

Martin Grötschel
László Lovász
Alexander Schrijver

Geometric Algorithms and Combinatorial Optimization

几何算法和组合最优化 [英]

Springer-Verlag
World Publishing Corp

Martin Grötschel
László Lovász
Alexander Schrijver

Geometric Algorithms and Combinatorial Optimization



Springer-Verlag
World Publishing Corp

Martin Grötschel
Institute of Mathematics
University of Augsburg
Memminger Straße 6
D-8900 Augsburg
Fed. Rep. of Germany

László Lovász
Department of Computer Science
Eötvös Loránd University
Budapest
Múzeum krt. 6-8
Hungary H-1088

Alexander Schrijver
Department of Econometrics
Tilburg University
P.O. Box 90153
NL-5000 LE Tilburg
The Netherlands

1980 Mathematics Subject Classification (1985 Revision):
primary 05-02, 11Hxx, 52-02, 90Cxx; secondary 05Cxx, 11H06,
11H55, 11J13, 52A43, 68Q25, 90C05, 90C10, 90C25, 90C27

ISBN 3-540-13624-X Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-13624-X Springer-Verlag New York Berlin Heidelberg

With 23 Figures

Library of Congress Cataloging-in-Publication Data
Grötschel, Martin.

Geometric algorithms and combinatorial optimization.

(Algorithms and combinatorics; 2)

Bibliography: p. Includes indexes.

1. Combinatorial geometry. 2. Geometry of numbers. 3. Mathematical optimization.

4. Programming (Mathematics) I. Lovász, László, 1948-. II. Schrijver, A. III. Title.

IV. Series. QA167.G76 1988 511'.6 87-36923

ISBN 0-387-13624-X (U.S.)

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in other ways, and storage in data banks. Duplication of this publication or parts thereof is only permitted under the provisions of the German Copyright Law of September 9, 1965, in its version of June 24, 1985, and a copyright fee must always be paid. Violations fall under the prosecution act of the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1988

Reprinted by World Publishing Corporation, Beijing, 1990
for distribution and sale in The People's Republic of China only
ISBN 7-5062-0768-0

Preface

Historically, there is a close connection between geometry and optimization. This is illustrated by methods like the gradient method and the simplex method, which are associated with clear geometric pictures. In combinatorial optimization, however, many of the strongest and most frequently used algorithms are based on the discrete structure of the problems: the greedy algorithm, shortest path and alternating path methods, branch-and-bound, etc. In the last several years geometric methods, in particular polyhedral combinatorics, have played a more and more profound role in combinatorial optimization as well.

Our book discusses two recent geometric algorithms that have turned out to have particularly interesting consequences in combinatorial optimization, at least from a theoretical point of view. These algorithms are able to utilize the rich body of results in polyhedral combinatorics.

The first of these algorithms is the **ellipsoid method**, developed for nonlinear programming by N. Z. Shor, D. B. Yudin, and A. S. Nemirovskii. It was a great surprise when L. G. Khachiyan showed that this method can be adapted to solve linear programs in polynomial time, thus solving an important open theoretical problem. While the ellipsoid method has not proved to be competitive with the simplex method in practice, it does have some features which make it particularly suited for the purposes of combinatorial optimization.

The second algorithm we discuss finds its roots in the classical "geometry of numbers", developed by Minkowski. This method has had traditionally deep applications in number theory, in particular in diophantine approximation. Methods from the geometry of numbers were introduced in integer programming by H. W. Lenstra. An important element of his technique, called **basis reduction**, goes in fact back to Hermite. An efficient version of basis reduction yields a polynomial time algorithm useful not only in combinatorial optimization, but also in fields like number theory, algebra, and cryptography.

A combination of these two methods results in a powerful tool for combinatorial optimization. It yields a theoretical framework in which the polynomial time solvability of a large number of combinatorial optimization problems can be shown quite easily. It establishes the algorithmic equivalence of problems which are "dual" in various senses.

Being this general, this method cannot be expected to give running times comparable with special-purpose algorithms. Our policy in this book is, therefore, not to attempt to obtain the best possible running times; rather, it is to derive just the polynomial time solvability of the problems as quickly and painlessly as

possible. Thus, our results are best conceived as “almost pure” existence results for polynomial time algorithms for certain problems and classes of problems.

Nevertheless, we could not get around quite a number of tedious technical details. We did try to outline the essential ideas in certain sections, which should give an outline of the underlying geometric and combinatorial ideas. Those sections which contain the technical details are marked by an asterisk in the list of contents. We therefore recommend, for a first reading, to skip these sections.

The central result proved and applied in this book is, roughly, the following. If K is a convex set, and if we can decide in polynomial time whether a given vector belongs to K , then we can optimize any linear objective function over K in polynomial time. This assertion is, however, not valid without a number of conditions and restrictions, and even to state these we have to go through many technical details. The most important of these is that the optimization can be carried out in an approximate sense only (as small compensation, we only need to test for membership in K in an approximate sense).

Due to the rather wide spread of topics and methods treated in this book, it seems worth while to outline its structure here.

Chapters 0 and 1 contain mathematical preliminaries. Of these, Chapter 1 discusses some non-standard material on the complexity of problems, efficiency of algorithms and the notion of oracles.

The main result, and its many versions and ramifications, are obtained by the ellipsoid method. Chapter 2 develops the framework necessary for the formulation of algorithmic problems on convex sets and the design of algorithms to solve these. A list of the main problems introduced in Chapter 2 can be found on the inner side of the back cover. Chapter 3 contains the description of (two versions of) the ellipsoid method. The statement of what exactly is achieved by this method is rather complicated, and the applications and specializations collected in Chapter 4 are, perhaps, more interesting. These range from the main result mentioned above to results about computing the diameter, width, volume, and other geometric parameters of convex sets. All these algorithms provide, however, only approximations.

Polyhedra encountered in combinatorial optimization have, typically, vertices with small integral entries and facets with small integral coefficients. For such polyhedra, the optimization problem (and many other algorithmic problems) can be solved in the exact sense, by rounding an approximate solution appropriately. While for many applications a standard rounding to some number of digits is sufficient, to obtain results in full generality we will have to use the sophisticated rounding technique of diophantine approximation. The basis reduction algorithm for lattices, which is the main ingredient of this technique, is treated in Chapter 5, along with several applications. Chapter 6 contains the main applications of diophantine approximation techniques. Besides strong versions of the main result, somewhat different combinations of the ellipsoid method with basis reduction give the strongly polynomial time solvability of several combinatorial optimization problems, and the polynomial time solvability of integer linear programming in fixed dimension, remarkable results of É. Tardos and H. W. Lenstra, respectively.

Chapters 7 to 10 contain the applications of the results obtained in the previous chapters to combinatorial optimization. Chapter 7 is an easy-to-read introduction to these applications. In Chapter 8 we give an in-depth survey of combinatorial optimization problems solvable in polynomial time with the methods of Chapter 6. Chapters 9 and 10 treat two specific areas where the ellipsoid method has resolved important algorithmic questions that so far have resisted direct combinatorial approaches: perfect graphs and submodular functions.

We are grateful to several colleagues for many discussions on the topic and text of this book, in particular to Bob Bixby, András Frank, Michael Jünger, Gerhard Reinelt, Éva Tardos, Klaus Truemper, Yoshiko Wakabayashi, and Zaw Win. We mention at this point that the technique of applying the ellipsoid method to combinatorial optimization problems was also discovered by R. M. Karp, C. H. Papadimitriou, M. W. Padberg, and M. R. Rao.

We have worked on this book over a long period at various institutions. We acknowledge, in particular, the support of the joint research project of the German Research Association (DFG) and the Hungarian Academy of Sciences (MTA), the Universities of Amsterdam, Augsburg, Bonn, Szeged, and Tilburg, Cornell University (Ithaca), Eötvös Loránd University (Budapest), and the Mathematical Centre (Amsterdam).

Our special thanks are due to Frau Theodora Konnerth for the efficient and careful typing and patient retyping of the text in \TeX .

March 1987

Martin Grötschel
László Lovász
Alexander Schrijver

New series!

Algorithms and Combinatorics

Editors: R. L. Graham, B. Korte, L. Lovász

Combinatorial mathematics has substantially influenced recent trends and developments in the theory of algorithms and its applications. Conversely, research on algorithms and their complexity has established new perspectives in discrete mathematics. This new series is devoted to the mathematics of these rapidly growing fields with special emphasis on their mutual interaction.

The series will cover areas in pure and applied mathematics as well as computer science, including: combinatorial and discrete optimization, polyhedral combinatorics, graph theory and its algorithmic aspects, network flows, matroids and their applications, algorithms in number theory, group theory etc., coding theory, algorithmic complexity of combinatorial problems, and combinatorial methods in computer science and related areas.

The main body of this series will be monographs ranging in level from first-year graduate up to advanced state-of-the-art research. The books will be conventionally type-set and bound in hard cover. In new and rapidly growing areas, collections of carefully edited monographic articles are also appropriate for this series. Occasionally there will also be "lecture-notes-type" volumes within the series, published as *Study and Research Texts* in soft cover and camera-ready form: this will be mainly an outlet for seminar notes, drafts of textbooks with essential novelty in their presentation, and preliminary drafts of monographs. Prospective readers of the series **Algorithms and Combinatorics** include scientists and graduate students working in discrete mathematics, operations research and computer science.

Volume 1: K. H. Borgwardt

The Simplex Method

A Probabilistic Analysis

Volume 2: M. Grötschel, L. Lovász, A. Schrijver

Geometric Algorithms and Combinatorial Optimization

Volume 3: K. Murota

Systems Analysis by Graphs and Matroids

Structural Solvability and Controllability

In preparation:

Volume 4: B. Korte, L. Lovász, R. Schrader

Greedoids

ISBN 3-540-18190-3

Volume 5: J. Nešetřil, V. Rödl (Eds.)

Mathematics of Ramsey Theory

ISBN 3-540-18191-1

Volume 6: A. Recski

Matroid Theory and its Applications

ISBN 3-540-15285-7

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo

Springer



K. H. Borgwardt

The Simplex Method

A Probabilistic Analysis

1987. 42 figures in 115 separate illustrations. XI, 268 pages.
(Algorithms and Combinatorics, Volume 1).
ISBN 3-540-17096-0

Contents: Introduction. - The Shadow-Vertex Algorithm. - The Average Number of Pivot Steps. - The Polynomiality of the Expected Number of Steps. - Asymptotic Results. - Problems with Nonnegativity Constraints. - Appendix. - References. - Subject Index.

M. Grötschel, L. Lovász, A. Schrijver

Geometric Algorithms and Combinatorial Optimization

1988. 23 figures. XII, 362 pages. (Algorithms and Combinatorics, Volume 2). ISBN 3-540-13624-X

Contents: Mathematical Preliminaries. - Complexity, Oracles, and Numerical Computation. - Algorithmic Aspects of Convex Sets: Formulation of the Problems. - The Ellipsoid Method. - Algorithms for Convex Bodies. - Diophantine Approximation and Basis Reduction. - Rational Polyhedra. - Combinatorial Optimization: Some Basic Examples. - Combinatorial Optimization: A Tour d'Horizon. - Stable Sets in Graphs. - Submodular Functions. - References. - Notation Index. - Author Index. - Subject Index.

K. Murota

Systems Analysis by Graphs and Matroids

Structural Solvability and Controllability

1987. 54 figures. IX, 281 pages. (Algorithms and Combinatorics, Volume 3). ISBN 3-540-17659-4

Contents: Introduction. - Preliminaries. - Graph-Theoretic Approach to the Solvability of a System of Equations. - Graph-Theoretic Approach to the Controllability of a Dynamical System. - Physical Observations for Faithful Formulations. - Matroid-Theoretic Approach to the Solvability of a System of Equations. - Matroid-Theoretic Approach to the Controllability of a Dynamical System. - Conclusion. - References. - Index.

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo

Springer



Table of Contents

Chapter 0. Mathematical Preliminaries	1
0.1 Linear Algebra and Linear Programming	1
Basic Notation	1
Hulls, Independence, Dimension	3
Eigenvalues, Positive Definite Matrices	4
Vector Norms, Balls	5
Matrix Norms	7
Some Inequalities	8
Polyhedra, Inequality Systems	9
Linear (Diophantine) Equations and Inequalities	11
Linear Programming and Duality	14
0.2 Graph Theory	16
Graphs	17
Digraphs	18
Walks, Paths, Circuits, Trees	19
Chapter 1. Complexity, Oracles, and Numerical Computation	21
1.1 Complexity Theory: \mathcal{P} and \mathcal{NP}	21
Problems	21
Algorithms and Turing Machines	22
Encoding	23
Time and Space Complexity	23
Decision Problems: The Classes \mathcal{P} and \mathcal{NP}	24
1.2 Oracles	26
The Running Time of Oracle Algorithms	26
Transformation and Reduction	27
\mathcal{NP} -Completeness and Related Notions	28
1.3 Approximation and Computation of Numbers	29
Encoding Length of Numbers	29
Polynomial and Strongly Polynomial Computations	32
Polynomial Time Approximation of Real Numbers	33

The sections and chapters marked with * are technical. We recommend that the reader skip these on the first reading.

1.4	Pivoting and Related Procedures	36
	Gaussian Elimination	36
	Gram-Schmidt Orthogonalization	40
	The Simplex Method	41
	Computation of the Hermite Normal Form	43
Chapter 2. Algorithmic Aspects of Convex Sets:		
	Formulation of the Problems	46
2.1	Basic Algorithmic Problems for Convex Sets	47
* 2.2	Nondeterministic Decision Problems for Convex Sets	56
Chapter 3. The Ellipsoid Method		
3.1	Geometric Background and an Informal Description	66
	Properties of Ellipsoids	66
	Description of the Basic Ellipsoid Method	73
	Proofs of Some Lemmas	76
	Implementation Problems and Polynomiality	80
	Some Examples	83
* 3.2	The Central-Cut Ellipsoid Method	86
* 3.3	The Shallow-Cut Ellipsoid Method	94
Chapter 4. Algorithms for Convex Bodies		
4.1	Summary of Results	102
* 4.2	Optimization from Separation	105
* 4.3	Optimization from Membership	107
* 4.4	Equivalence of the Basic Problems	114
* 4.5	Some Negative Results	118
* 4.6	Further Algorithmic Problems for Convex Bodies	122
* 4.7	Operations on Convex Bodies	128
	The Sum	128
	The Convex Hull of the Union	129
	The Intersection	129
	Polars, Blockers, Antiblockers	131
Chapter 5. Diophantine Approximation and Basis Reduction		
5.1	Continued Fractions	134
5.2	Simultaneous Diophantine Approximation: Formulation of the Problems	138
5.3	Basis Reduction in Lattices	139
* 5.4	More on Lattice Algorithms	150

Chapter 6. Rational Polyhedra	157
6.1 Optimization over Polyhedra: A Preview	157
* 6.2 Complexity of Rational Polyhedra	162
* 6.3 Weak and Strong Problems	170
* 6.4 Equivalence of Strong Optimization and Separation	174
* 6.5 Further Problems for Polyhedra	181
* 6.6 Strongly Polynomial Algorithms	188
* 6.7 Integer Programming in Bounded Dimension	192
Chapter 7. Combinatorial Optimization: Some Basic Examples	197
7.1 Flows and Cuts	197
7.2 Arborescences	201
7.3 Matching	203
7.4 Edge Coloring	208
7.5 Matroids	210
7.6 Subset Sums	218
7.7 Concluding Remarks	221
* Chapter 8. Combinatorial Optimization: A Tour d'Horizon	225
* 8.1 Blocking Hypergraphs and Polyhedra	225
* 8.2 Problems on Bipartite Graphs	229
* 8.3 Flows, Paths, Chains, and Cuts	233
* 8.4 Trees, Branchings, and Rooted and Directed Cuts	242
Arborescences and Rooted Cuts	242
Trees and Cuts in Undirected Graphs	247
Dicuts and Dijoins	251
* 8.5 Matchings, Odd Cuts, and Generalizations	254
Matching	255
<i>b</i>-Matching	257
<i>T</i>-Joins and <i>T</i>-Cuts	259
Chinese Postmen and Traveling Salesmen	262
* 8.6 Multicommodity Flows	266
* Chapter 9. Stable Sets in Graphs:	272
* 9.1 Odd Circuit Constraints and <i>t</i> -Perfect Graphs	273
* 9.2 Clique Constraints and Perfect Graphs	276
Antiblockers of Hypergraphs	284
* 9.3 Orthonormal Representations	285
* 9.4 Coloring Perfect Graphs	296
* 9.5 More Algorithmic Results on Stable Sets	299

* Chapter 10. Submodular Functions	304
* 10.1 Submodular Functions and Polymatroids	304
* 10.2 Algorithms for Polymatroids and Submodular Functions	308
Packing Bases of a Matroid	311
* 10.3 Submodular Functions on Lattice, Intersecting, and Crossing Families	313
* 10.4 Odd Submodular Function Minimization and Extensions	325
References	331
Notation Index	347
Author Index	351
Subject Index	355

Five Basic Problems (see inner side of the back cover)

Chapter 0

Mathematical Preliminaries

This chapter summarizes mathematical background material from linear algebra, linear programming, and graph theory used in this book. We expect the reader to be familiar with the concepts treated here. We do not recommend to go thoroughly through all the definitions and results listed in the sequel – they are mainly meant for reference.

0.1 Linear Algebra and Linear Programming

In this section we survey notions and well-known facts from linear algebra, linear programming, polyhedral theory, and related fields that will be employed frequently in subsequent chapters. We have also included a number of useful inequalities and estimates. The material covered here is standard and can be found in several textbooks. As references for linear algebra we mention FADDEEV and FADDEVA (1963), GANTMACHER (1959), LANCASTER and TISMENETSKY (1985), MARCUS and MINC (1964), STRANG (1980). For information on linear programming and polyhedral theory see for instance CHVÁTAL (1983), DANTZIG (1963), GASS (1984), GRÜNBAUM (1967), ROCKAFELLAR (1970), SCHRIJVER (1986), STOER and WITZGALL (1970).

Basic Notation

By \mathbb{R} (\mathbb{Q} , \mathbb{Z} , \mathbb{N} , \mathbb{C}) we denote the set of real (rational, integral, natural, complex) numbers. The set \mathbb{N} of natural numbers does not contain zero. \mathbb{R}_+ (\mathbb{Q}_+ , \mathbb{Z}_+) denotes the nonnegative real (rational, integral) numbers. For $n \in \mathbb{N}$, the symbol \mathbb{R}^n (\mathbb{Q}^n , \mathbb{Z}^n , \mathbb{N}^n , \mathbb{C}^n) denotes the set of vectors with n components (or n -tuples or n -vectors) with entries in \mathbb{R} (\mathbb{Q} , \mathbb{Z} , \mathbb{N} , \mathbb{C}). If E and R are sets, then R^E is the set of mappings of E to R . If E is finite, it is very convenient to consider the elements of R^E as $|E|$ -vectors where each component of a vector $x \in R^E$ is indexed by an element of E , i. e., $x = (x_e)_{e \in E}$. For $F \subseteq E$, the vector $\chi^F \in \mathbb{R}^E$ defined by $\chi_e^F = 1$ if $e \in F$ and $\chi_e^F = 0$ if $e \in E \setminus F$ is called the **incidence vector** of F .

Addition of vectors and multiplication of vectors with scalars are as usual. With these operations, \mathbb{R}^n and \mathbb{Q}^n are vector spaces over the fields \mathbb{R} and \mathbb{Q} , respectively, while \mathbb{Z}^n is a module over the ring \mathbb{Z} . A vector is always considered as a **column vector**, unless otherwise stated. The superscript “ T ”

denotes **transposition**. So, for $x \in \mathbb{R}^n$, x^T is a row vector, unless otherwise stated. \mathbb{R}^n is endowed with a (Euclidean) **inner product** defined as follows:

$$x^T y := \sum_{i=1}^n x_i y_i \quad \text{for } x, y \in \mathbb{R}^n.$$

For a real number α , the symbol $\lfloor \alpha \rfloor$ denotes the largest integer not larger than α (the **floor** or **lower integer part** of α), $\lceil \alpha \rceil$ denotes the smallest integer not smaller than α (the **ceiling** or **upper integer part** of α) and $\lceil \alpha \rceil := \lceil \alpha - \frac{1}{2} \rceil$ denotes the integer nearest to α . If $a = (a_1, \dots, a_n)^T$ and $b = (b_1, \dots, b_n)^T$ are vectors, we write $a \leq b$ if $a_i \leq b_i$ for $i = 1, \dots, n$.

For two sets M and N , the expression $M \subseteq N$ means that M is a subset of N , while $M \subset N$ denotes strict containment, i. e., $M \subseteq N$ and $M \neq N$. We write $M \setminus N$ for the set-theoretical difference $\{x \in M \mid x \notin N\}$, $M \Delta N$ for the **symmetric difference** $(M \setminus N) \cup (N \setminus M)$, and 2^M for the set of all subsets of M , the so-called **power set** of M . For $M, N \subseteq \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, we use the following standard terminology for set operations: $M + N := \{x + y \mid x \in M, y \in N\}$, $\alpha M := \{\alpha x \mid x \in M\}$, $-M := \{-x \mid x \in M\}$, $M - N := M + (-N)$.

For any set R , $R^{m \times n}$ denotes the set of $m \times n$ -matrices with entries in R . For a matrix $A \in R^{m \times n}$, we usually assume that the row index set of A is $\{1, \dots, m\}$ and that the column index set is $\{1, \dots, n\}$. Unless specified otherwise, the elements or entries of $A \in R^{m \times n}$ are denoted by a_{ij} , $1 \leq i \leq m$, $1 \leq j \leq n$; we write $A = (a_{ij})$. Vectors with n components are also considered as $n \times 1$ -matrices.

If I is a subset of the row index set M of a matrix A and J a subset of the column index set N of A , then A_{IJ} denotes the submatrix of A induced by those rows and columns of A whose indices belong to I and J , respectively. Instead of A_{MJ} (A_{IN} resp.) we frequently write $A_{\cdot J}$ ($A_{I \cdot}$, resp.). A submatrix A of the form A_{II} is called a **principal submatrix** of A . If $K = \{1, \dots, k\}$ then A_{KK} is called the k -th **leading principal submatrix** of A . A_i is the i -th row of A (so it is a row vector), and A_j is the j -th column of A .

Whenever we do not explicitly state whether a number, vector, or matrix is integral, rational, or complex it is implicitly assumed to be real. Moreover, we often do not specify the dimensions of vectors and matrices explicitly. When operating with them, we always assume that their dimensions are compatible.

The **identity matrix** is denoted by I or, if we want to stress its dimension, by I_n . The symbol 0 stands for any appropriately sized matrix which has all entries equal to zero, and similarly for any zero vector. The symbol $\mathbf{1}$ denotes a vector which has all components equal to one. The j -th **unit vector** in \mathbb{R}^n , whose j -th component is one while all other components are zero, is denoted by e_j . If $x = (x_1, \dots, x_n)^T$ is a vector then the $n \times n$ -matrix with the entries x_1, \dots, x_n on the main diagonal and zeros outside the main diagonal is denoted by $\text{diag}(x)$. If $A \in R^{m \times p}$ and $B \in R^{m \times q}$ then (A, B) (or just $(A \ B)$ if this does not lead to confusion) denotes the matrix in $R^{m \times (p+q)}$ whose first p columns are the columns of A and whose other q columns are those of B .

The **determinant** of an $n \times n$ -matrix A is denoted by $\det(A)$. The **trace** of an $n \times n$ -matrix A , denoted by $\text{tr}(A)$, is the sum of the elements of the main diagonal of the matrix A .

When using functions like \det , tr , or diag we often omit the brackets if there is no danger of confusion, i. e., we frequently write $\det A$ instead of $\det(A)$ etc.

The inverse matrix of an $n \times n$ -matrix A is denoted by A^{-1} . If a matrix has an inverse matrix then it is called **nonsingular**, and otherwise **singular**. An $n \times n$ -matrix A is nonsingular if and only if $\det A \neq 0$.

Hulls, Independence, Dimension

A vector $x \in \mathbb{R}^n$ is called a **linear combination** of the vectors $x_1, x_2, \dots, x_k \in \mathbb{R}^n$ if, for some $\lambda \in \mathbb{R}^k$,

$$x = \sum_{i=1}^k \lambda_i x_i.$$

If, in addition,

$$\left. \begin{array}{l} \lambda \geq 0 \\ \lambda^T \mathbf{1} = 1 \end{array} \right\} \text{ we call } x \text{ a } \left\{ \begin{array}{l} \text{conic} \\ \text{affine} \\ \text{convex} \end{array} \right\} \text{ combination}$$

of the vectors x_1, x_2, \dots, x_k . These combinations are called **proper** if neither $\lambda = 0$ nor $\lambda = e_j$ for some $j \in \{1, 2, \dots, k\}$. For a nonempty subset $S \subseteq \mathbb{R}^n$, we denote by

$$\left\{ \begin{array}{l} \text{lin}(S) \\ \text{cone}(S) \\ \text{aff}(S) \\ \text{conv}(S) \end{array} \right\} \text{ the } \left\{ \begin{array}{l} \text{linear} \\ \text{conic} \\ \text{affine} \\ \text{convex} \end{array} \right\} \text{ hull of the elements of } S,$$

that is, the set of all vectors that are linear (conic, affine, convex) combinations of finitely many vectors of S . For the empty set, we define $\text{lin}(\emptyset) := \text{cone}(\emptyset) := \{0\}$ and $\text{aff}(\emptyset) := \text{conv}(\emptyset) := \emptyset$.

A subset $S \subseteq \mathbb{R}^n$ is called

$$\left\{ \begin{array}{l} \text{a linear subspace} \\ \text{a cone} \\ \text{an affine subspace} \\ \text{a convex set} \end{array} \right\} \text{ if } \left\{ \begin{array}{l} S = \text{lin}(S) \\ S = \text{cone}(S) \\ S = \text{aff}(S) \\ S = \text{conv}(S) \end{array} \right\}.$$

A subset $S \subseteq \mathbb{R}^n$ is called **linearly (affinely) independent** if none of its members is a proper linear (affine) combination of elements of S ; otherwise S is called **linearly (affinely) dependent**. It is well-known that a linearly (affinely) independent subset of \mathbb{R}^n contains at most n elements ($n+1$ elements). For any set $S \subseteq \mathbb{R}^n$, the **rank** of S (**affine rank** of S) denoted by $\text{rank}(S)$ ($\text{arank}(S)$), is the cardinality of the largest linearly (affinely) independent subset of S . For any subset $S \subseteq \mathbb{R}^n$, the **dimension** of S , denoted by $\text{dim}(S)$, is the cardinality of a largest affinely independent subset of S minus one, i. e., $\text{dim}(S) = \text{arank}(S) - 1$. A set $S \subseteq \mathbb{R}^n$ with $\text{dim}(S) = n$ is called **full-dimensional**.

The **rank** of a matrix A (notation: $\text{rank}(A)$), is the rank of the set of its column vectors. This is known to be equal to the rank of the set of its row vectors. An $m \times n$ -matrix A is said to have **full row rank** (**full column rank**) if $\text{rank } A = m$ ($\text{rank } A = n$).

Eigenvalues, Positive Definite Matrices

If A is an $n \times n$ -matrix, then every complex number λ with the property that there is a nonzero vector $u \in \mathbb{C}^n$ such that $Au = \lambda u$ is called an **eigenvalue** of A . The vector u is called an **eigenvector** of A associated with λ . The function $f(\lambda) := \det(\lambda I_n - A)$ is a polynomial of degree n , called the **characteristic polynomial** of A . Thus the equation

$$\det(\lambda I_n - A) = 0$$

has n (complex) roots (multiple roots counted with their multiplicity). These roots are the (not necessarily distinct) n eigenvalues of A .

We will often consider **symmetric matrices** (i. e., $n \times n$ -matrices $A = (a_{ij})$ with $a_{ij} = a_{ji}$, $1 \leq i \leq j \leq n$). It is easy to see that all eigenvalues of real symmetric matrices are real numbers.

There are useful relations between the eigenvalues $\lambda_1, \dots, \lambda_n$ of a matrix A , its determinant and its trace, namely

$$(0.1.1) \quad \det A = \prod_{i=1}^n \lambda_i,$$

$$(0.1.2) \quad \operatorname{tr} A = \sum_{i=1}^n \lambda_i.$$

An $n \times n$ -matrix A is called **positive definite** (**positive semidefinite**) if A is symmetric and if $x^T A x > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$ ($x^T A x \geq 0$ for all $x \in \mathbb{R}^n$). If A is positive definite then A is nonsingular and its inverse is also positive definite. In fact, for a symmetric $n \times n$ -matrix A the following conditions are equivalent:

- (0.1.3)
- (i) A is positive definite.
 - (ii) A^{-1} is positive definite.
 - (iii) All eigenvalues of A are positive real numbers.
 - (iv) $A = B^T B$ for some nonsingular matrix B .
 - (v) $\det A_k > 0$ for $k = 1, \dots, n$, where A_k is the k -th leading principal submatrix of A .

It is well known that for any positive definite matrix A , there is exactly one matrix among the matrices B satisfying (0.1.3) (iv) that is itself positive definite. This matrix is called the (**square**) **root** of A and is denoted by $A^{1/2}$.

Positive semidefinite matrices can be characterized in a similar way, namely, for a symmetric $n \times n$ -matrix A the following conditions are equivalent:

- (0.1.4)
- (i) A is positive semidefinite.
 - (ii) All eigenvalues of A are nonnegative real numbers.
 - (iii) $A = B^T B$ for some matrix B .
 - (iv) $\det A_{II} \geq 0$ for all principal submatrices A_{II} of A .
 - (v) There is a positive definite principal submatrix A_{II} of A with $|I| = \operatorname{rank} A$.

Vector Norms, Balls

A function $N : \mathbb{R}^n \rightarrow \mathbb{R}$ is called a **norm** if the following three conditions are satisfied:

- (0.1.5) (i) $N(x) \geq 0$ for $x \in \mathbb{R}^n$, $N(x) = 0$ if and only if $x = 0$,
 (ii) $N(\alpha x) = |\alpha|N(x)$ for all $x \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$,
 (iii) $N(x + y) \leq N(x) + N(y)$ for all $x, y \in \mathbb{R}^n$ (**triangle inequality**).

Every norm N on \mathbb{R}^n induces a **distance** d_N defined by

$$d_N(x, y) := N(x - y) \quad \text{for } x, y \in \mathbb{R}^n.$$

For our purposes four norms will be especially important:

$$\|x\| := \sqrt{x^T x} = \left(\sum_{i=1}^n x_i^2 \right)^{1/2} \quad (\text{the } l_2\text{- or Euclidean norm}).$$

(This norm induces the **Euclidean distance** $d(x, y) = \|x - y\|$. Usually, the Euclidean norm is denoted by $\|\cdot\|_2$. But we use it so often that we write simply $\|\cdot\|$.)

$$\|x\|_1 := \sum_{i=1}^n |x_i| \quad (\text{the } l_1\text{- or 1-norm}),$$

$$\|x\|_\infty := \max_{1 \leq i \leq n} |x_i| \quad (\text{the } l_\infty\text{- or maximum norm}),$$

$$\|x\|_A := \sqrt{x^T A^{-1} x},$$

where A is a positive definite $n \times n$ -matrix. Recall that A induces an inner product $x^T A^{-1} y$ on \mathbb{R}^n . Norms of type $\|\cdot\|_A$ are sometimes called **general Euclidean** or **ellipsoidal norms**. We always consider the space \mathbb{R}^n as a Euclidean space endowed with the Euclidean norm $\|\cdot\|$, unless otherwise specified. So all notions related to distances and the like are defined via the Euclidean norm.

For all $x \in \mathbb{R}^n$, the following relations hold between the norms introduced above:

$$(0.1.6) \quad \|x\| \leq \|x\|_1 \leq \sqrt{n} \|x\|,$$

$$(0.1.7) \quad \|x\|_\infty \leq \|x\| \leq \sqrt{n} \|x\|_\infty,$$

$$(0.1.8) \quad \|x\|_\infty \leq \|x\|_1 \leq n \|x\|_\infty.$$

If A is a positive definite $n \times n$ -matrix with smallest eigenvalue λ and largest eigenvalue Λ then

$$(0.1.9) \quad \sqrt{\lambda} \|x\| \leq \|x\|_{A^{-1}} \leq \sqrt{\Lambda} \|x\| \quad \text{for all } x \in \mathbb{R}^n.$$