Optimization Methods

in operations research and systems analysis

K V Mital

Optimization Methods

in operations research and
systems analysis

K. V. MITAL
University of Roorkee

WILEY EASTERN LIMITED
New Delhi Bangalore Bombay Calcutta

10 9 8 7 6 5 4 3

ISBN 0 85226 582 4

Published by Ravi Acharya for Wiley Eastern Limited, 4835/24 Ansari Road, Daryaganj, New Delhi 110002 and printed by Pramodh Kapur at Raj Bandhu Industrial Company, C 61 Maya Puri II, New Delhi 110064. Printed in India.

OPTIMIZATION METHODS

To Father
whose memory lingers

Preface

The subject which started as operations research during the last World War in the early forties has been growing theoretically and also in its applications to a variety of problems in diverse fields, such as, engineering, management and economics. In its more comprehensive sense which includes survey and data collection, mathematical modelling, solutions of large mathematical problems and improvements through feedback of results, the subject has come to be known as systems analysis. The mathematical contents of this subject concerned with optimization of objectives may be grouped more expressively under optimization methods which form the subject matter of this volume.

This book is an elementary mathematical introduction to linear and nonlinear programming, dynamic programming, geometric programming, direct search methods and theory of games. It has grown out of lectures given to M.Sc. and M.E. classes and to short-term courses under the Refresher Courses Department and the Quality Improvement Programme at the University of Roorkee during the last six years. Only deterministic problems have been dealt with. Stochastic problems have not been touched. The book is intended to serve as a suitable text for students of mathematics, operations research, engineering, economics or management whose courses of study include some or all of the topics treated here. Most of the chapters can be studied independently of each other. A knowledge of algebra (including matrices), calculus and geometry as is usually given in the B.Sc. and the B.E. courses in India is assumed. Chapters I and II provide the additional necessary topics in mathematics. Convex sets have been treated in some detail as they are not included in the usual mathematics courses given to our students but are fundamental to the theory of mathematical programming.

Bibliography at the end lists a number of books, mostly recent, on the various topics discussed in this volume. A short bibliographical note at the end of each chapter is meant to guide the reader to a few standard books which he may profitably consult either along with the present book or subsequently for more advanced study. No references are made to research papers, as it is seldom profitable for students to go direct to them without acquiring a working knowledge of the established concepts. A short historical note at the end of some of the chapters is meant to acquaint the student with the pioneer workers in the field. For original research papers and credits, books listed in the bibliography can be usefully consulted.

In their application to real life, problems in systems analysis and operations research usually involve large number of variables, parameters, equations and constraints. The problems generally involve too much numerical work which can be handled only by the digital computer. For this reason the methods of solution are computer oriented. The criterion of suitability of a method is often the economy and efficiency with which it can be programmed on the computer. In this book we are not concerned with computer programming. The illustrative examples in the text and also the problems at the end of each chapter are small enough to be solved by hand and may not apparently justify the methods recommended to solve them. But the student should not lose sight of the fact that the problems are only illustrative and the methods are really designed for large problems of the same type.

It is my pleasant duty to acknowledge gratefully the generous help I have received from many colleagues and friends in the preparation of this book. I am particularly grateful to Professor C. Prasad, Dr. O.P. Varshney and Dr. A.P. Gupta for their assistance in chapters I, II and VI; to Dr. U.S. Gupta for chapter III; to Dr. C. Mohan for chapters VIII and X; to Dr. R.K. Gupta for chapter IX; and to Dr. Bal Krishna for critically reading through chapter I. It is a truism that a teacher learns through his students. I am thankful to all my students of M.Sc. and M.E. classes and all those participants of special short-term courses who have attended my lectures over the last many years. Without their knowing it, they taught me a great deal and have contributed in some measure to the writing of this book.

K. V. MITAL

Roorkee
June 26, 1976

Contents

II EXTREMA OF FUNCTIONS

III LINEAR PROGRAMMING

I / Mathematical Preliminaries

EUCLIDEAN SPACE

1 Vectors and vector spaces

A mathematical model of a system may contain n variables $x_1, x_2, \ldots, x_n$, each of which may vary within a subset of real numbers R. A collection of n real numbers, taken in order, such that the first number is the value of x_1, the second of x_2, and so on, is called an ordered n-tuple of real numbers. We may denote an ordered n-tuple by a single symbol $\mathbf{X}$, so that

$$\mathbf{X}=(x_1, x_2, \ldots, x_n),$$

and a set of such n-tuples by R_n, so that

$$R_n=\{\mathbf{X} \mid \mathbf{X}=(x_1, x_2, \ldots, x_n)\}.$$

In order to deal with such sets it is convenient to establish an analogy with geometrical concepts which are easy to visualize. We therefore assume that $\mathbf{X}$ and R_n satisfy certain postulates which are generalizations of notions familiar in two- and three-dimensional geometry, and then call $\mathbf{X}$ a vector and R_n a vector space. We start with general definitions of vector and vector space.

DEFINITION 1. *Let V be a set such that if $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \in V$ and $a, b \in R$, then the following postulates (defining the binary operation of sum and the operation of product with a real number) hold.*

Sum:

(i) $\mathbf{X}+\mathbf{Y} \in V$;

(ii) $\mathbf{X}+\mathbf{Y}=\mathbf{Y}+\mathbf{X}$;

(iii) $(\mathbf{X}+\mathbf{Y})+\mathbf{Z}=\mathbf{X}+(\mathbf{Y}+\mathbf{Z})$;

(iv) *There exists an element $\mathbf{0} \in V$, called the null or zero vector, such that $\mathbf{X}+\mathbf{0}=\mathbf{X}$;*

(v) *There exists an element $-\mathbf{X} \in V$, called the additive inverse of $\mathbf{X}$, such that $\mathbf{X}+(-\mathbf{X})=\mathbf{0}$;*

Product:

(vi) $a\mathbf{X} \in V$;

(vii) $a(\mathbf{X}+\mathbf{Y})=a\mathbf{X}+a\mathbf{Y}$;
(viii) $(a+b)\,\mathbf{X}=a\mathbf{X}+b\mathbf{X}$;
(ix) $(ab)\,\mathbf{X}=a\,(b\mathbf{X})$;
(x) $1\mathbf{X}=\mathbf{X}$.

Then V is called a vector space and its elements are called vectors. Throughout this chapter V shall denote a vector space.

Example: Let V be a set of all polynomials in x of degree n or less.

$$V=\{f_1(x), f_2(x), \ldots, f_i(x), \ldots\},$$

where

$$f_i(x)=\sum_{j=1}^{n} a_{ij}\, x^j,\ a_{ij}\in R.$$

If the two operations be the usual operations of sum and product by a real number, then it can be verified that the postulates (i) to (x) hold, and so V is a vector space. The additive inverse of $f_i(x)$ and the zero vector can be easily identified.

Example: Let $\mathbf{X}=(x_1, x_2, \ldots, x_n)$ be an ordered n-tuple of real numbers and R_n be the set of all such n-tuples. If we define the sum of two n-tuples as

$$\mathbf{X}+\mathbf{Y}=(x_1, x_2, \ldots, x_n)+(y_1, y_2, \ldots, y_n)$$
$$=(x_1+y_1, x_2+y_2, \ldots, x_n+y_n),$$

and the product with a real number a as

$$a\mathbf{X}=(ax_1, ax_2, \ldots, ax_n),$$

then it can be verified that R_n is a vector space. The zero vector of the space is $(0, 0, \ldots, 0)$.

In matrix operations it is convenient to regard $\mathbf{X}$ as a column vector:

$$\mathbf{X}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=[x_1\, x_2 \ldots x_n]'$$

Its transpose $\mathbf{X}'$ is a row vector. To indicate that $\mathbf{X}$ has n components it is customary to call $\mathbf{X}$ an *n-vector*.

DEFINITION 2. *A subset W of V is called a subspace of the vector space V if W is itself a vector space with respect to the operations of sum and product defined in V.*

Example: Let $W\subseteq R_n$ (of last example) such that $W=\{\mathbf{X}\,|\,\mathbf{X}=(x_1, 0, x_3, \ldots, x_n)\}$. Then W is a subspace of R_n. For, if $\mathbf{X}, \mathbf{Y}\in W$,

$$\mathbf{X}+\mathbf{Y}=(x_1+y_1,\ 0,\ x_3+y_3, \ldots, x_n+y_n)\in W,$$
$$a\mathbf{X}=(ax_1, 0, ax_3, \ldots, ax_n)\in W,\ a\in R,$$

and similarly all other postulates can be seen to hold.

2 Linear dependence

DEFINITION 3. *Let* $\mathbf{X}_i$, $1 \leqslant i \leqslant m$, *be vectors of* V. *Then* $\mathbf{X}$ *is called a linear combination of the vectors* $\mathbf{X}_i$ *if*

$$\mathbf{X} = \sum_{i=1}^{m} a_i \mathbf{X}_i, \; a_i \in R.$$

THEOREM 1. *The set* W *of all linear combinations of* $\mathbf{X}_i$, $1 \leqslant i \leqslant m$, *of vector space* V *is a subspace of* V.

Proof. Let $\mathbf{X} = \sum_{i=1}^{m} a_i \mathbf{X}_i$, $\mathbf{Y} = \sum_{i=1}^{m} b_i \mathbf{X}_i$, so that $\mathbf{X}, \mathbf{Y} \in W$. Then

$$\mathbf{X} + \mathbf{Y} = \sum_{i=1}^{m} (a_i + b_i) \mathbf{X}_i \in W,$$

$$\lambda \mathbf{X} = \sum_{i=1}^{m} (\lambda a_i) \mathbf{X}_i \in W, \; \lambda \in R.$$

It follows (see problem 1) that W is a vector space and hence a subspace of V. Proved.

W is said to be spanned by (or a span of) $\mathbf{X}_i$.

DEFINITION 4. *The vectors* $\mathbf{X}_i$, $1 \leqslant i \leqslant m$, *of* V *are said to be linearly dependent if there exist real numbers* a_i, *not all zero, such that*

$$\sum_{i=1}^{m} a_i \mathbf{X}_i = 0.$$

If, however, this is so only if $a_i = 0$ for all i, then the vectors are said to be *linearly independent*.

Example: Let $\mathbf{X}_1 = [2 \; -1 \; 3 \; 2]'$, $\mathbf{X}_2 = [1 \; 2 \; 2 \; -4]'$, $\mathbf{X}_3 = [4 \; 3 \; 7 \; -6]'$. Since $\mathbf{X}_1 + 2\mathbf{X}_2 - \mathbf{X}_3 = 0$, the vectors are linearly dependent. But $\mathbf{X}_1$ and $\mathbf{X}_2$ are linearly independent. So are $\mathbf{X}_2$, $\mathbf{X}_3$.

To test whether the n-vectors $\mathbf{X}_i$, $1 \leqslant i \leqslant m$, are linearly independent or not, one has to examine the equations

$$\sum_{i=1}^{m} a_i \mathbf{X}_i = 0,$$

or putting $\mathbf{X}_i$ as $[x_{1i} \; x_{2i} \ldots x_{ni}]'$,

$$\begin{aligned} a_1 x_{11} + a_2 x_{12} + \ldots + a_m x_{1m} &= 0, \\ a_1 x_{21} + a_2 x_{22} + \ldots + a_m x_{2m} &= 0, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot & \quad \cdot \quad \cdot \\ a_1 x_{n1} + a_2 x_{n2} + \ldots + a_m x_{nm} &= 0, \end{aligned}$$

and investigate whether values of a_i, not all zero, exist which satisfy these equations, or in other words, whether the n equations in m unknowns a_i have a nontrivial solution. (A solution $a_i = 0$ for all i is called a trivial solution.) We shall discuss the solution of such equations later in this chapter.

3 Dimension of a vector space, basis

DEFINITION 5. *V is said to be of dimension m if there exists at least one set of m linearly independent vectors in V, while every set of $m+1$ vectors in V is linearly dependent. The linearly independent set is called a basis of V.*

THEOREM 2. *A set of m linearly independent vectors in a vector space V of dimension m spans V.*

Proof. Let $\mathbf{Y}_i$, $1 \leqslant i \leqslant m$, be m linearly independent vectors in V, and let $\mathbf{X}$ be any vector in V. Since V is of dimension m, the $m+1$ vectors $\mathbf{X}$, $\mathbf{Y}_i$ must be linearly dependent. Hence

$$a_0 \mathbf{X} + \sum_{i=1}^{m} a_i \mathbf{Y}_i = 0,$$

where $a_0 \neq 0$. For, $a_0 = 0$ will imply that $\mathbf{Y}_i$, $1 \leqslant i \leqslant m$, are linearly dependent vectors which, by hypothesis, they are not. It follows that

$$\mathbf{X} = -\sum_{i=1}^{m} (a_i/a_0)\, \mathbf{Y}_i$$

which means that $\mathbf{X}$ is a linear combination of $\mathbf{Y}_i$. Since $\mathbf{X}$ is any vector in V the theorem is proved.

It can be seen that the set of linearly independent vectors spanning a vector space is not unique. Consequently the basis of a vector space is also not unique. But once the basis is chosen every vector of the vector space has a unique linear combination expression in terms of the chosen basis.

4 Euclidean space

DEFINITION 6. *The inner product $\langle \mathbf{X}, \mathbf{Y} \rangle$ of any two vectors $\mathbf{X}$ and $\mathbf{Y}$ of V is a real number satisfying the following properties.*

(i) $\langle \mathbf{X}, \mathbf{Y} \rangle = \langle \mathbf{Y}, \mathbf{X} \rangle$;

(ii) $\langle \mathbf{X}+\mathbf{Z}, \mathbf{Y} \rangle = \langle \mathbf{X}, \mathbf{Y} \rangle + \langle \mathbf{Z}, \mathbf{Y} \rangle$, $\mathbf{Z} \in V$;

(iii) $\langle a\mathbf{X}, \mathbf{Y} \rangle = a \langle \mathbf{X}, \mathbf{Y} \rangle$, $a \in R$;

(iv) $\langle \mathbf{X}, \mathbf{X} \rangle > 0$ if $\mathbf{X} \neq 0$, $\langle \mathbf{X}, \mathbf{X} \rangle = 0$ if $\mathbf{X} = 0$.

Two nonzero vectors are said to be *orthogonal* if their inner product is zero.

DEFINITION 7. *A vector space with an inner product defined on it is called an Euclidean space.*

For vectors of the vector space R_n, the expression

$$X'Y=\sum_{i=1}^{n} x_i y_i$$

satisfies the definition of inner product. With this definition R_n becomes a Euclidean space. This Euclidean space, if of dimension n, shall be denoted by E_n. On account of its importance in the present work we give afresh the definition of E_n which may be understood without reference to general definitions of vectors and vector spaces given above.

DEFINITION 8. *Let R_n be a set of ordered n-tuples of real numbers. For every pair of n-tuples $\mathbf{X}, \mathbf{Y} \in R_n$, let*

(i) *Sum*: $\mathbf{X}+\mathbf{Y}=\mathbf{Y}+\mathbf{X}=(x_1+y_1, x_2+y_2, \ldots, x_n+y_n) \in R_n$;

(ii) *Product*: $a\mathbf{X}=(ax_1, ax_2, \ldots, ax_n) \in R_n, a \in R$;

(iii) *Inner product*: $\mathbf{X}'\mathbf{Y}=\mathbf{Y}'\mathbf{X}=x_1y_1+x_2y_2+\ldots+x_ny_n \in R$;

be defined. Then the n-tuples are called vectors and R_n is called a Euclidean space. Also let

(iv) *There be at least one set of n linearly independent vectors in R_n. Then R_n is a Euclidean space of dimension n which we shall denote as E_n.*

It should be noticed that the additional condition 'every set of $n+1$ vectors in R_n is linearly dependent' which was included in definition 5 has been dropped in (iv) above. The reason is that in this case it is implied and its explicit statement will be superfluous (see problem 10).

Example: The set of column vectors $[1\,0\,0]'$, $[0\,1\,0]'$, $[0\,0\,1]'$ is a basis of R_3. For, these vectors are linearly independent, and any vector $[x_1\ x_2\ x_3]'$ of R_3 can be expressed as a linear combination of these vectors as follows.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=x_1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}+x_2\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}+x_3\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

This basis is called the canonical or the natural basis of R_3. Another basis of R_3 is $[1\,0\,0]'$, $[1\,1\,0]'$, $[1\,1\,1]'$. For,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=a\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}+b\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}+c\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

where $c=x_3$, $b=x_2-x_3$, $a=x_1-x_2$.