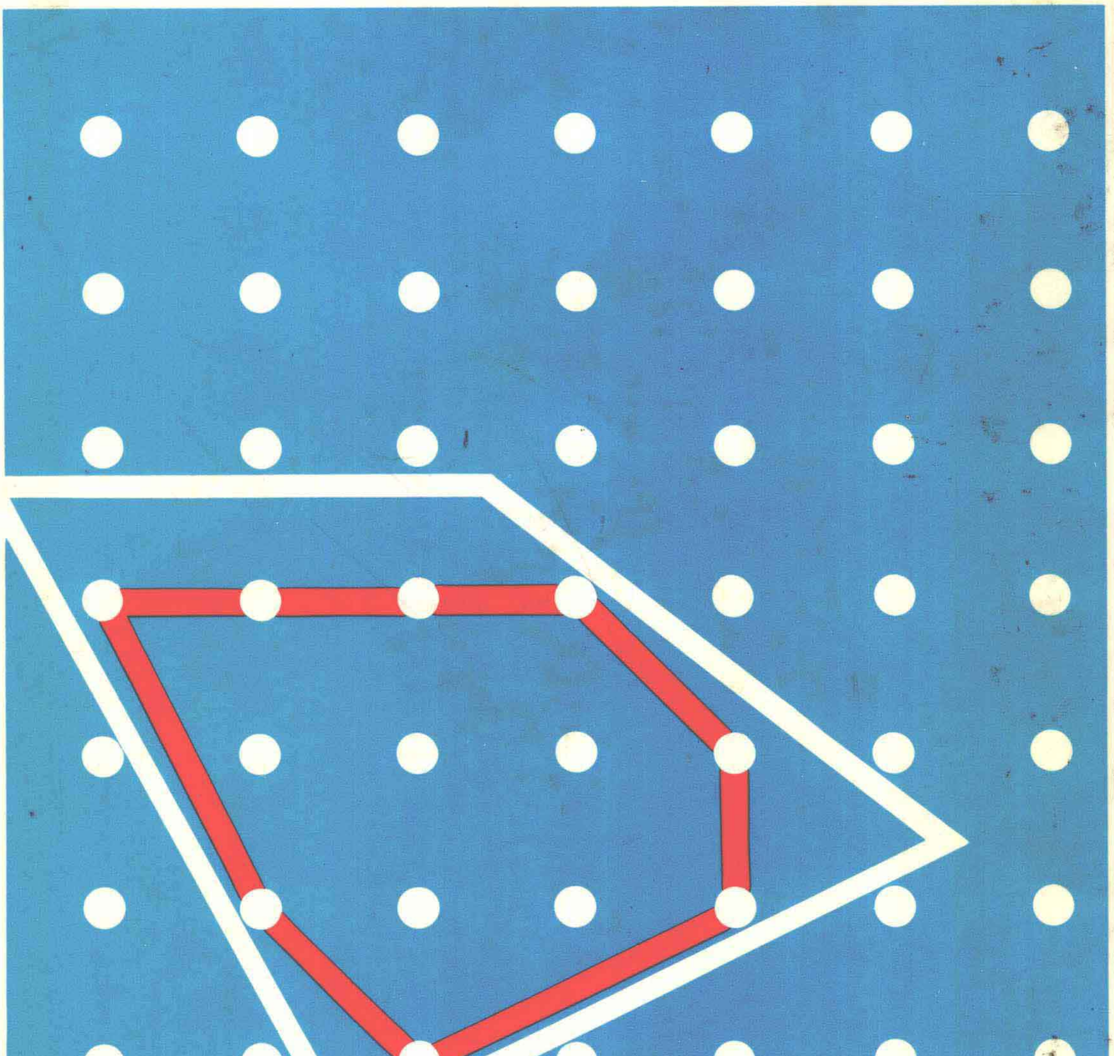
THEORY OF LINEAR AND INTEGER PROGRAMMING



Alexander Schrijver

WILEY–INTERSCIENCE SERIES IN DISCRETE MATHEMATICS AND OPTIMIZATION

THEORY OF LINEAR AND INTEGER PROGRAMMING

ALEXANDER SCHRIJVER

Department of Econometrics,
Tilburg University
and
Centrum voor Wiskunde en Informatica,
Amsterdam

A Wiley-Interscience Publication

JOHN WILEY & SONS

Chichester · New York · Brisbane · Toronto · Singapore

A. Schrijver,
Department of Econometrics,
Tilburg University,
P.O. Box 90153,
5000 LE Tilburg,
The Netherlands.

Library of Congress Cataloging-in-Publication Data:

Schrijver, Alexander.
Theory of linear and integer programming.

(Wiley–Interscience series in discrete mathematics)
'A Wiley–Interscience publication.'
Bibliography: p.
Includes index.
1. Linear programming. 2. Integer programming.
I. Title. II. Series.
T57.74.S53 1986 519.7′2 85–12314
ISBN 0 471 90854 1

British Library Cataloguing in Publication Data:

Schrijver, Alexander
Theory of linear and integer programming.
1. Linear programming 2. Integer programming
I. Title
519.7′2 T57.74

ISBN 0 471 90854 1

Printed and Bound in Great Britain by Anchor Brendan Ltd, Tiptree, Essex

Preface

There exist several excellent books on linear and integer programming. Yet, I did not feel it superfluous to write the present book. Most of the existing books focus on the, very important, algorithmic side of linear and integer programming. The emphasis of this book is on the more theoretical aspects, and it aims at complementing the more practically oriented books.

Another reason for writing this book is that during the last few years several interesting new results have been obtained, which are not yet all covered by books: Lovász's basis reduction method, Khachiyan's ellipsoid method and Karmarkar's method for linear programming, Borgwardt's analysis of the average speed of the simplex method, Tardos' and Megiddo's algorithms for linear programming, Lenstra's algorithm for integer linear programming, Seymour's decomposition theorem for totally unimodular matrices, and the theory of total dual integrality.

Although the emphasis is on theory, this book does not exclude algorithms. This is motivated not only by the practical importance of algorithms, but also by the fact that the complexity analysis of problems and algorithms has become more and more a theoretical topic as well. In particular, the study of polynomial-time solvability has led to interesting theory. Often the polynomial-time solvability of a certain problem can be shown theoretically (e.g. with the ellipsoid method); such a proof next serves as a motivation for the design of a method which is efficient in practice. Therefore we have included a survey of methods known for linear and integer programming, together with a brief analysis of their running time. Our descriptions are meant for a quick understanding of the method, and might be, in many cases, less appropriate for a direct implementation.

The book also arose as a prerequisite for the forthcoming book *Polyhedral Combinatorics*, dealing with polyhedral (i.e. linear programming) methods in combinatorial optimization. Dantzig, Edmonds, Ford, Fulkerson, and Hoffman have pioneered the application of polyhedral methods to combinatorial optimization, and now combinatorial optimization is dissolubly connected to (integer) linear programming. The book *Polyhedral Combinatorics* describes these connections, which heavily lean on results and methods discussed in the present book. For a better understanding, and to make this book self-contained, we have illustrated some of the results by combinatorial applications.

Several friends and colleagues have helped and inspired me in preparing this book. It was Cor Baayen who stimulated me to study discrete mathematics, especially combinatorial optimization, and who advanced the idea of compiling a monograph on polyhedral methods in combinatorial optimization. During leaves of absence spent in Oxford and Szeged (Hungary) I enjoyed the hospitality of Paul Seymour and Laci Lovász. Their explanations and insights have helped me considerably in understanding polyhedral combinatorics and integer linear programming. Concurrently with the present book, I was involved with Martin Grötschel and Laci Lovász in writing the book *The Ellipsoid Method and Combinatorial Optimization* (Springer-Verlag, Heidelberg). Although the plans of the two books are distinct, there is some overlap, which has led to a certain cross-fertilization. I owe much to the pleasant cooperation with my two co-authors. Also Bob Bixby, Bill Cook, Bill Cunningham, Jack Edmonds, Werner Fenchel, Bert Gerards, Alan Hoffman, Antoon Kolen, Jaap Ponstein, András Sebö, Éva Tardos, Klaus Trümper and Laurence Wolsey have helped me by pointing out to me information and ideas relevant to the book, or by reading and criticizing parts of the manuscript. The assistance of the staff of the library of the Mathematical Centre, in particular of Carin Klompen, was important in collecting many ancient articles indispensable for composing the historical surveys.

Thanks are due to all of them. I also acknowledge hospitality and/or financial support given by the following institutions and organizations: the Mathematical Centre/Centrum voor Wiskunde en Informatica, the Netherlands organization for the advancement of pure research Z.W.O., the University of Technology Eindhoven, the Bolyai Institute of the Attila József University in Szeged, the University of Amsterdam, Tilburg University, and the Institut für Ökonometrie und Operations Research of the University of Bonn.

Finally, I am indebted to all involved in the production of this book. It has been a pleasure to work with Ian McIntosh and his colleagues of John Wiley & Sons Limited. In checking the galley proofs, Theo Beekman, Jeroen van den Berg, Bert Gerards, Stan van Hoesel, Cor Hurkens, Hans Kremers, Fred Nieuwland, Henk Oosterhout, Joke Sterringa, Marno Verbeek, Hein van den Wildenberg, and Chris Wildhagen were of great assistance, and they certainly cannot be blamed for any surviving errors.

ALEXANDER SCHRIJVER

Contents

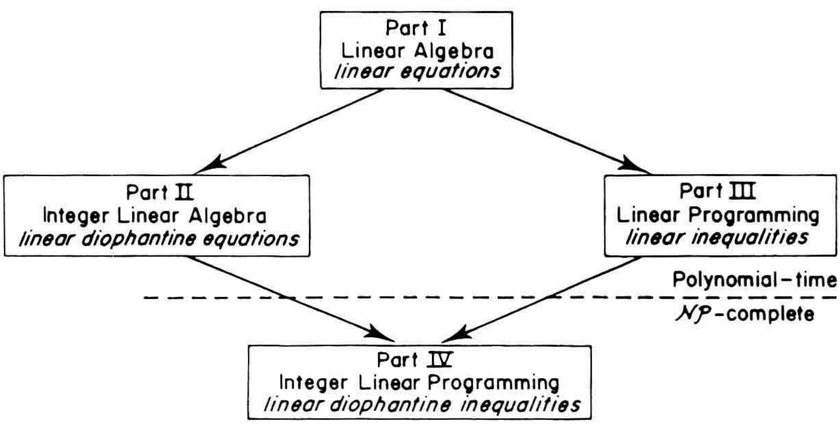
1

Introduction and preliminaries

After the introduction in Section 1.1, we discuss general preliminaries (Section 1.2), preliminaries on linear algebra, matrix theory and Euclidean geometry (Section 1.3), and on graph theory (Section 1.4).

1.1. INTRODUCTION

The structure of the theory discussed in this book, and of the book itself, may be explained by the following diagram.



In Part I, 'Linear Algebra', we discuss the theory of linear spaces and of systems of linear equations, and the complexity of solving these systems. The theory and methods here are to a large extent standard, and therefore we do not give an extensive treatment. We focus on some less standard results, such as sizes of solutions and the running time of the Gaussian elimination method. It is shown that this method is a *polynomial-time* method, i.e. its running time is bounded by a polynomial in the size of the input data.

In Part II, 'Lattices and Linear Diophantine Equations', our main problem is to solve systems of *linear diophantine equations*, i.e. to solve systems of linear equations in integer variables. The corresponding geometric notion is that of a *lattice*. The existence of solutions here is characterized with the help of the *Hermite normal form*. One linear diophantine equation can be solved in polynomial time with the classical *Euclidean algorithm*. More generally, also *systems* of linear diophantine equations can be solved in polynomial time, with methods due to Frumkin and Votyakov, von zur Gathen and Sieveking, and Kannan and Bachem.

Also in Part II we discuss the problem of *diophantine approximation*. The *continued fraction method* approximates a real number by a rational number with low denominator, and is related to the Euclidean algorithm. Its extension to more dimensions, i.e. approximating a real vector by a rational vector whose entries have one common low denominator, can be done with Lovász's *basis reduction method* for lattices. These techniques are also useful in linear and integer programming, as we shall see in Parts III and IV.

In Part III, 'Polyhedra, Linear Inequalities, and Linear Programming', our main problems are the following:

(1) – solving systems of linear inequalities;
– solving systems of linear equations in nonnegative variables;
– solving *linear programming* problems.

These three problems are equivalent in the sense that any method for one of them directly yields methods for the other two. The geometric notion corresponding to the problems is that of a *polyhedron*. Solutions of the problems (1) are characterized by *Farkas' lemma* and by the *Duality theorem of linear programming*.

The *simplex method* is the famous method for solving problems (1); it is fast in practice, and polynomial-time 'on the average', but no version of it could be proved to have polynomially bounded running time also in the worst case. It was for some time an open problem whether the problems (1) can be solved in polynomial time, until in 1979 Khachiyan showed that this is possible with the *ellipsoid method*. Although it does not seem to be a practical method, we spend some time on this method, as it has applications in combinatorial optimization. We also discuss briefly another polynomial-time method, due to Karmarkar.

The problems discussed in Parts I–III being solvable in polynomial time, in Part IV 'Integer Linear Programming' we come to a field where the problems in general are less tractable, and are *$\mathcal{NP}$-complete*. It is a general belief that these problems are not solvable in polynomial time. The problems in question are:

(2) – solving systems of linear diophantine inequalities, i.e. solving linear inequalities in integers;

- solving systems of linear equations in nonnegative integer variables;
- solving *integer linear programming* problems.

Again, these three problems are equivalent in the sense that any method for one of them yields also methods for the other two. Geometrically, the problems correspond to the intersection of a lattice and a polyhedron. So the problems discussed in Parts II and III meet here.

The theory we shall discuss includes that of characterizing the convex hull P_I of the integral vectors in a polyhedron P. The case $P = P_I$ generally gives rise to better-to-handle integer linear programming problems. This occurs when P is defined by a *totally unimodular* matrix, or, more generally, by a *totally dual integral* system of inequalities. Inter alia, we shall discuss (but not prove) a deep theorem of Seymour characterizing total unimodularity.

If P is not-necessarily equal to P_I, we can characterize P_I with the *cutting plane method*, founded by Gomory. This method is not a polynomial-time method, but it yields some insight into integer linear programming. We also discuss the result of Lenstra that for each fixed number of variables, the problems (2) are solvable in polynomial time.

The theory discussed in Part IV is especially interesting for combinatorial optimization.

Before Parts I–IV, we discuss in the present chapter some preliminaries, while in Chapter 2 we briefly review the complexity theory of problems and algorithms. In particular, we consider polynomiality as a complexity criterion.

1.2. GENERAL PRELIMINARIES

Some general notation and terminology is as follows. If α is a real number, then

(3) $\lfloor \alpha \rfloor$ and $\lceil \alpha \rceil$

denote the lower integer part and the upper integer part, respectively, of α.

The symbols $\mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$ denote the sets of integers, rationals, and real numbers, respectively. $\mathbb{Z}_+, \mathbb{Q}_+$ and $\mathbb{R}_+$ are the restrictions of these sets to the nonnegatives. We denote, for real numbers α and β,

(4) $\alpha | \beta$ if and only if α divides β, i.e. if and only if $\beta = \lambda\alpha$ for some integer λ.

Moreover, $\alpha \equiv \beta (\bmod \gamma)$ means $\gamma | (\alpha - \beta)$.

If $\alpha_1, \ldots, \alpha_n$ are rational numbers, not all equal to 0, then the largest rational number γ dividing each of $\alpha_1, \ldots, \alpha_n$ exists, and is called the *greatest common divisor* or *g.c.d.* of $\alpha_1, \ldots, \alpha_n$, denoted by

(5) $\text{g.c.d.}\{\alpha_1, \ldots, \alpha_n\}$

(so the g.c.d. is always positive). The numbers $\alpha_1, \ldots, \alpha_n$ are *relatively prime* if g.c.d. $\{\alpha_1, \ldots, \alpha_n\} = 1$.

We write $f(x) = O(g(x))$ for real-valued functions f and g, if there exists a constant C such that $f(x) \leqslant Cg(x)$ for all x in the domain.

If we consider an optimization problem like

(6) $\max\{\varphi(x)|x\in A\}$

where A is a set and $\varphi: A\rightarrow\mathbb{R}$, then any element x of A is called a *feasible solution* for the maximization problem. If A is nonempty, problem (6) is called *feasible*, otherwise *infeasible*. Similarly, a set of conditions is *feasible* (or *solvable*) if the conditions can be fulfilled all at the same time. Otherwise, they are called *infeasible* (or *unsolvable*). Any instance satisfying the conditions is called a *feasible solution*.

If the maximum (6) is attained, we say that the maximum *exists*, *is finite*, or *is bounded*. A feasible solution attaining the maximum is called an *optimum* (or *optimal*) *solution*. The maximum value then is the *optimum value*.

Similar terminology is used for minima.

A constraint is *valid* for a set S if each element in S satisfies this constraint.

'Left-hand side' and 'right-hand side' are sometimes abbreviated to *LHS* and *RHS*.

1.3. PRELIMINARIES FROM LINEAR ALGEBRA, MATRIX THEORY, AND EUCLIDEAN GEOMETRY

We assume familiarity of the reader with the elements of linear algebra, such as linear (sub)space, linear (in)dependence, rank, determinant, matrix, non-singular matrix, inverse, etc. As background references we mention Birkhoff and Mac Lane [1977], Gantmacher [1959], Lancaster and Tismenetsky [1985], Lang [1966a], Nering [1963], Strang [1980].

If $a=(\alpha_1,\ldots,\alpha_n)$ and $b=(\beta_1,\ldots,\beta_n)$ are row vectors, we write $a\leqslant b$ if $\alpha_i\leqslant\beta_i$ for $i=1,\ldots,n$. Similarly for column vectors. If A is a matrix, and x,b,y, and c are vectors, then when using notation like

(7) $Ax=b,\quad Ax\leqslant b,\quad yA=c$

we implicitly assume compatibility of sizes of A,x,b,y, and c. So as for (7), if A is an $m\times n$-matrix, then x is a column vector of dimension n, b is a column vector of dimension m, y is a row vector of dimension m, and c is a row vector of dimension n.

Similarly, if c and x are vectors, and if we use

(8) cx

then c is a row vector and x is a column vector, with the same number of components. So (8) can be considered as the inner product of c and x.

An *n-vector* is an n-dimensional vector.

If a is a row vector and β is a real number, then $ax=\beta$ and $ax\leqslant\beta$ are called a *linear equation* and a *linear inequality*, respectively. If vector x_0 satisfies a linear inequality $ax\leqslant\beta$, then the inequality is called *tight* (*for* x_0) if $ax_0=\beta$.

If A is a matrix, and b is a column vector, we shall call $Ax=b$ a *system of linear equations*, and $Ax\leqslant b$ a *system of linear inequalities*. The matrix A is called the *constraint matrix* of the system.

A system of linear inequalities can have several alternative forms, like

(9) $Ax \geqslant b$ (for $(-A)x \leqslant -b$)

$Ax \leqslant b, Cx \leqslant d$ $\left(\text{for} \begin{bmatrix} A \\ C \end{bmatrix} x \leqslant \begin{pmatrix} b \\ d \end{pmatrix}\right)$

$Ax = b$ (for $Ax \leqslant b, -Ax \leqslant -b$)

and so on.

If $A'x \leqslant b'$ arises from $Ax \leqslant b$ by deleting some (or none) of the inequalities in $Ax \leqslant b$, then $A'x \leqslant b'$ is called a *subsystem* of $Ax \leqslant b$. Similarly for systems of linear equations.

The identity matrix is denoted by I, where the order usually is clear from the context. If δ is a real number, then an *all-δ vector* (*all-δ matrix*) is a vector (matrix) with all entries equal to δ. So an *all-zero* and an *all-one vector* have all their entries equal to 0 and 1, respectively. $\mathbf{0}$ and 0 stand for all-zero vectors or matrices, and $\mathbf{1}$ stands for an all-one vector, all of appropriate dimension.

The transpose of a matrix A is denoted by A^{T}. We use $\|\cdot\|$ or $\|\cdot\|_2$ for the *Euclidean norm*, i.e.

(10) $$\|x\| := \|x\|_2 := \sqrt{x^{\mathsf{T}}x}.$$

$d(x, y)$ denotes the *Euclidean distance* of vectors x and y (i.e. $d(x, y) := \|x - y\|_2$), and $d(x, P)$ the *Euclidean distance* between x and a set P (i.e. $d(x, P) := \inf\{d(x, y) | y \in P\}$).

The *ball* with *centre* x and *radius* ρ is the set

(11) $$B(x, \rho) := \{y | d(x, y) \leqslant \rho\}.$$

A point $x \in \mathbb{R}^n$ is an *internal point* of $S \subseteq \mathbb{R}^n$ if there exists an $\varepsilon > 0$ such that

(12) $$B(x, \varepsilon) \subseteq S.$$

Other norms occurring in this text are the l_1- and the l_∞-norms:

(13) $$\|x\|_1 := |\xi_1| + \cdots + |\xi_n|$$
$$\|x\|_\infty := \max\{|\xi_1|, \ldots, |\xi_n|\}$$

for $x = (\xi_1, \ldots, \xi_n)$ or $x = (\xi_1, \ldots, \xi_n)^{\mathsf{T}}$.

An $m \times n$-matrix A is said to have *full row rank* (*full column rank*, respectively) if rank $A = m$ (rank $A = n$, respectively).

A *row submatrix* of a matrix A is a submatrix consisting of some rows of A. Similarly, a *column submatrix* of A consists of some columns of A.

A matrix $A = (\alpha_{ij})$ is called *upper triangular* if $\alpha_{ij} = 0$ whenever $i > j$. It is *lower triangular* if $\alpha_{ij} = 0$ whenever $i > j$. It is *strictly upper triangular* if $\alpha_{ij} = 0$ whenever $i \geqslant j$. It is *strictly lower triangular* if $\alpha_{ij} = 0$ whenever $i \leqslant j$. It is a *diagonal matrix* if $\alpha_{ij} = 0$ whenever $i \neq j$. The square diagonal matrix of order n, with the numbers $\delta_1, \ldots, \delta_n$ on its main diagonal, is denoted by

(14) $$\operatorname{diag}(\delta_1, \ldots, \delta_n).$$

For any subset T of $\mathbb{R}$, a vector (matrix) is called a *T-vector* (*T-matrix*) if its entries all belong to T. A vector or matrix is called *rational* (*integral*, respectively) if its entries all are rationals (integers, respectively).

A linear equation $ax = \beta$ or a linear inequality $ax \leqslant \beta$ is *rational* (*integral*) if a and β are rational (integral). A system of linear equations $Ax = b$ or inequalities $Ax \leqslant b$ is *rational* (*integral*) if A and b are rational (integral). A *rational polyhedron* is a polyhedron determined by rational linear inequalities, i.e. it is $\{x \in \mathbb{R}^n \mid Ax \leqslant b\}$ for some rational system $Ax \leqslant b$ of linear inequalities.

Lattice point is sometimes used as a synonym for integral vector. A vector or matrix is *$1/k$-integral* if its entries all belong to $(1/k)\mathbb{Z}$, i.e. if all entries are integral multiples of $1/k$.

Scaling a vector means multiplying the vector by a nonzero real number.

For any finite set S, we identify the function $x: S \to \mathbb{R}$ with the corresponding vector in $\mathbb{R}^S$. If $T \subseteq S$, the *incidence vector* or *characteristic vector* of T is the $\{0,1\}$-vector in $\mathbb{R}^S$, denoted by χ_T, satisfying

$$\text{(15)} \qquad \begin{aligned} &\chi_T(s) = 1 \quad \text{if } s \in T \\ &\chi_T(s) = 0 \quad \text{if } s \in S \backslash T. \end{aligned}$$

If S and T are finite sets, an $S \times T$*-matrix* is a matrix with rows and columns indexed by S and T, respectively. If A is an $S \times T$-matrix and $b \in \mathbb{R}^T$, the product $Ab \in \mathbb{R}^S$ is defined by:

$$\text{(16)} \qquad (Ab)_s := \sum_{t \in T} \alpha_{s,t} \beta_t$$

for $s \in S$ (denoting $A = (\alpha_{s,t})$ and $b = (\beta_t)$).

If $\mathscr{C}$ is a collection of subsets of a set S, the *incidence matrix* of $\mathscr{C}$ is the $\mathscr{C} \times S$-matrix M whose rows are the incidence vectors of the sets in $\mathscr{C}$. So

$$\text{(17)} \qquad \begin{aligned} &M_{T,s} = 1 \quad \text{if } s \in T \\ &M_{T,s} = 0 \quad \text{if } s \notin T \end{aligned}$$

for $T \in \mathscr{C}$, $s \in S$.

The *support* of a vector is the set of coordinates at which the vector is nonzero.

The *linear hull* and the *affine hull* of a set X of vectors, denoted by lin.hull X and aff.hull X, are given by

$$\text{(18)} \qquad \begin{aligned} \text{lin.hull}\, X &= \{\lambda_1 x_1 + \cdots + \lambda_t x_t \mid t \geqslant 0; x_1, \ldots, x_t \in X; \lambda_1, \ldots, \lambda_t \in \mathbb{R}\} \\ \text{aff.hull}\, X &= \{\lambda_1 x_1 + \cdots + \lambda_t x_t \mid t \geqslant 1; x_1, \ldots, x_t \in X; \lambda_1, \ldots, \lambda_t \in \mathbb{R}; \\ &\qquad \lambda_1 + \cdots + \lambda_t = 1\}. \end{aligned}$$

A set C of vectors is *convex* if it satisfies:

$$\text{(19)} \qquad \text{if } x, y \in C \text{ and } 0 \leqslant \lambda \leqslant 1, \text{ then } \lambda x + (1 - \lambda) y \in C.$$

The *convex hull* of a set X of vectors in the smallest convex set containing X, and is denoted by conv.hull X; so

$$\text{(20)} \qquad \begin{aligned} \text{conv.hull}\, X = \{&\lambda_1 x_1 + \cdots + \lambda_t x_t \mid t \geqslant 1; x_1, \ldots, x_t \in X; \\ &\lambda_1, \ldots, \lambda_t \geqslant 0; \lambda_1 + \cdots + \lambda_t = 1\}. \end{aligned}$$

A (*convex*) *cone* is a nonempty set of vectors C satisfying

(21) if $x, y \in C$ and $\lambda, \mu \geqslant 0$, then $\lambda x + \mu y \in C$.

The *cone generated by* a set X of vectors is the smallest convex cone containing X, and is denoted by cone X; so

(22) $$\text{cone } X = \{\lambda_1 x_1 + \cdots + \lambda_t x_t \mid t \geqslant 0; x_1, \ldots, x_t \in X; \lambda_1, \ldots, \lambda_t \geqslant 0\}.$$

If $S \subseteq \mathbb{R}^n$, then a function $f: S \to \mathbb{R}$ is *convex* if S is convex and $f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda) f(y)$ whenever $x, y \in S$ and $0 \leqslant \lambda \leqslant 1$. f is *concave* if $-f$ is convex.

Pivoting

If A is a matrix, say

(23) $$A = \begin{bmatrix} \alpha & b \\ c & D \end{bmatrix}$$

where α is a nonzero number, b is a row vector, c is a column vector, and D is a matrix, then *pivoting* over the *pivot element* (1, 1) means replacing A by the matrix

(24) $$\begin{bmatrix} -\alpha^{-1} & \alpha^{-1} b \\ \alpha^{-1} c & D - \alpha^{-1} cb \end{bmatrix}.$$

Pivoting over any other element of A is defined similarly.

Some inequalities

We recall the following well-known (in)equalities (cf. Beckenbach and Bellman [1983]). First the *Cauchy–Schwarz inequality*: if $c, d \in \mathbb{R}^n$ then

(25) $$c^{\mathrm{T}} d \leqslant \|c\| \cdot \|d\|.$$

If $b_1, \ldots, b_m$ are column vectors in $\mathbb{R}^n$, and B is the $n \times m$-matrix with columns $b_1, \ldots, b_m$, then

(26) $\sqrt{\det B^{\mathrm{T}} B}$ = the area of the parallelepiped spanned by $b_1, \ldots, b_m$.

This implies the *Hadamard inequality*:

(27) $$\sqrt{\det B^{\mathrm{T}} B} \leqslant \|b_1\| \cdot \cdots \cdot \|b_m\|.$$

In particular, if B is a square matrix, then

(28) $$|\det B| \leqslant \|b_1\| \cdot \cdots \cdot \|b_m\|.$$

(26) also implies that if A denotes the matrix with columns $b_1, \ldots, b_{m-1}$, and c is a vector orthogonal to $b_1, \ldots, b_{m-1}$, where c is in the space spanned by $b_1, \ldots, b_m$, then

(29) $$\sqrt{\det B^{\mathrm{T}} B} = \frac{|c^{\mathrm{T}} b_m|}{\|c\|} \sqrt{\det A^{\mathrm{T}} A}.$$