G})$	The edge set of the complement graph $\bar{G}$, or the set of missing edges of G, of cardinality $\bar{m}$
$\mathscr{E}(y)$	$= \sum_{\{i,j\} \in \bar{E}(\bar{G})} y_{ij} E^{ij}$
$\mathscr{M}(y)$	$= \sum_{\{i,j\} \in \bar{E}(\bar{G})} y_{ij} M^{ij}$
$\deg(i)$	The degree of node i of a graph
deg	The vector consisting of the degrees of all nodes of a graph
K_n	The complete graph on n nodes
conv(S)	The convex hull of set S
aff(S)	The affine hull of set S
int(S)	The interior of set S
relint(S)	The relative interior of set S
cl(S)	The closure of set S
∂S	The boundary of set S
rbd(S)	The relative boundary of set S
K°	The polar of cone K
$\setminus$	The set theoretic difference
$N_S(\hat{x})$	The normal cone of set S at point $\hat{x}$
$T_S(\hat{x})$	The tangent cone of set S at point $\hat{x}$
face(x, S)	The minimal face of set S containing x
$A \succ \mathbf{0}$	Real matrix A is symmetric positive definite
$A \succeq \mathbf{0}$	Real matrix A is symmetric positive semidefinite

Chapter 1
Mathematical Preliminaries

In this chapter, we briefly review some of the mathematical preliminaries that will be needed throughout the monograph. These include a brief review of the most pertinent concepts and results in the theories of vector spaces, matrices, convexity, and graphs. Proofs of several of these results are included to make this chapter as self-contained as possible.

1.1 Vector Spaces

The notion of a vector space plays an important role in Euclidean geometry. In this monograph we are interested only in finite-dimensional real vector spaces.

Let $\mathscr{V}$ be a nonempty set equipped with the operations of addition and scalar multiplication. Then $\mathscr{V}$ is a real vector space (or a real linear space) if the following conditions are satisfied:

1. $x+y=y+x$ for all x,y in $\mathscr{V}$.
2. $x+(y+z)=(x+y)+z$ for all x,y,z in $\mathscr{V}$.
3. There exists a unique $\mathbf{0}\in\mathscr{V}$ such that $x+\mathbf{0}=x$ for all $x\in\mathscr{V}$.
4. For each $x\in\mathscr{V}$, there exists a unique $(-x)\in\mathscr{V}$ such that $x+(-x)=\mathbf{0}$.
5. $(\alpha+\beta)x=\alpha x+\beta x$ for all α,β in $\mathbb{R}$ and all $x\in\mathscr{V}$.
6. $\alpha(x+y)=\alpha x+\alpha y$ for all α in $\mathbb{R}$ and all x,y in $\mathscr{V}$.
7. $(\alpha\beta)x=\alpha(\beta x)$ for all α,β in $\mathbb{R}$ and all x in $\mathscr{V}$.
8. $1x=x$ for all $x\in\mathscr{V}$.

The vector spaces of interest to us are the ones where $\mathscr{V}=\mathbb{R}^n$ and $\mathscr{V}=\mathscr{S}^n$, the set of $n\times n$ real symmetric matrices. The elements of real vector space $\mathscr{V}$ are called *vectors*. If the origin $\mathbf{0}$ is of no particular interest to us, then the elements of $\mathscr{V}$ are called *points*. Let $\mathscr{V}$ be a real vector space and let $\mathscr{V}'\subset\mathscr{V}$. If $\mathscr{V}'$ is a real vector space in its own right, then $\mathscr{V}'$ is called a *linear subspace*, or a subspace for short, of $\mathscr{V}$. It is easy to see that a nonempty subset of $\mathscr{V}$ is a subspace of $\mathscr{V}$ iff it is closed under linear combinations.

A.Y. Alfakih, *Euclidean Distance Matrices and Their Applications in Rigidity Theory*, https://doi.org/10.1007/978-3-319-97846-8_1

An *inner product* on a real vector space $\mathscr{V}$, denoted by $\langle .,. \rangle$, is a real-valued function on $\mathscr{V} \times \mathscr{V}$ that satisfies the following properties:

1. $\langle x,x \rangle \geq 0$ for all $x \in \mathscr{V}$ and $\langle x,x \rangle = 0$ iff $x = \mathbf{0}$.
2. $\langle \alpha x + \beta y, z \rangle = \alpha \langle x,z \rangle + \beta \langle y,z \rangle$ for all x,y,z in $\mathscr{V}$ and all α, β in $\mathbb{R}$.
3. $\langle x,y \rangle = \langle y,x \rangle$ for all x,y in $\mathscr{V}$.

A vector space on which an inner product is defined is called an *inner product space*.

A *norm* on a real vector space $\mathscr{V}$, denoted by $||x||$, is a function $\mathscr{V} \to \mathbb{R}$ that satisfies the following properties:

1. $||x|| \geq 0$ for all $x \in \mathscr{V}$, and $||x|| = 0$ iff $x = \mathbf{0}$.
2. $||\alpha x|| = |\alpha| \, ||x||$ for all x in $\mathscr{V}$ and all α in $\mathbb{R}$.
3. $||x+y|| \leq ||x|| + ||y||$ for all x,y in $\mathscr{V}$.

A vector space equipped with a norm is called a *normed vector space*. Every inner product naturally induces a norm of the form $||x|| = \langle x,x \rangle^{1/2}$. Our interest in this monograph is in the norm in $\mathbb{R}^n$ induced by $\langle x,y \rangle = x^T y$ and the norm in $\mathscr{S}^n$ induced by $\langle X,Y \rangle = \text{trace}(XY)$.

Theorem 1.1 (Cauchy–Schwarz inequality) *Let x and y be two vectors in a real vector space $\mathscr{V}$ equipped with inner product $\langle .,. \rangle$. Then*

$$|\langle x,y \rangle| \leq \langle x,x \rangle^{1/2} \langle y,y \rangle^{1/2},$$

where equality holds if and only if $x - \alpha y = \mathbf{0}$ for some scalar α.

Proof. Let $||x||^2 = \langle x,x \rangle$. Then, it follows from the definition of inner product that $\langle x - ty, x - ty \rangle = t^2 ||y||^2 - 2t \langle x,y \rangle + ||x||^2 \geq 0$ for all $t \in \mathbb{R}$. Now if $y = \mathbf{0}$, then the result follows trivially. Thus, assume that $||y||^2 \neq 0$ and let

$$\hat{t} = \frac{\langle x,y \rangle}{||y||^2}.$$

Then

$$\langle x - \hat{t}y, x - \hat{t}y \rangle = ||x||^2 - \frac{\langle x,y \rangle^2}{||y||^2} \geq 0.$$

Consequently, $\langle x,y \rangle^2 \leq ||y||^2 ||x||^2$, with equality iff $x - \hat{t}y = \mathbf{0}$.

□

Cauchy–Schwarz inequality is used to establish the continuity of the inner product.

Lemma 1.1 *Let $\{x^k\}_{k \in \mathbb{N}}$, $\{y^k\}_{k \in \mathbb{N}}$ be two sequences in $\mathscr{V}$ that converge to x and y, respectively. Then*

$$\lim_{k \to \infty} \langle x^k, y^k \rangle = \langle x,y \rangle.$$

That is, the inner product $\langle x,y \rangle$ is a continuous function.

Proof.

$$\begin{aligned}
|\langle x^k, y^k\rangle - \langle x,y\rangle| &= |\langle x^k, y^k\rangle - \langle x^k, y\rangle + \langle x^k, y\rangle - \langle x,y\rangle| \\
&= |\langle x^k, (y^k - y)\rangle + \langle (x^k - x), y\rangle| \\
&\leq |\langle x^k, (y^k - y)\rangle| + |\langle (x^k - x), y\rangle| \\
&\leq ||x^k||\, ||y^k - y|| + ||x^k - x||\, ||y||,
\end{aligned}$$

where the last inequality follows from Cauchy–Schwarz inequality. Now $||x^k||$ is bounded by the convergence of $\{x^k\}$. Thus, $\langle x^k, y^k\rangle \to \langle x,y\rangle$ as $x^k \to x$ and $y^k \to y$. □

Let $\mathscr{V}_1$ and $\mathscr{V}_2$ be two vector spaces and let $T : \mathscr{V}_1 \to \mathscr{V}_2$ be a linear transformation. The *adjoint* of T, denoted by T^*, is the unique transformation $T^* : \mathscr{V}_2 \to \mathscr{V}_1$ that satisfies

$$\langle y, T(x)\rangle = \langle T^*(y), x\rangle \text{ for all } x \in \mathscr{V}_1 \text{ and for all } y \in \mathscr{V}_2.$$

For example, let $\text{Diag}(x)$ denote the diagonal matrix formed by vector x and let $\text{diag}(A)$ denote the vector consisting of the diagonal entries of matrix A. Let $\mathbb{R}^n$ be endowed with inner product $\langle x,y\rangle = x^T y$ and let $\mathscr{S}^n$ be endowed with the trace inner product $\langle A,B\rangle = \text{trace}(AB)$. Further, let $T : \mathbb{R}^n \to \mathscr{S}^n$, where $T(x) = \text{Diag}(x)$. Then for any $A \in \mathscr{S}^n$, we have

$$\text{trace}(A\text{Diag}(x)) = \sum_{i=1}^{n} a_{ii} x_i = x^T \text{diag}(A).$$

Therefore, the adjoint of T is $T^* : \mathscr{S}^n \to \mathbb{R}^n$, where $T^*(A) = \text{diag}(A)$.

1.2 Matrix Theory

In this monograph we deal only with real matrices. Let A be an $n \times n$ matrix. The matrix obtained from A by deleting $n-k$ rows and $n-k'$ columns, where $1 \leq k, k' \leq n$, is a $k \times k'$ *submatrix* of A. A *principal submatrix* of A is the square submatrix obtained from A by deleting similarly indexed rows and columns; i.e., if the ith row of A is deleted, then so is the ith column. The determinant of a principal submatrix of A is called a *principal minor* of A. The *kth leading principal submatrix* of A is the square submatrix obtained by deleting the last $n-k$ columns and rows of A. Note that the nth leading principal submatrix of A is A itself. The determinant of the kth leading principal submatrix of A is called the *kth leading principal minor* of A. It easily follows that an $n \times n$ matrix has $2^n - 1$ principal minors and n leading principal minors.

1.2.1 The Characteristic and the Minimal Polynomials

Let A be an $n \times n$ matrix and let x be a nonzero vector in $\mathbb{R}^n$. Then x is said to be an *eigenvector* of A if

$$Ax = \lambda x,$$

for some scalar λ, in which case, λ is said to be the *eigenvalue* of A corresponding to x. The pair (λ, x) is called an *eigenpair* of A. An immediate consequence of this definition is that the eigenvalues of A are the roots of the polynomial

$$\chi_A(\lambda) = \det(A - \lambda I). \tag{1.1}$$

$\chi_A(\lambda)$ is called the *characteristic polynomial* of A. An important fact to bear in mind is that A and A^T have the same characteristic polynomial and hence the same eigenvalues. Since $\chi_A(\lambda)$ is of degree n, it follows that A has n eigenvalues some of which may be complex even if A is real. On the other hand, A may or may not have n linearly independent eigenvectors. A is said to be *diagonalizable* if there exists a nonsingular matrix S such that

$$A = S\Lambda S^{-1},$$

where Λ is the diagonal matrix consisting of the eigenvalues of A, in which case, the columns of S are the eigenvectors of A, and the rows of S^{-1} are the eigenvectors of A^T. It is easy to see that A is diagonalizable if and only if it has n linearly independent eigenvectors. As will be shown next, real symmetric matrices are always diagonalizable, and more importantly, they are diagonalizable by orthogonal matrices.

Theorem 1.2 *Let A be an $n \times n$ real symmetric matrix. Then A has n real eigenvalues and n orthonormal eigenvectors.*

Proof. Let $Ax = \lambda x$. Then $A\bar{x} = \bar{\lambda}\bar{x}$, where $\bar{x}$ is the complex conjugate of x. Therefore, $\bar{x}^T Ax - x^T A\bar{x} = (\lambda - \bar{\lambda})\, \bar{x}^T x = 0$. Thus, $\bar{\lambda} = \lambda$ since $\bar{x}^T x \neq 0$. Hence, λ is real and thus x can be chosen real.

Now let $Ax^1 = \lambda_1 x^1$ and $Ax^2 = \lambda_2 x^2$, where $\lambda_1 \neq \lambda_2$. Then ${x^2}^T Ax^1 - {x^1}^T Ax^2 = (\lambda_1 - \lambda_2)\, {x^2}^T x^1 = 0$. Thus ${x^2}^T x^1 = 0$. Hence, the eigenvectors of A corresponding to distinct eigenvalues are orthogonal. Now if an eigenvalue λ of A is repeated, then the eigenvectors corresponding to λ can be chosen to be orthogonal.

□

Theorem 1.3 (The Spectral Theorem) *Let A be a real $n \times n$ matrix. Then A is symmetric if and only if*

$$A = Q\Lambda Q^T, \tag{1.2}$$

where Λ is the diagonal matrix consisting of the eigenvalues of A and Q is an orthogonal matrix whose columns are the corresponding eigenvectors.

Equation (1.2) is called the *spectral decomposition* of A. Let $(\lambda_1, q^1), \ldots, (\lambda_n, q^n)$ be the eigenpairs of A. Then Eq. (1.2) can be written as $A = \sum_{i=1}^n \lambda_i q^i (q^i)^T$.

Rayleigh–Ritz Theorem gives a variational characterization of the largest and the smallest eigenvalues of a real symmetric matrix.

Theorem 1.4 (Rayleigh–Ritz) *Let A be an $n \times n$ real symmetric matrix and let $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of A. Further, let x^1 and x^n be eigenvectors of A corresponding to λ_1 and λ_n, respectively. Then*

$$\lambda_1 = \max_{x \neq \mathbf{0}} \frac{x^T A x}{x^T x} \text{ and } \lambda_n = \min_{x \neq \mathbf{0}} \frac{x^T A x}{x^T x}.$$

Moreover, the maximum is attained at x^1 and the minimum is attained at x^n.

Proof. We only present a proof of the maximum case. The proof of the minimum case is similar. Let $A = Q\Lambda Q^T$ be the spectral decomposition of A. Then for all $x \neq \mathbf{0}$ we have

$$\frac{x^T A x}{x^T x} = \frac{y^T \Lambda y}{y^T y} = \frac{\sum_{i=1}^n \lambda_i y_i^2}{\sum_{i=1}^n y_i^2} \leq \lambda_1$$

since $\sum_{i=1}^n \lambda_i y_i^2 \leq \sum_{i=1}^n \lambda_1 y_i^2$. The result follows since $(x^1)^T A x^1 = \lambda_1 (x^1)^T x^1$.

□

We will find the following corollary of Rayleigh–Ritz Theorem useful in later chapters.

Corollary 1.1 *Let A be an $n \times n$ real symmetric matrix. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A with corresponding orthonormal eigenvectors $q^1, \ldots, q^n$. Assume that $\lambda_1 > \lambda_i$ for all $i = 2, \ldots, n$ and let x be a unit vector such that $x^T A x = \lambda_1$. Then $x = \pm q^1$.*

Proof. Let $x = \sum_{i=1}^n \alpha_i q^i$. Then $\sum_{i=1}^n \alpha_i^2 = 1$ and $\lambda_1 = \sum_{i=1}^n \alpha_i^2 \lambda_i$. Hence,

$$\sum_{i=2}^n \alpha_i^2 \lambda_i = \lambda_1 (1 - \alpha_1^2) = \lambda_1 \sum_{i=2}^n \alpha_i^2.$$

Therefore, $\sum_{i=2}^n \alpha_i^2 (\lambda_1 - \lambda_i) = 0$. But $(\lambda_1 - \lambda_i) > 0$ for all $i = 2, \ldots, n$. Hence, $\alpha_2 = \cdots = \alpha_n = 0$ and thus $x = \pm q^1$.

□

Theorem 1.5 *Let A be a real symmetric $n \times n$ matrix and let $\mathscr{L}$ be a k-dimensional subspace of $\mathbb{R}^n$ such that $x^T A x \leq 0$ for all $x \in \mathscr{L}$. Then A has at least k nonpositive eigenvalues.*

Proof. Let $q^1, \ldots, q^n$ be orthonormal eigenvectors of A with corresponding eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. Let $S = \text{span}\ \{q^1, \ldots, q^{n-k+1}\}$. Then $\dim(S) = n - k + 1$. Thus $\mathscr{L} \cap S \neq \emptyset$ and hence let x be a unit vector in $\mathscr{L} \cap S$. Hence, $\lambda_{n-k+1} \leq x^T A x \leq 0$. Therefore, $\lambda_n \leq \cdots \leq \lambda_{n-k+1} \leq 0$.

□

The inertia of a real symmetric matrix A is the ordered triple (n_+, n_-, n_0), where n_+, n_-, and n_0 are, respectively, the numbers of positive, negative, and zero eigenvalues of A. Thus, $\text{rank}(A) = n_+ + n_-$.

Theorem 1.6 (Sylvester Law of Inertia) *Let A be an $n \times n$ real symmetric matrix and let S be a nonsingular $n \times n$ matrix. Then A and SAS^T have the same inertia.*

Theorem 1.7 (Cauchy Interlacing Theorem) *Let $\mu_1 \geq \cdots \geq \mu_n$ be the eigenvalues of an $n \times n$ real symmetric matrix A. Let B be any $(n-1) \times (n-1)$ principal submatrix of A and let $\lambda_1 \geq \cdots \geq \lambda_{n-1}$ be the eigenvalues of B. Then the eigenvalues of A are interlaced by those of B; i.e.,*

$$\mu_k \geq \lambda_k \geq \mu_{k+1} \text{ for } k = 1, \ldots, n-1.$$

The coefficients of the characteristic polynomial can be expressed in terms of the principal minors.

Theorem 1.8 *Let A be an $n \times n$ matrix and let c_k be the coefficient of λ^k in $\chi_A(\lambda)$, the characteristic polynomial of A. Then for $k \leq n-1$, we have*

$$c_k = (-1)^k \times \text{the sum of all principal minors of } A \text{ of order } n-k.$$

Proof. Let e^i be the ith standard unit vector in $\mathbb{R}^n$ and let $A_{.j}$ be the jth column of A. Then

$$\chi_A(\lambda) = \det(A - \lambda I) = \det(\ [(A_{.1} - \lambda e^1)\ \ (A_{.2} - \lambda e^2)\ \ \cdots\ \ (A_{.n} - \lambda e^n)\ ]). \quad (1.3)$$

Since the determinant is linear in each column separately, the coefficient of λ^k in (1.3) is

$$c_k = (-1)^k \sum \det(\ [x(1)\ x(2) \ldots x(k)\ x(k+1) \ldots x(n)]\ ), \quad (1.4)$$

where the sum is taken over all possible ways to replace $x(j)$ with $A_{.j}$ or e^j such that the total number of e^j's is k. For instance, $\det([e^1\ e^2 \ldots e^k\ A_{.k+1} \ldots A_{.n}])$ is one of the terms in the sum in (1.4). But this term is precisely the principal minor of A of order $n-k$, obtained by deleting the first k rows and columns. Hence, the sum in (1.4) is the sum of all principal minors of A of order $n-k$.

□

The coefficient c_n is called the *leading coefficient* of $\chi_A(\lambda)$ and it is equal to $(-1)^n$. Also, Theorem 1.8 immediately implies that $c_0 = \det(A)$, $c_{n-1} = (-1)^{n-1}$ trace(A). Note that the determinant and the trace are equal, respectively, to the product and the sum of the eigenvalues counting multiplicities.

Theorem 1.9 (Cayley–Hamilton) *Every square matrix satisfies its characteristic polynomial.*

The *geometric multiplicity* of an eigenvalue λ is equal to the maximum number of linearly independent eigenvectors corresponding to λ. On the other hand, the *algebraic multiplicity* of λ is equal to the number of times λ is repeated as a root of the characteristic polynomial. Note that the geometric multiplicity of an eigenvalue is always less than or equal to its algebraic multiplicity. Moreover, matrix A is diagonalizable if and only if, for each eigenvalue of A, the geometric multiplicity is equal to the algebraic multiplicity.

Let $\lambda_1,\dots,\lambda_k$ be the distinct eigenvalues of A with respective algebraic multiplicities m_1, …, m_k. Then

$$\chi_A(\lambda) = (\lambda_1 - \lambda)^{m_1} \dots (\lambda_k - \lambda)^{m_k}.$$

Hence, Cayley–Hamilton Theorem implies that

$$\chi_A(A) = (\lambda_1 I - A)^{m_1} \dots (\lambda_k I - A)^{m_k} = \mathbf{0}.$$

A polynomial is called *monic* if its leading coefficient is 1. The *minimal polynomial* of A, denoted by $m_A(\lambda)$, is the smallest degree monic polynomial that annihilates A, i.e., $m_A(A) = \mathbf{0}$. Consequently,

$$m_A(\lambda) = (\lambda_1 - \lambda)^{r_1} \dots (\lambda_k - \lambda)^{r_k},$$

where $r_i \leq m_i$ for all $i = 1,\dots,k$.

Theorem 1.10 *Let A be an $n \times n$ matrix. Then A is diagonalizable if and only if its minimal polynomial, $m_A(\lambda)$, is the product of linear terms, i.e., iff $r_1 = \cdots = r_k = 1$.*

The norm of matrix A, denoted by $||A||$, is a real-valued function that satisfies the following three properties: (i) $||A|| \geq 0$ for all A and $||A|| = 0$ iff $A = \mathbf{0}$, (ii) $||\alpha A|| = |\alpha|\ ||A||$ for all A and for all scalars α, and (iii) $||A+B|| \leq ||A|| + ||B||$ for all A and B. In addition, if a matrix norm satisfies the property that $||AB|| \leq ||A||\ ||B||$ for all A and B, then this norm is said to be *consistent* or *submultiplicative*.

The *Frobenius norm* of an $m \times n$ real matrix A, denoted by $||A||_F$, is defined by $||A||_F = \sqrt{\text{trace}(A^T A)}$. It is not hard to show that the Frobenius norm is submultiplicative. Every vector norm $||x||$ *induces* a matrix norm as follows:

$$||A|| = \max_{x \neq \mathbf{0}} \frac{||Ax||}{||x||}.$$

Thus, $||A||\,||x|| \geq ||Ax||$ for any x. Accordingly, every induced matrix norm is submultiplicative since

$$||ABx|| \leq ||A||\,||B||\,||x|| \text{ for any } x \in \mathbb{R}^n.$$

Furthermore, it follows from Rayleigh–Ritz Theorem that $||A||_2$, the matrix norm induced by the Euclidean vector norm, is given by $||A||_2 = \sqrt{\lambda_{\max}(A^T A)}$. Consequently, $||A||_2 \leq ||A||_F$ for any matrix A.

1.2.2 The Perron Theorem

A vector x in $\mathbb{R}^n$ is said to be *positive*, denoted by $x > \mathbf{0}$, if $x_i > 0$ for all $i = 1,\dots,n$. Similarly, an $n \times n$ real matrix A is said to be *positive (nonnegative)*, denoted by $A > \mathbf{0}\ (\geq \mathbf{0})$, if $a_{ij} > 0\ (\geq 0)$ for all $i, j = 1,\dots,n$. A nonnegative matrix A is said to be *primitive* if $A^k > 0$ for some positive integer k. Clearly, positive matrices form a subset of primitive matrices. The *spectral radius* of A is $\rho(A) = \max\{|\lambda_i| : \lambda_i$ is an eigenvalue of $A\}$.

Theorem 1.11 (Perron) *Let A be an $n \times n$ primitive matrix and let $\rho(A)$ be its spectral radius. Then*

1. *There exists an eigenvalue $\lambda_1 = \rho(A)$ with a corresponding eigenvector $x^1 > \mathbf{0}$.*
2. *λ_1 has algebraic multiplicity 1.*
3. *x^1 is the only positive eigenvector of A.*
4. *$\lambda_1 > |\lambda|$ for all eigenvalues $\lambda \neq \lambda_1$ of A.*

The eigenpair (λ_1, x^1) is called the Perron eigenpair.

To keep the proof simple, we assume that matrix A is symmetric. Also, we assume that A is positive. We comment later on the proof when A is primitive.

Proof. Assume that A is symmetric and positive. Thus $\rho(A) > 0$. Let $\lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of A. Notice that $\lambda_1 > 0$ since trace$(A) > 0$. Therefore, $\rho(A)$ is either equal to λ_1 or $|\lambda_n|$. Let y^1 and y^n be the normalized eigenvectors corresponding to λ_1 and λ_n, respectively, and let $x^1 = |y^1|$, i.e., $x^1_i = |y^1_i|$ for $i = 1, \ldots, n$. Then

$$|\lambda_n| = |(y^n)^T A y^n| = |\sum_j a_{ij} y^n_i y^n_j| \leq \sum_j a_{ij} |y^n_i|\,|y^n_j| \leq \lambda_1,$$

where the last inequality follows from Rayleigh–Ritz Theorem. Thus $\lambda_1 = \rho(A)$. Moreover,

$$\lambda_1 = |\lambda_1| = |(y^1)^T A y^1| \leq (x^1)^T A x^1 \leq \lambda_1.$$

Therefore, $(x^1)^T A x^1 = \lambda_1$ and thus $Ax^1 = \lambda_1 x^1$. Furthermore, since $x^1 \geq \mathbf{0}$ and since $x^1 = Ax^1/\lambda_1$, it follows that $x^1 > 0$. This proves Statement 1.

To prove Statement 2, assume that there exists y such that $Ay = \lambda_1 y$. Let

$$\alpha = \min\{\frac{x^1_i}{y_i} : y_i > 0\} = \frac{x^1_{i_0}}{y_{i_0}} > 0.$$

Let $z = x^1 - \alpha y$. Then $z \geq \mathbf{0}$ and $z_{i_0} = 0$. Assume that $z \neq \mathbf{0}$. Then $Az = \lambda_1 z$ and hence, $z = Az/\lambda_1$ must be $> \mathbf{0}$ since $z \geq \mathbf{0}$, a contradiction. Therefore, $z = \mathbf{0}$ and y is a multiple of x^1 and hence the geometric multiplicity of λ_1 is 1. Statement 2, thus, follows since A is diagonalizable.

Statement 3 follows from the fact that eigenvectors corresponding to distinct eigenvalues of a real symmetric matrix are orthogonal. Indeed, assume, to the contrary, that there exists an eigenpair (λ, y), where $\lambda \neq \lambda_1$ and $y > \mathbf{0}$. Thus, on one hand $y^T x^1 > 0$, and on the other hand, y and x^1 are orthogonal, a contradiction.

To prove Statement 4, it suffices to prove that $\lambda_n \neq -\lambda_1$. To this end, assume to the contrary that $Ay^n = -\lambda_1 y^n$. Then $(y^n)^T x^1 = 0$. Moreover, $A^2 y = \lambda_1^2 y$ and $A^2 x^1 = \lambda_1^2 x^1$. Thus, for the symmetric positive matrix A^2, the algebraic multiplicity of the Perron eigenvalue λ_1^2 is 2, a contradiction.

□

The proof of the case where A is primitive uses the above proof applied to the symmetric positive matrix A^k. It also uses the following facts. First, if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, then $\lambda_1^k, \ldots, \lambda_n^k$ are the eigenvalues of A^k. Consequently, if $|\lambda_1|^k \geq \cdots \geq |\lambda_n|^k$, then $|\lambda_1| \geq \cdots \geq |\lambda_n|$. Second, since the algebraic multiplicity of λ_1^k is 1, and thus its geometric multiplicity is also 1, it follows that the Perron eigenvector x^1 of A^k is also a Perron eigenvector of A. Furthermore, $\lambda_1 > 0$ since $Ax^1 = \lambda_1 x^1$.

1.2.3 The Null Space, the Column Space, and the Rank

Let A be an $m \times n$ real matrix. The *null space* of A is $\text{null}(A) = \{x \in \mathbb{R}^n : Ax = \mathbf{0}\}$ and the *column space* of A is $\text{col}(A) = \{y \in \mathbb{R}^m : y = Az \text{ for some } z \text{ in } \mathbb{R}^n\}$. Both null$(A)$ and col(A) are, respectively, subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$ of dimensions $n - \text{rank}(A)$ and rank(A). The subspace null(A^T) is often called the *left null space* of A.

Every k-dimensional subspace $\mathscr{L}$ of $\mathbb{R}^n$ can be represented either as the column space of an $n \times k$ matrix A, or as the null space of an $(n-k) \times n$ matrix B. The columns of A form a basis of $\mathscr{L}$, while the rows of B form a basis of $\mathscr{L}^{\perp}$, the orthogonal complement of $\mathscr{L}$ in $\mathbb{R}^n$.

The *Moore–Penrose inverse* of A, denoted by $A^{\dagger}$, is the unique matrix that satisfies: (i) $AA^{\dagger}A = A$, (ii) $A^{\dagger}AA^{\dagger} = A^{\dagger}$, (iii) $(A^{\dagger}A)^T = A^{\dagger}A$, and (iv) $(AA^{\dagger})^T = AA^{\dagger}$. Obviously, if A is nonsingular, then $A^{\dagger} = A^{-1}$. The following two facts are easy to verify. First, if A has full column rank, then $A^{\dagger} = (A^TA)^{-1}A^T$. Second, if $A = Q\Lambda Q^T$ where Q is orthogonal, then $A^{\dagger} = Q\Lambda^{\dagger}Q^T$.

A matrix P satisfying $P^2 = P$ is called a *projection matrix*. If such P is symmetric, then it is called an *orthogonal projection matrix*. Otherwise, it is called an *oblique projection matrix*. It easily follows that $AA^{\dagger}$ is the orthogonal projection matrix onto col(A). Notice that $AA^{\dagger}$ is symmetric. Thus, if the system of equations $Ax = b$ is consistent, i.e., if $AA^{\dagger}b = b$, then $x = A^{\dagger}b + (I - A^{\dagger}A)z$, where z is an arbitrary vector, is a solution of this system since $Ax = AA^{\dagger}b = b$.

The following technical result will be used repeatedly in this monograph.

Proposition 1.1 *Let A and B be two real symmetric square matrices such that $AB = \mathbf{0}$. Further, assume that A is singular and let U be the matrix whose columns form an orthonormal basis of null(A). Then $B = U\Phi U^T$, where $\Phi = U^TBU$.*

Proof. Let $A = W\Lambda W^T$ be the spectral decomposition of A, where Λ is the diagonal matrix consisting of the nonzero eigenvalues of A. Therefore, $W^TB = \mathbf{0}$ since $W\Lambda$ has full column rank. Let $Q = [W\ U]$. Then Q is orthogonal and $Q^TBQ = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Phi \end{bmatrix}$, where $\Phi = U^TBU$. Consequently, $B = U\Phi U^T$.

□

Let A and B be two $m \times n$ matrices. Then $\text{col}([A\ B]) = \text{col}(A) + \text{col}(B)$. Hence, dim $(\text{col}([A\ B]))$ = $\dim(\text{col}(A))$ + $\dim(\text{col}(B))$ − $\dim(\text{col}(A) \cap \text{col}(B))$. Therefore,

$$\text{rank}([A\ B]) = \text{rank}(A) + \text{rank}(B) - \dim(\text{col}(A) \cap \text{col}(B)). \tag{1.5}$$

On the other hand, $\text{col}(A+B) = \{y : y = (A+B)z \text{ for some } z \in \mathbb{R}^n\} = \{y : y = [A\ B]\begin{bmatrix} z \\ z \end{bmatrix} \text{ for some } z \in \mathbb{R}^n\} \subseteq \text{col}([A\ B])$. Therefore,

$$\text{rank}(A+B) \leq \text{rank}([A\ B]). \tag{1.6}$$

As a result, it follows from Eqs. (1.5) and 1.6 that

$$\text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B). \tag{1.7}$$

The following theorem establishes a necessary and sufficient condition for equality to hold in Eq. (1.7).

Theorem 1.12 (Marsaglia and Styan [140]) *Let A and B be two $m \times n$ matrices. Further, let $\alpha = dim(col(A) \cap col(B))$ and $\beta = dim(col(A^T) \cap col(B^T))$. Then*

$$rank(A+B) = rank(A) + rank(B)$$

if and only if $\alpha = \beta = 0$.

We present the proof for the case where A and B are symmetric.
Proof. It follows from Eqs. (1.5) and 1.6 that

$$\text{rank}(A+B) \leq \text{rank}([A\ B]) \leq \text{rank}(A) + \text{rank}(B) - \alpha \leq \text{rank}(A) + \text{rank}(B).$$

Thus, if $\text{rank}(A+B) = \text{rank}(A) + \text{rank}(B)$, then $\alpha = 0$.

To prove the reverse direction, let $A = W_1 \Lambda_1 W_1^T$ and $B = W_2 \Lambda_2 W_2^T$ be the spectral decompositions of A and B, where Λ_1 and Λ_2 are the diagonal matrices consisting of the nonzero eigenvalues of A and B. Thus, W_1 and W_2 are orthonormal bases of $\text{col}(A)$ and $\text{col}(B)$. Let $\Lambda = \begin{bmatrix} \Lambda_1 & \mathbf{0} \\ \mathbf{0} & \Lambda_2 \end{bmatrix}$. Then $\text{rank}(\Lambda) = \text{rank}(A) + \text{rank}(B)$.

If $\alpha = 0$, then $W = [W_1\ W_2]$ has full column rank and hence, has a left inverse $W^\dagger = (W^T W)^{-1} W^T$; i.e., $W^\dagger W = I$. Moreover,

$$A + B = [W_1\ W_2] \begin{bmatrix} \Lambda_1 & \mathbf{0} \\ \mathbf{0} & \Lambda_2 \end{bmatrix} \begin{bmatrix} W_1^T \\ W_2^T \end{bmatrix}.$$

Thus,

$$\text{rank}(A+B) = \text{rank}(W \Lambda W^T) \geq \text{rank}(W^\dagger W \Lambda W^T (W^\dagger)^T) = \text{rank}(\Lambda)$$

and the proof is complete.

□

As an illustration of Theorem 1.12, let A and B be two real symmetric matrices such that $AB = \mathbf{0}$. Then clearly col(B) is perpendicular to col(A). Thus, by Theorem 1.12, rank($A+B$) = rank(A) + rank(B).

A real-valued function $f(x)$ is said to be *lower semicontinuous* if the set $\{x : f(x) > a\}$ is open for every $a \in \mathbb{R}$.

Lemma 1.2 *Let $\mathscr{S}^{mn}$ denote the set of $m \times n$ real matrices and let $S \subseteq \mathbb{R}^n$. Let $A(x) : S \to \mathscr{S}^{mn}$ be continuous. Then rank$(A(x))$ is lower semicontinuous.*

Proof. Let $a \in \mathbb{R}$ and let $S' = \{x \in S : \text{rank}(A(x)) > a\}$. If $S' = \emptyset$, then S' is open. Therefore, assume that $S' \neq \emptyset$ and let $x^0 \in S'$. Assume that $\text{rank}(A(x^0)) = k$, then there exists a $k \times k$ submatrix of $A(x^0)$, say $A_{\mathscr{I}\mathscr{J}}(x^0)$, such that $\det(A_{\mathscr{I}\mathscr{J}}(x^0)) \neq 0$. Hence, by the continuity of the determinant function, there exists a neighborhood U of x^0 such that $\det(A_{\mathscr{I}\mathscr{J}}(x)) \neq 0$ for all $x \in U$. Consequently, $\text{rank}(A(x)) \geq k > a$ for all $x \in U$ and thus $U \subset S'$. As a result, S' is open and the result follows. □

Hence, for a sufficiently small perturbation of A, rank(A) either stays the same or increases. That is, for a sufficiently small neighborhood U of x^0, $\text{rank}(A(x)) \geq \text{rank}(A(x^0))$ for all $x \in U$.

We end this section with a useful property of rank-2 symmetric matrices. Vectors u and v in $\mathbb{R}^n$ are *parallel* if $u = cv$ for some nonzero scalar c. Thus, u and v are parallel if $u = v = \mathbf{0}$.

Proposition 1.2 *Let a and b be two nonzero, nonparallel vectors in $\mathbb{R}^n$, $n \geq 2$, and let $C = ab^T + ba^T$. Then C has exactly one positive eigenvalue λ_1 and one negative eigenvalue λ_n, where*

$$\lambda_1 = a^T b + ||a||\,||b|| \text{ and } \lambda_n = a^T b - ||a||\,||b||.$$

Here, $||.||$ is the Euclidean norm.

Proof. Assume that $n = 2$ and let the eigenvectors of C be of the form $xa + yb$, where x and y are scalars. Then $C(xa + yb) = \lambda(xa + yb)$ leads to the following system of equations:

$$\begin{bmatrix} a^T b & ||b||^2 \\ ||a||^2 & a^T b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}. \tag{1.8}$$

Hence, the eigenvalues of C are precisely the eigenvalues of $\begin{bmatrix} a^T b & ||b||^2 \\ ||a||^2 & a^T b \end{bmatrix}$, which are $\lambda_1 = a^T b + ||a||\,||b||$ and $\lambda_r = a^T b - ||a||\,||b||$.

Now assume that $n \geq 3$ and let $u^1, \ldots, u^{n-2}$ be an orthonormal basis of the null space of $\begin{bmatrix} a^T \\ b^T \end{bmatrix}$. Then obviously, $u^1, \ldots, u^{n-2}$ are orthonormal eigenvectors of C corresponding to eigenvalue 0. Thus, we have two remaining eigenvectors of C of the form $xa + yb$, where x and y satisfy Eq. (1.8). Therefore, the remaining two eigenvalues of C are λ_1 and λ_n as given above. The fact that $\lambda_1 > 0$ and $\lambda_n < 0$ follows from Cauchy–Schwarz inequality since a and b are nonzero and nonparallel. □

1.2.4 Hadamard and Kronecker Products

Let A and B be two $m \times n$ matrices. The *Hadamard product* of A and B, denoted by $A \circ B$, is the $m \times n$ matrix C such that $c_{ij} = a_{ij} b_{ij}$ for all $i = 1, \ldots, m$ and $j = 1, \ldots, n$. An $n \times n$ symmetric matrix A is said to be *positive definite (positive semidefinite)* if and only if all of its eigenvalues are positive (nonnegative). Chapter 2 is devoted to a detailed study of these matrices.

Theorem 1.13 (Schur Product Theorem) *Let A and B be two $n \times n$ symmetric positive semidefinite matrices. Then $A \circ B$ is symmetric positive semidefinite.*

It should be pointed out that Schur Product Theorem follows from Theorem 1.15 below. Let A and B be two $m \times n$ and $p \times q$ matrices, respectively. The *Kronecker product* of A and B, denoted by $A \otimes B$, is the $mp \times nq$ matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}.$$

The following basic lemma follows immediately from the definition.

Lemma 1.3 *Let A, B, C, and D be matrices of appropriate sizes. Then*

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

Let A and B be two $n \times n$ and $m \times m$ matrices and let (λ, x) and (μ, y) be two eigenpairs of A and B, respectively. Then it immediately follows from Lemma 1.3 that $(\lambda\mu, x \otimes y)$ is an eigenpair of $A \otimes B$. Moreover, every eigenvalue of $A \otimes B$ is of the form $\lambda_i \mu_j$ where λ_i and μ_j are eigenvalues of A and B. Hence, we have the following two theorems.

Theorem 1.14 *Let A and B be two $n \times n$ and $m \times m$ matrices. Then*

1. $\det(A \otimes B) = (\det(A))^m \, (\det(B))^n$,
2. $\mathrm{trace}(A \otimes B) = \mathrm{trace}(A) \, \mathrm{trace}(B)$.

Theorem 1.15 *Let A and B be two symmetric positive semidefinite matrices. Then $A \otimes B$ is positive semidefinite.*

Schur Product Theorem follows from Theorem 1.15 since $A \circ B$ is a principal submatrix of $A \otimes B$. See Chap. 2 for more details.

For more details concerning the topics discussed in this section, see, e.g., [112, 113, 41].

1.3 Graph Theory

In this monograph, we are interested in connected simple (no loops and no multiple edges) graphs. For a simple graph $G = (V,E)$, we denote by $V(G)$ its node set and by $E(G)$ its edge set. We assume that $V(G) = \{1,\ldots,n\}$ and that the number of edges of G is m. The complement graph of G is $\bar{G} = (V(G), \bar{E}(\bar{G}))$, where $\{i, j\} \in \bar{E}(\bar{G})$ iff $i \neq j$ and $\{i, j\} \notin E(G)$. The cardinality of $\bar{E}(\bar{G})$ is denoted by $\bar{m}$ and hence $\bar{m} = n(n-1)/2 - m$. The edges of $\bar{G}$ are referred to as the *missing edges* of G.

The *adjacency matrix* of G is the $n \times n$ symmetric $(0-1)$ matrix $A = (a_{ij})$ such that $a_{ij} = 1$ iff $\{i, j\} \in E(G)$. The *degree* of node i, denoted by $\deg(i)$, is the number of edges incident with i. The vector consisting of all the degrees is denoted by deg. As a result, $\deg = Ae$, where e is the vector of all 1's in $\mathbb{R}^n$. Nodes of degree one are called *leaves*. It is easy to see that the sum of the degrees in a graph is equal to twice the number of its edges, i.e., $e^T \deg = 2m$.

Let us orient each edge $\{i, j\}$ arbitrarily as (i, j) and refer to i and j as the *tail* and the *head* of (i, j), respectively. The *node-edge incidence matrix* of G is the $n \times m$ matrix $M = (m_{ij})$ such that

$$m_{ij} = \begin{cases} 1 & \text{if node } i \text{ is the tail of edge } j, \\ -1 & \text{if node } i \text{ is the head of edge } j, \\ 0 & \text{otherwise.} \end{cases}$$

Obviously, $M^T e = \mathbf{0}$. Moreover, it is easy to prove that $\text{rank}(M) = n-1$ if and only if G is connected. The matrix

$$L = \text{Diag}(\deg) - A = \text{Diag}(Ae) - A$$

is called the *Laplacian* of G. It is a well-known result in algebraic graph theory [42] that $L = MM^T$. Consequently, $Le = \mathbf{0}$ and L is positive semidefinite. Moreover, $\text{rank}(L) = n-1$ iff G is connected.

Graph G is *complete* if its adjacency matrix is $A = E - I$, where E is the $n \times n$ matrix of all 1's and I is the identity matrix of order n. That is, G is complete if every two of its nodes are adjacent. The complete graph on n nodes is denoted by K_n. A *clique* of G is a complete subgraph of G. Graph G is said to be *k-vertex connected* if either $G = K_{k+1}$, or $|V(G)| \geq k+2$ and the deletion of any $k-1$ nodes leaves G connected. Connected graphs with no cycles are called *trees* . It is not hard to see that if T is a tree on n nodes, then T has $n-1$ edges. Thus for a tree T, $e^T \deg = 2(n-1)$.

A graph is said to be *series-parallel* [47] if it can be obtained from an edge by a sequence of series and parallel extensions. A *series extension* is the subdivision of an edge, while a *parallel extension* is the addition of a new edge joining two adjacent nodes.

Graph G is said to be *chordal* [44, 89] if every cycle of G of length ≥ 4 has a chord, that is, an edge connecting two nonconsecutive nodes on the cycle. Chordal graphs are also known as *triangulated*, *monotone transitive*, *rigid circuit*, or *perfect elimination graphs*.

Among the many different characterizations of chordal graphs, the following two are the most useful for our purposes. An ordering $\pi(1), \ldots, \pi(n)$ of the vertices of a graph G is called a *perfect elimination ordering* if for each $i = 1, \ldots, n-1$, the set of vertices $\pi(j)$, $j > i$ that are adjacent to $\pi(i)$, induce a clique in G.

Theorem 1.16 (Fulkerson and Gross [82]) *Graph G is chordal if and only if it has a perfect elimination ordering.*

It should be pointed out that such a perfect elimination ordering can be obtained in polynomial time [161].

Let G_1 and G_2 be two graphs and let $K_1 \subset V(G_1)$ and $K_2 \subset V(G_2)$ be two cliques of the same cardinality. We say that G is a *clique sum* of G_1 and G_2 if it is obtained from G_1 and G_2 by identifying K_1 and K_2, and then deleting duplicate edges in the clique.

Theorem 1.17 (Dirac [72]) *Graph G is a chordal graph if and only if it is a clique sum of complete graphs.*

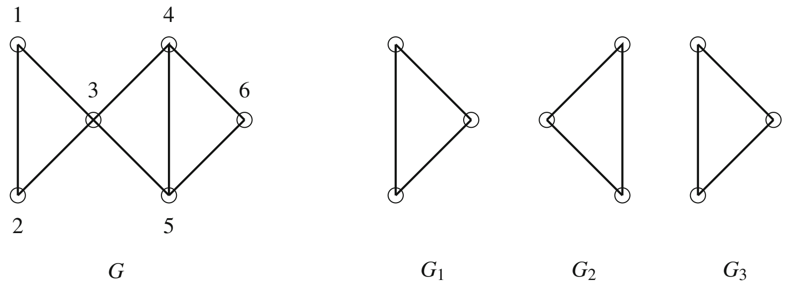

Fig. 1.1 The chordal graph of Example 1.1

Example 1.1 *Consider the chordal graph G depicted in Fig. 1.1. Clearly, the ordering $1, 2, \ldots, 6$ is a perfect elimination ordering of G. Also, it is clear that G is a clique sum of the three complete graphs G_1, G_2, and G_3.*

A graph is *planar* if it can be drawn in the plane with no two of its edges crossing. Every planar graph admits a planar drawing in which all edges are straight line segments (Fáry's Theorem [78]). A drawing of a connected planar graph divides the plane into regions or faces. The unbounded face is called the *outer face* and all other bounded faces are called *inner faces*.

1.4 Convexity Theory

Convex sets play a prominent role in this monograph. For excellent references on the topics discussed in this section, see, e.g., [160, 109, 166]. Let $\mathscr{V}$ be a finite-