Klaus Glashoff
Sven-Åke Gustafson

Applied Mathematical Sciences 45

Linear Optimization and Approximation

Springer-Verlag
New York Heidelberg Berlin

Klaus Glashoff
Sven-Åke Gustafson

Linear Optimization and Approximation

An Introduction to the Theoretical Analysis and Numerical Treatment of Semi-infinite Programs

With 20 Illustrations

Springer-Verlag
New York Heidelberg Berlin

Klaus Glashoff
Universität Hamburg
Institut für Angewandte Mathematik
2 Hamburg 13
Bundestrasse 55
Federal Republic of Germany

Sven-Åke Gustafson
Department of Numerical Analysis and Computing Sciences
Royal Institute of Technology
S-10044 Stockholm 70
Sweden

and

Centre for Mathematical Analysis
Australian National University
P.O. Box 4
Canberra, ACT 2600
Australia

AMS Subject Classifications: 90C05, 49D35

Library of Congress Cataloging in Publication Data
Glashoff, Klaus, 1947–
Linear optimization and approximation.
(Applied mathematical sciences ; v. 45)
Translation of: Einführung in die lineare Optimierung.
Includes bibliographical references and index.
1. Mathematical optimization. 2. Duality theory.
(Mathematics) I. Gustafson, Sven-Ake, 1938–
II. Title. III. Series: Applied mathematical sciences
(Springer-Verlag New York Inc.) ; v. 45.
QA1.A647 vol. 45 510s [519.7'2] 83-647
[QA402.5]

Printed and bound by R.R. Donnelley & Sons, Harrisonburg, VA.
Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-**90857**-9 Springer-Verlag New York Heidelberg Berlin
ISBN 3-540-**90857**-9 Springer-Verlag Berlin Heidelberg New York

Preface

A linear optimization problem is the task of minimizing a linear real-valued function of finitely many variables subject to linear constraints; in general there may be infinitely many constraints. This book is devoted to such problems. Their mathematical properties are investigated and algorithms for their computational solution are presented. Applications are discussed in detail.

Linear optimization problems are encountered in many areas of applications. They have therefore been subject to mathematical analysis for a long time. We mention here only two classical topics from this area: the so-called uniform approximation of functions which was used as a mathematical tool by Chebyshev in 1853 when he set out to design a crane, and the theory of systems of linear inequalities which has already been studied by Fourier in 1823.

We will not treat the historical development of the theory of linear optimization in detail. However, we point out that the decisive breakthrough occurred in the middle of this century. It was urged on by the need to solve complicated decision problems where the optimal deployment of military and civilian resources had to be determined. The availability of electronic computers also played an important role. The principal computational scheme for the solution of linear optimization problems, the simplex algorithm, was established by Dantzig about 1950. In addition, the fundamental theorems on such problems were rapidly developed, based on earlier published results on the properties of systems of linear inequalities.

Since then, the interest of mathematicians and users in linear optimization has been sustained. New classes of practical applications are being introduced continually and special variants of the simplex algorithm and related schemes have been used for the computational treatment of practical problems of ever-growing size and complexity. The theory of "classical" linear optimization problems (with only finitely many linear constraints) had almost reached its final form around 1950; see e.g. the excellent book by A. Charnes, W. W. Cooper and A. Henderson (1953). Simultaneously there were great efforts devoted to the generalization and extension of the theory of linear optimization to new areas. Thus *nonlinear* optimization problems were attacked at an early date. (This area plays only a marginal role in our book.) Here, connections were found with the classical theory of Lagrangian multipliers as well as to the duality principles of mechanics. The latter occurred in the framework of convex analysis.

At the same time the theory of *infinite* linear optimization came into being. It describes problems with infinitely many variables and constraints. This theory also found its final form rapidly; see the paper by R. J. Duffin (1956).

A special but important class of infinite linear optimization problems are those problems where the number of variables is finite but the number of linear inequality constraints is arbitrary, i.e. may be infinite. This type of problem, which constitutes a natural generalization of the classical linear optimization problem, appears in the solution of many concrete examples. We have already mentioned the calculation of uniform approximation of functions which plays a major role in the construction of computer representations of mathematical expressions. Uniform approximation can also be successfully used in the numerical treatment of differential equations originating in physics and technological problems.

Using an investigation by Haar from 1924 as a point of departure, A. Charnes, W. W. Cooper and K. O. Kortanek in 1962 gave the fundamental mathematical results of the last-mentioned class of linear optimization problems (with the exception of those questions which were already settled by Duffin's theory).

This class of optimization problems, often called *semi-infinite programs*, will be the main topic of the present book. The "classical" linear optimization problems, called *linear programs*, will occur naturally as a special case.

Whether the number of inequality constraints is finite is a matter of minor importance in the mathematical theory of linear optimization problems. The great advantage of treating such a general class of problems, encompassing so many applications, need not, fortunately, be achieved by means of a correspondingly higher level of mathematical sophistication. In our account we have endeavored to use mathematical tools which are as simple as possible. To understand this book it is only necessary to master the fundamentals of linear algebra and n-dimensional analysis. (This theory is summarized in §2.) Since we have avoided all unnecessary mathematical abstractions, geometrical arguments have been used as much as possible. In this way we have escaped the temptation to complicate simple matters by introducing the heavy apparatus of functional analysis.

The central concept of our book is that of *duality*. Duality theory is not investigated for its own sake but as an effective tool, in particular for the numerical treatment of linear optimization problems.

Therefore all of Chapter II has been devoted to the concept of weak duality. We give some elementary arguments which serve to illustrate the fundamental ideas (primal and dual problems). This should give the reader a feeling for the numerical aspects of duality. In Chapter III we discuss some applications of weak duality to uniform approximation where the emphasis is again placed on numerical aspects.

The duality theory of linear optimization is investigated in Chapter IV. Here we prove theorems on the existence of solutions to the optimization problems considered. We also treat the so-called strong duality, i.e. the question of equality of the values of the primal and dual problems. The "geometric" formulation of the dual problem, introduced here, will be very useful for the presentation of the simplex algorithm which is described in the chapter to follow.

In Chapter V we describe in great detail the principle of the exchange step which is the main building block of the simplex algorithm. Here we dispense with the computational technicalities which dominate many presentations of this scheme. The nature of the simplex algorithm can be explained very clearly using duality theory and the language of matrices and without relying on "simplex tableaux", which do not appear in our text.

In Chapter VI we treat the numerical realization of the simplex algorithm. It requires that a sequence of linear systems of equations be solved. Our presentation includes the stable variants of the simplex method which have been developed during the last decade.

In Chapter VII we present a method for the computational treatment of a general class of linear optimization problems with infinitely many constraints. This scheme was described for the first time in Gustafson (1970). Since then it has been successfully used for the solution of many practical problems, e.g. uniform approximation over multidimensional domains (also with additional linear side-conditions), calculation of quadrature rules, control problems, and so on.

In Chapter VIII we apply the ideas of the preceding three chapters to the special problem of uniform approximation over intervals. The classical Remez algorithm is studied and set into the general framework of linear optimization.

The concluding Chapter IX contains several worked examples designed to elucidate the general approach of this book. We also indicate that the ideas behind the computational schemes described in our book can be applied to an even more general class of problems.

The present text is a translated and extended version of Glashoff-Gustafson (1978). Chapters VIII and IX are completely new and Chapter IV is revised. More material has been added to Chapters III and VII. These changes and additions have been carried out by the second author, who is also responsible for the translation into English. Professor Harry Clarke, Asian Institute of Technology, Bangkok, has given valuable help with the latter task.

We hope that this book will provide theoretical and numerical insights which will help in the solution of practical problems from many disciplines. We also believe that we have clearly demonstrated our conviction that mathematical advances generally are inspired by work on real world problems.

Table of Contents

		Page
Preface		v
CHAPTER I.	INTRODUCTION AND PRELIMINARIES	1
	§1. Optimization Problems	1
	§2. Some Mathematical Prerequisites	5
	§3. Linear Optimization Problems	10
CHAPTER II.	WEAK DUALITY	20
	§4. Duality Lemma and Dual Problem	20
	§5. State Diagrams and Duality Gaps	30
CHAPTER III.	APPLICATIONS OF WEAK DUALITY IN UNIFORM APPROXIMATION	37
	§6. Uniform Approximation	37
	§7. Polynomial Approximation	46
CHAPTER IV.	DUALITY THEORY	58
	§8. Geometric Interpretation of the Dual Problem	58
	§9. Solvability of the Dual Problem	69
	§10. Separation Theorem and Duality	75
	§11. Supporting Hyperplanes and Duality	82
CHAPTER V.	THE SIMPLEX ALGORITHM	92
	§12. Basic Solutions and the Exchange Step	93
	§13. The Simplex Algorithm and Discretization	105
CHAPTER VI.	NUMERICAL REALIZATION OF THE SIMPLEX ALGORITHM	115
	§14. Stable Variants of the Simplex Algorithm	115
	§15. Calculating a Basic Solution	128
CHAPTER VII.	A GENERAL THREE-PHASE ALGORITHM	134
	§16. Nonlinear Systems Derived From Optimality Conditions	135
	§17. A General Computational Scheme	141
CHAPTER VIII.	APPROXIMATION PROBLEMS BY CHEBYSHEV SYSTEMS	153
	§18. General Properties of Chebyshev Systems	153
	§19. One-sided Approximation and Generalized Quadrature Rules of the Gaussian Type	158
	§20. Computing the Best Approximation in the Uniform Norm	168
CHAPTER IX.	EXAMPLES AND APPLICATIONS OF SEMI-INFINITE PROGRAMMING	175
	§21. A Control Problem with Distributed Parameters	175
	§22. Operator Equations of Monotonic Type	181
	§23. An Air Pollution Abatement Problem	184
	§24. Nonlinear Semi-Infinite Programs	188
References		193
Index		196

Chapter I
Introduction and Preliminaries

§1. OPTIMIZATION PROBLEMS

Optimization problems are encountered in many branches of technology, in science, and in economics as well as in our daily life. They appear in so many different shapes that it is useless to attempt a uniform description of them or even try to classify them according to one principle or another. In the present section we will introduce a few general concepts which occur in all optimization problems. Simple examples will elucidate the presentation.

(1) Example: Siting of a power plant. Five major factories are located at $P_1, P_2, \ldots, P_5$. A power plant to supply them with electricity is to be built and the problem is to determine the optimal site for this plant. The transmission of electrical energy is associated with energy losses which are proportional to the amount of transmitted energy and to the distance between power plant and energy consumer. One seeks to select the site of the plant so that the combined energy loss is rendered a minimum. $P_1, P_2, \ldots, P_5$ are represented by points in the plane with the coordinates

$$P_1 = (x_1, y_1), \ldots, P_5 = (x_5, y_5).$$

The distance between the two points $P = (x,y)$, $\bar{P} = (\bar{x}, \bar{y})$ is given by

$$d(P,\bar{P}) = \{(x-\bar{x})^2 + (y-\bar{y})^2\}^{1/2}.$$

Denote the transmitted energy quantities by $E_1, \ldots, E_5$. Our siting problem may now be formulated. We seek, within a given domain G of the plane, a point $\bar{P}$ such that the following function assumes its minimal value at $\bar{P}$:

$$E_1 d(P,P_1) + E_2 d(P,P_2) + \ldots + E_5 d(P,P_5).$$

In order to introduce some terminology we reformulate this task. We define the real-valued function f of two real variables x,y through

$$f(x,y) = E_1\{(x-x_1)^2 + (y-y_1)^2\}^{1/2} + \ldots + E_5\{(x-x_5)^2 + (y-y_5)^2\}^{1/2}.$$

We then arrive at the optimization problem: Determine numbers $\bar{x},\bar{y}$ such that $\bar{P} = (\bar{x},\bar{y}) \in G$ and

$$f(\bar{x},\bar{y}) \leq f(x,y) \quad \text{for all} \quad (x,y) \in G.$$

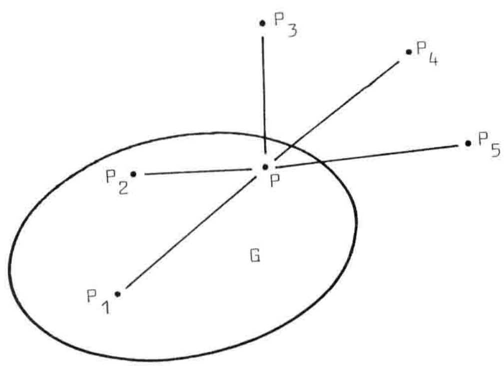

Fig. 1.1. Siting of power plant

All important concepts associated with optimization problems may be illustrated by this example: f is called a *preference function*, G is the *permissible set*, and the points of G are called *permissible* or *feasible*. Thus the optimization problem means that one should seek a permissible point such that f assumes its minimal value with respect to the permissible set. If such a point does exist, it is called an *optimal point* (for the problem considered), or *optimal solution*, or *minimum point* of f in G.

In the analysis of an optimization problem it is important to verify that an optimal solution does exist, i.e. that the problem is solvable. This is not always the case. As an illustration of this fact we note that the functions $f_1(x) = -x$ and $f_2(x) = e^{-x}$ do not have any minimum points in the set of all real numbers. On the other hand, if an optimization problem is solvable, a minimum point may not be *unique*. In many applications it is required to determine *all* minimum points which the preference function has in the permissible set.

It is of course of no use to formulate a task, appearing in economics or technology, as an optimization problem when this problem cannot be solved. A formulation as an optimization problem is thus advantageous only when the mathematical structure of this task can be investigated and suitable theoretical and computational tools can be brought to bear. Oftentimes, "applications" to economics or management are proposed whereby very complicated optimization problems are constructed but it is not pointed out that neither theoretical nor numerical treatment of the problem appears to be within reach, now or in the near future. It should always be remembered that only *some* of the relevant factors can be incorporated when a decision problem is formulated as an optimization problem. There are always decision criteria which cannot be quantified and whose inclusion into a mathematical model is of doubtful value. Thus, in the siting problem discussed above, there are many political and ecological factors which cannot be accounted for in a mathematical model. This indicates that there is, in principle, a limit of what can be gained by the mathematization of social processes. This difficulty cannot, as a rule, be overcome by resorting to more complicated models (control theory, game theory, etc.) even if it sometimes may be concealed. The situation is quite different for technical systems. Since nowadays the mathematization and also the "optimization" of social processes are pushed forward with great energy, we find the critical remark above to be justified.

(2) Example: Production model. We consider a firm which produces or consumes n goods $G_1,\ldots,G_n$ (e.g. raw materials, labor, capital, environmental pollutants). An *activity* of the firm is represented by n numbers $(a_1,\ldots,a_n)$ where a_r indicates the amount of good G_r which is produced or consumed when the activity P is taking place with intensity 1 (measured in suitable units). We assume that the firm can select various activities P_s. Thus the firm's technology has the property that to each s in a fixed index set S (which may be finite or infinite) there are n numbers $(a_1(s),\ldots,a_n(s))$. A *production plan* of the firm is defined by selecting a (finite) number of activities $P_{s_1},\ldots,P_{s_q}$ and prescribing that they are carried out with the *intensities* $x_1,\ldots,x_q$, where $x_i \geq 0$, $i = 1,2,\ldots,q$. We assume that the production process is linear, i.e. for the given production plan the amount of good G_r which is produced or consumed is given by

$$a_r(s_1)x_1 + a_r(s_2)x_2 + \ldots + a_r(s_q)x_q.$$

We shall further assume that the activity P_s causes the profit (or cost) $b(s)$. Hence the profit achieved by the chosen production plan is given by

$$b(s_1)x_1 + b(s_2)x_2 + \ldots + b(s_q)x_q. \qquad (3)$$

The optimization problem of the firm is to maximize its profit by proper choice of its production plan, i.e. it must select finitely many activities $P_{s_1},\ldots,P_{s_q}$ and the corresponding intensities $x_1,x_2,\ldots,x_q$ such that the expression (3) assumes the greatest value possible.

The choice of activities and intensities is restricted by the fact that only finite amounts of the goods $G_1,\ldots,G_n$ are available. In practice this is true only for *some* of the goods but for simplicity of presentation we want to assume that *all* goods can only be obtained in limited amounts:

$$a_r(s_1)x_1 + a_r(s_2)x_2 + \ldots + a_r(s_q)x_q \leq c_r, \quad r = 1,2,\ldots,n. \qquad (4)$$

Thus (4) defines n *side-conditions* which constrain the feasible activities and intensities. The optimization problem can thus be cast into the form: Determine a finite subset $\{s_1,\ldots,s_q\}$ of the index set S and the real numbers $x_1,\ldots,x_q$ such that the expression (3) is rendered a maximum under the constraints (4) and the further side-conditions

$$x_i \geq 0, \quad i = 1,2,\ldots,q.$$

(5) Remark. A *maximization* problem is transformed into an equivalent *minimization* problem by multiplying its preference function by -1.

(6) The general optimization problem. Let M be a fixed set and let f be a real-valued function defined on M. We seek an element $\bar{x}$ in M such that

$$f(\bar{x}) \leq f(x) \quad \text{for all} \quad x \in M.$$

M is called the *feasible* or *permissible set* and f is termed the *preference function*. We remark here that the feasible set is, as a rule, not explicitly given but is defined through side-conditions (often called *constraints*), as in Example (2).

(7) Definition. The number v given by

$$v = \{\inf f(x) \mid x \in M\}$$

is called the *value* of the corresponding optimization problem. If M is the empty set, i.e. there are no feasible points, the optimization problem is said to be *inconsistent* and we put $v = \infty$. If feasible points do exist we term the optimization problem *feasible* or *consistent*. If $v = -\infty$, the optimization problem is said to be "unbounded from below". Thus every minimization problem must be in one and only one of the following three "states" IC, B, UB:

IC = Inconsistent; the feasible set is empty and the value of the problem is $+\infty$.

B = Bounded; there are feasible points and the value is finite.

UB = Unbounded; there are feasible points, the preference function is unbounded from below, and the value is $-\infty$.

The value of a *maximization problem* is $-\infty$ in the state IC, finite in state B, and $+\infty$ in the state UB.

§2. SOME MATHEMATICAL PREREQUISITES

The successful study of this book requires knowledge of some elementary concepts of mathematical analysis as well as linear algebra. We shall summarize the notations and some mathematical tools in this section.

(1) Vectors. We denote the field of real numbers by R, and by R^n the n-dimensional space of all n-tuples of real numbers

$$\begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix}. \tag{2}$$

In R^n, the usual vector space operations are defined: componentwise addition of vectors and multiplication by scalars (i.e. real numbers).

We assume that the reader is familiar with the concepts of "linear independence", "basis", and "subspace". The zero vector of R^n is written 0. n-tuples of the form (2) are also referred to as "points".

(3) Matrices. An $m \times n$ matrix A $(m \geq 1)$ is a rectangular array of $m \cdot n$ real numbers a_{ik} $(i = 1,2,\ldots,m,\ k = 1,2,\ldots,n)$,

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

The numbers a_{ik} are termed the *elements* of the matrix A and a_{ik} is situated in row number i and column number k. To each given matrix A we define its *transpose* A^T by

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

Every vector $x \in R^n$ may be considered an $n \times 1$ matrix. In order to save space we write, instead of (2),

$$x^T = (x_1, x_2, \ldots, x_n).$$

We note that $(A^T)^T = A$. The reader is supposed to know elementary matrix operations (addition and multiplication of matrices).

(4) Linear mappings. Every $m \times n$ matrix A defines a linear mapping of R^n into R^m whereby every vector $x \in R^n$ is mapped onto a vector $y \in R^m$ via

$$y = Ax. \tag{5}$$

Using the definition of matrix multiplication we find that the components of y are to be calculated according to

$$y_i = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n, \quad 1 \le i \le m.$$

Denote the column vectors of A by $a_1, a_2, \ldots, a_n$. Then we find

$$Ax = a_1x_1 + a_2x_2 + \ldots + a_nx_n. \tag{6}$$

Equation (6) thus means that the vector y is a linear combination of the column vectors of A.

(7) Linear systems of equations. Now let a fixed y be given in (5). The task of determining x in (5) is one of the fundamental problems of linear algebra. (5) is called a linear system of equations with n unknowns $x_1, x_2, \ldots, x_n$ and m equations. We assume that the solvability theory of (5) (existence and uniqueness of solutions) is known to the reader. An example: from (6) we conclude that (5) is solvable for each $y \in R^m$ if the column vectors of A span all of R^m, i.e. if A has the *rank* m. It is equally simple to verify that (5) has *at most* one solution if the column vectors of A are linearly independent. The case when A is a *square* matrix, $n \times n$, is of particular interest. Then (5)

has an equal number of equations and unknowns. Then the linear system $Ax = y$ has a unique solution $x \in R^n$ for each $y \in R^n$ if and only if the column vectors $a_1, a_2, \ldots, a_n$ of A form a basis of R^n, i.e. are linearly independent. Then the matrix A is said to be *regular* (or *nonsingular*). In this case there exists a $n \times n$ matrix A^{-1} with the properties

$$A^{-1}(Ax) = x, \quad A(A^{-1}x) = x, \quad \text{all} \quad x \in R^n.$$

A^{-1} is called the *inverse* of A and the linear system of equations (5) has the unique solution

$$x = A^{-1}y.$$

(8) Hyperplanes. A vector $y \in R^n$ and a number $\eta \in R$ are given. Then we denote by the *hyperplane* $H(y;\eta)$ the set of all points $x \in R^n$ such that

$$y^T x = y_1x_1 + y_2x_2 + \ldots + y_nx_n = \eta.$$

y is called the *normal vector* of the hyperplane. For any two vectors x and z in $H(y;\eta)$ we have

$$y^T(x-z) = 0.$$

A hyperplane $y^T x = \eta$ partitions R^n into three disjoint sets, namely $H(y;\eta)$ and the two "open half-spaces"

$$A_1 = \{x \mid y^T x < \eta\}$$
$$A_2 = \{x \mid y^T x > \eta\}.$$

The linear system of equations (5) also admits the interpretation that the vector x must be in the intersection of the hyperplanes $H(a^i;y_i)$, $(i = 1,2,\ldots,m)$, where $a^1,\ldots,a^m$ here are the row-vectors of the matrix A. Sets of the form $A_1 \cup H(\eta;y)$ and $A_2 \cup H(\eta;y)$ are termed *closed half-spaces*. They consist of all points $x \in R^n$ such that

$$y^T x \leq \eta \quad \text{or} \quad y^T x \geq \eta,$$

respectively.

(9) Vector norms. We shall associate with each vector $x \in R^n$ a real number $||x||$. The mapping $x \to ||x||$ shall obey the following laws:

(i) $||x|| \geq 0$, all $x \in R^n$ and $||x|| = 0$ for $x = 0$ only;
(ii) $||\lambda x|| = |\lambda|\ ||x||$, all $x \in R^n$, all $\lambda \in R$;
(iii) $||x+y|| \leq ||x|| + ||y||$, all $x \in R^n$, $y \in R^n$.

Then $||x||$ will be called the *norm* of the vector.

Exercise: Show that the following mappings define vector norms on R^n:

$$x \to |x_1| + |x_2| +\ldots+ |x_n|$$
$$x \to \max\{|x_1|,|x_2|,\ldots,|x_n|\}.$$

The most well-known norm is the Euclidean norm, which will be treated in the next subsection.

(10) Scalar product and Euclidean norm. The *scalar product* of two vectors x and y is defined to be the real number

$$x^Ty = y^Tx = x_1y_1 + x_2y_2 +\ldots+ x_ny_n.$$

The real number

$$|x| = \sqrt{x^Tx} = (x_1^2 + x_2^2 +\ldots+ x_n^2)^{1/2}$$

is called the *Euclidean norm* or *length* or *absolute value* of the vector x. The reader should verify that the mapping $x \to |x|$ defines a norm in the sense of (9). It is also easy to establish the "parallelogram law"

$$|x+y|^2 + |x-y|^2 = 2(|x|^2 + |y|^2) \quad \text{for all } x,y \in R^n.$$

(11) Some topological fundamentals. We define the *distance* between two points x,y in R^n to be given by $|x-y|$. The set $K_r(a)$ consisting of all points whose distance to a is less than r, a fixed positive number, is termed the *open sphere* with *center* a and *radius* r. Thus

$$K_r(a) = \{x \in R^n \mid |x-a| < r\}.$$

We are now in a position to introduce the fundamental topological structure of R^n. A point a is said to be an *inner point* of a subset $A \subset R^n$ if there is a sphere $K_r(a)$ which in its entirety belongs to A, $K_r(a) \subset A$. We will use the symbol $\overset{\circ}{A}$ for the set of all inner points of A. $\overset{\circ}{A}$ is also called the *interior* of A. A is termed *open* if $A = \overset{\circ}{A}$. The point a is said to be a *boundary point* of the set A if every sphere $K_r(a)$ contains both points in A and points which do not belong to A. The set of all boundary points of A is called the *boundary* of A

and is denoted bd A. The union of A and its boundary is called the *closure* of A and is denoted $\bar{A}$. The set A is said to be *closed* if $A = \bar{A}$. The following relations always hold.

$$\overset{\circ}{A} \subset A \subset \bar{A}, \quad \text{bd } A = \bar{A} \setminus \overset{\circ}{A}.$$

The topological concepts introduced above have been defined using the Euclidean norm. This norm will be most often used in the sequel. However, one may define spheres in terms of other norms and in this way arrive at the fundamental topological concepts "inner points", "open sets", and so on, in the same manner as above. Fortunately it is possible to prove that all norms on R^n are *equivalent* in the sense that they generate the same topological structure on R^n: A set which is open with respect to one norm remains open with respect to all other norms. In order to establish this assertion one first verifies that if $||\cdot||_1$ and $||\cdot||_2$ are two norms on R^n there are two positive constants c and C such that

$$c||x||_1 \leq ||x||_2 \leq C||x||_1 \quad \text{for all} \quad x \in R^n.$$

Based on these fundamental structures one can now define the main concept of convergence of sequences and continuity of functions in the usual way. We suppose here the reader is familiar with these concepts.

(12) Compact sets. A subset $A \subset R^n$ is said to be *bounded* when there is a real number $r > 0$ such that $A \subset K_r(0)$. Closed bounded subsets of R^n will be termed *compact*.

Compact subsets A of R^n have the following important property: Every infinite sequence $\{x_i\}_{i>1}$ of points in the set A has a convergent subsequence $\{x_{i_k}\}_{k>1}$. If $f: R^n \to R^m$ is a continuous mapping, then the image f(A) of every compact set A is compact also. From this statement we immediately arrive at the following result which also may be looked upon as an existence statement for optimization problems:

(13) Theorem of Weierstrass. Let A be a nonempty compact subset of R^n and f a real-valued continuous function defined on A. Then f assumes its maximum and minimum value on A, i.e. there exist points $\bar{x} \in A$ and $\tilde{x} \in A$ such that

$$f(\bar{x}) = \max\{f(x) \mid x \in A\}$$

and

$$f(\tilde{x}) = \min\{f(x) \mid x \in A\}.$$

It is recommended that the reader, as an exercise, carry out the proof of this simple but important theorem.