MATHEMATICAL PROGRAMMING

Theory and Methods

S. M. Sinha

ELSEVIER

MATHEMATICAL PROGRAMMING
THEORY AND METHODS

This Page Intentionally Left Blank

MATHEMATICAL PROGRAMMING
THEORY AND METHODS

S. M. SINHA
Formerly Professor of Operational Research
University of Delhi, India

ELSEVIER
ELSEVIER
A division of
Reed Elsevier India Private Limited

Mathematical Programming: Theory and Methods
Sinha, SM

ELSEVIER
A division of
Reed Elsevier India Private Limited

First Edition 2006

ISBN 13: 978-81-312-0114-5 (PB)
ISBN 10: 81-312-0114-7 (PB)

ISBN 13: 978-81-312-0376-7 (HB)
ISBN 10: 81-312-0376-X (HB)

Published by Elsevier, a division of Reed Elsevier India Private Limited,
17-A/1, Main Ring Road, Lajpat Nagar-IV, New Delhi-110024, INDIA.

Laser typeset by Chitra Computers, Delhi

Printed and bound at Rajkamal Electric Press, Delhi.

DEDICATED TO

Mira Sinha, my wife, in memoriam

and to

Susmita, Madhumita and Amit

This Page Intentionally Left Blank

Preface

This book is a result of my teaching mathematical programming to graduate students of the University of Delhi for over thirty years. In preparing this book, a special care has been made to see that it is self-contained and that it is suitable both as a text and as a reference.

The book is divided in three parts. In Part 1, some mathematical topics have been reviewed to aid the reader in understanding the material discussed and to equip him with the ability to pursue further study. In Part 2, linear programming from its origin to the latest developments have been included and Part 3 deals with non-linear and dynamic programming including some special topics of recent developments. Several examples and exercises of varying difficulties have been provided to enable the reader to have a clear understanding of the concepts and methods discussed in the text.

I am indebted to Prof. George B. Dantzig, who first introduced me to stochastic programming, way back in 1961, when it was emerging as an important area of development. I am grateful to Prof. J. Medhi who originally suggested me to write a book on Mathematical Programming and to Prof. Roger J.B. Wets for his encouragement in taking up this project. Special thanks are due to my colleagues Prof. K. Sen, Prof. Manju Agarwal and Prof. J.M. Gupta for their various suggestions and help during the preparation of the manuscript. I am also thankful to my many students, particularly Dr. C.K. Jaggi and Dr. G.C. Tuteja for their numerous help. I am specially thankful to my son Dr. Amit K. Sinha for his help and continuous insistence for completing the book.

Finally, I must thank Mr. Sanjay Banerjee, Managing Director, Elsevier, South and South-East Asia, for getting the book published.

Delhi *S.M. Sinha*

This Page Intentionally Left Blank

Contents

Part – 2

LINEAR PROGRAMMING

Part – 3

NONLINEAR AND DYNAMIC PROGRAMMING

This Page Intentionally Left Blank

CHAPTER 1

Introduction

1.1. Background and Historical Sketch

Since the beginning of the history of mankind, man has been confronted with the problem of deciding a course of action that would be the best for him under the circumstances. This process of making optional judgement according to various criteria is known as the science of decision making. Unfortunately, there was no scientific method of solution for such an important class of problems until very recently. It is only in 1930s that a systematic approach to the decision problem started developing, mainly due to the advent of the 'New-Deal' in the United States and similar attempts in other parts of the world to curve the great economic depression prevailing throughout the world during this period: As a result during the 1940s, a new science began to emerge out.

About the same time, during World War II, the military management in the United Kingdom called upon a group of scientists from different disciplines to use their scientific knowledge for providing assistance to several strategic and tactical war problems. The encouraging results achieved by the British scientists soon motivated the military management of the U.S.A. to start on similar activities. The methodology applied by these scientists to achieve their objectives was named as Operations Research (O.R.) because they were dealing with research on military operations. Many of the theories of O.R. were in fact developed in direct response to practical requirements from military problems during the war. Following the end of the war, there have been increasing applications of operations research techniques in business and industry, commerce and managements and many other areas of our present day activities and from its impressive progress, it can aptly be said that one of the most remarkable development of the present century is the development of Operations Research techniques of which perhaps the most important is mathematical programming.

A mathematical programming problem is a special class of decision problem where we are concerned with the efficient use of limited resources to meet desired objectives. Mathematically the problem can be stated as,

$$\begin{aligned} &\text{Maximize} && f(X) \\ &\text{Subject to} && g_i(X) \le 0 \quad i = 1, 2, \ldots\ldots. m. \\ &&& X \ge 0 \end{aligned}$$

where $X = (x_1, x_2 .. x_n)^T \in R^n$, $f(X)$, $g_i(X)$, $i = 1, 2....m$ are real valued functions of X.

If the functions $f(X)$ and $g_i(X)$ are all linear, the problem is known as a linear programming problem otherwise it is said to be a nonlinear program.

Programming problems however, have long been of interest to economists. It can be traced back to the eighteenth century when economists began to describe economic systems in mathematical terms. In fact a crude example of a linear programming model can be found in the 'Tableau economique' of Qnesney, who attempted to interrelate the rolls of the landlord, the peasant and the artisan [355]. During the next 150 years there was however little in the way of exploitation of a linear type model, although it did appear as a part of Walrasian equilibrium model of an economy in 1874. The first major impetus to the construction of practical mathematical models to describe an economy came about in the 1930s when a group of Austrian and German economists started work on generalizations of the linear technology of Walras. This work stimulated the mathematician Von Neuman [488] to develop a linear model of expanding economy, which proved to be an outstanding one. A more practical approach was suggested by Leontief [303] who developed input–output models of an economy where one is concerned with determining the level of outputs that each of the various industries should produce to meet the specified demand for that product.

1.2. Linear Programming

Linear programming, as it is known today, arose during World War II out of the empirical programming needs of the Air Force and the possibility of generalizing the simple practical structure of the Leontief model to this end. It was George B. Dantzig, a member of the U.S. Air Force, who formulated the general linear programming problem and devised the simplex method of solution around 1947. [96]

As indicated earlier, the general linear programming problem has the following mathematical form: Find $x_1, x_2, .. x_n$ which satisfies the conditions

$$\begin{aligned} &a_{11} x_1 + a_{12} x_2 + .. + a_{in} x_n \le b_1 \\ &a_{21} x_1 + a_{22} x_2 + .. + a_{2n} x_n \le b_2 \\ &\vdots \\ &a_{m1} x_1 + a_{m2} x_2 + .. + a_{mn} x_n \le b_m \\ &\quad x_1 \ge 0, x_2 \ge 0 .. \quad x_n \ge 0. \end{aligned}$$

and maximize a linear function

$$c_1 x_1 + c_2 x_2 + .. + c_n x_n$$

where a_{ij}, b_i and c_j are assumed to be known constants.

It might seem that a solution to the above problem could be obtained by the method of calculus and in particular by the method of Lagrange multipliers. Unfortunately, the ordinary calculus can rarely be applied to programming problems since their optimal solutions lie on the boundary of the feasible region-infact at corner points of the boundary. Hence we need to develop a new procedure which will exploit the special feature of linear programming problems. Simplex method developed by Dantzig is such a method which is the most general and powerful enough to solve a large class of real life problems that can be formulated as linear programming problems. It is an iterative which yields an exact optimal solution of the problem in a finite number of steps.

Soon after the World War II, tremendous potentialities of linear programming in different fields of activity were realized and there followed throughout the business and scientific world, a rapidly expanding interest in the areas and methods of programming. The development of the powerful simplex method and the advent of high speed digital computers gave a large impetus to this rapid increase in interest and a surprisingly large class of decisions problems, particularly arising in industry and business, could now be formulated as linear programming problems. The bibliography composed by Riley and Gass [376] comprise hundreds of case study references. Much of the theoretical and computational development in linear programming is due to Dantzig, Gale, Kuhn and Tucker, Charnes and Cooper, and P. Wolfe.

1.3. Illustrative Examples

Before proceeding to discuss the theory of linear programming, we shall present a few examples to illustrate the applications of linear programming.

Example 1: Diet Problem

One simple example, which has become the classical illustration in linear programming, is the minimum cost diet problem. The problem is concerned in finding a diet that meets certain nutritional requirements, at minimum cost.

To be specific, let there be n different foods F_j, (j = 1, 2 .. n) from which a diet is to be selected and let b_i be the minimum daily requirement of nutrient N_i (i = 1, 2 .. m) such as proteins, calories, minerals, vitamins, etc., as suggested by the dietitian. Let x_j be the number of units of food F_j to be included in the diet, c_j be the cost per unit of food F_j and a_{ij} be the amount of nutrient N_i contained in one unit of food F_j.

The total amount of nutrient N_i, contained in such a food is

$$a_{i1} x_1 + a_{i2} x_2 + .. + a_{in} x_n$$

and this must be at least b_i and the cost of the food is

$$c_1 x_1 + c_2 x_2 + .. + c_n x_n,$$

which we would like to minimize.

Obviously, the number of units of different food will not be negative and we should have, $x_j \geq 0.$, $j = 1, 2 .. n$.

The problem therefore is

$$\text{Minimize } \sum_{j=1}^{n} c_j x_j$$

$$\text{Subject to } \sum_{j=1}^{n} a_{ij} x_j \geq b_i \qquad i = 1, 2,..m$$

$$x_j \geq 0, \qquad j = 1, 2,..n$$

which is a linear programming problem.

The diet problem is perhaps the earliest linear programming problem considered. It was George J. Stiegler [444] who first solved the problem in 1945 by trial and error method and later Dantzig and Laderman (unpublished) solved the same problem in 1947 by the simplex method. It is interesting to note Stiegler's comment that "the procedure is experimental because there does not appear to be any direct method of finding the minimum of a linear function subject to linear conditions." The development of the simplex method by Dantzig, only two years later is therefore a great and significant contribution in the field of linear programming.

It should be noted that the diet problem as formulated, could certainly not serve to feed human begins for any length of time as it does not provide any variations in the diet. It can, however, be profitably used to animal feeding and for selecting suitable diets for patients in large hospitals or for an army. It can also be applied to space research where however, instead of minimizing the cost we are required to minimize the weight/volume of food.

Example 2: Activity Analysis Problem

Suppose that a manufacturer produces a number of products through n different activities (plants) from fixed amount of m resources available to him. Also, suppose that there is unlimited demand of the products in the market.

Let a_{ij} be the number of units of the ith resources required to produce one unit of the jth product,

b_i be the available quantity of the ith resource, and

c_j be the profit from the sale of one unit of the jth product

We now wish to find a production schedule x_j ($j = 1, 2, .. n$) which will maximize the total income without exceeding the given supply of resources. The problem therefore is

$$\text{Minimize } \sum_{j=1}^{n} c_j x_j$$

$$\text{Subject to } \sum_{j=1}^{n} a_{ij}x_j \le b_i \qquad i = 1, 2,..m$$

$$x_j \le 0, \qquad j = 1, 2,..n$$

which again is a linear programming problem.

Example 3: Transportation Problem

Suppose that a manufacturer wishes to send a number of units of a product from several of his warehouses (origins) to a number of retail outlets (destinations, markets).

Let there be m origins and n destinations and let a_i be the total amount of the product available for shipment at the ith origin and b_j be the total demand of the product at the jth destination.

Let x_{ij} be the quantity of the product shipped from the ith origin to the jth destination and c_{ij} be the cost of shipping one unit from the ith origin to the jth destination.

The problem now is to determine a shipping schedule $x_{ij} \ge 0$, such that

(a) the demand b_j at the jth destination is satisfied
(b) the supply a_i at the ith origin is not exceeded
(c) the total shipping cost is minimum.

The problem therefore is, the linear programming problem:

$$\text{Minimize } \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}x_{ij}$$

$$\text{Subject to } \sum_{j=1}^{n} x_{ij} \le a_i, \qquad i = 1, 2, .. m$$

$$\sum_{i=1}^{m} x_{ij} \ge b_j, \qquad j = 1, 2, .. n$$

$$x_{ij} \ge 0, \qquad \text{for all } i, j$$

The transportation model has very wide applications and has received more attention than any other linear programming problem. Due to the simple structure of the problem, many special computational schemes have been designed for it. It is also to be noted that many linear programming problems which have nothing to do with transportation can be formulated as transportation problems.

The transportation problem has been discussed in detail in Chapter 19.

1.4. Graphical Solutions

To illustrate some basic features of linear programming, let us consider some problems involving only two variables which permit graphical solutions.

Example 1: Maximize $x_1 + 2x_2$

Subject to

$$x_1 + x_2 \le 5$$
$$-x_1 + 2x_2 \le 7$$
$$x_1 + 4x_2 \ge 4 \tag{1.2}$$
$$x_1 - x_2 \le 1.$$
$$x_1 \ge 0, x_2 \ge 0. \tag{1.3}$$

The nonnegativity conditions imply that the points should lie in the first quadrant of the (x_1, x_2) plane. We then consider the inequalities as equations and draw straight lines. On each of these lines drawn, we indicate by arrows the area any point in which will satisfy that particular inequality (condition). The intersection of all these half planes is the polygon ABCDE (Figure 1.1). Any point within this polygon or on its boundary satisfies all the conditions. An infinite number of such points exist. Our problem is to find a point (or points) within this polygon or on its boundary that maximizes the linear form $x_i + 2x_2$.

Consider the parallel straight lines

$$x_1 + 2x_2 = z$$

where z is a parameter.

For z = 4 or 6, the line cuts through the polygon ABCDE, implying that the value of z can be taken higher than z = 6. For z = 9, the straight line passes only through the corner point D of the polygon indicating that for any value of z higher than 9, the straight line passes beyond the polygon. It does not intersect the polygon at all and hence there is no point satisfying the conditions and yielding a value of z > 9.

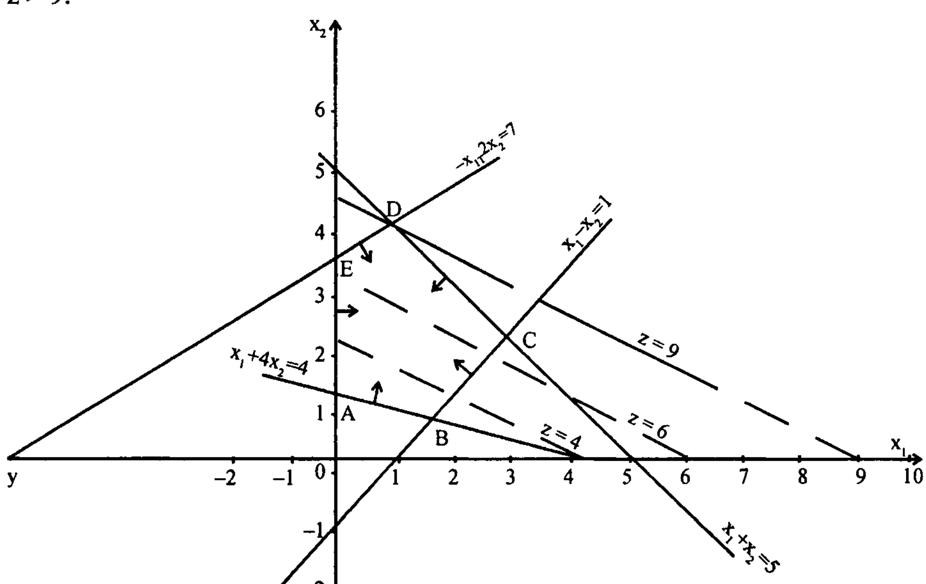


Figure 1.1

Hence the maximum value of the linear form $x + 2x_2$ subject to the conditions (1.2), (1.3) is 9 and attained at the point D whose coordinate are $x_1 = 1$, $x_2 = 4$ which can be easily obtained from the solution of the system

$$x_1 + x_2 = 5$$
$$-x_1 + 2x_2 = 7$$

Example 2. In example 1, the linear form attained its maximum at a unique point D. It is however not necessary that a linear programming problem will always have a unique solution.

If we maximize the linear form $x_1 + x_2$ subject to the same conditions as in example 1 we note that the linear form attains its maximum 5 at corner points C and D and also at every point on the line CD.

Example 3. Consider the problem

$$\begin{array}{ll} \text{Minimize} & 2x_1 + x_1 \\ \text{Subject to} & x_1 + x_2 \leq 4 \\ & 3x_1 + x_2 \geq 5 \\ & x_1 + 3x_2 \geq 3 \\ & x_1 \qquad \leq 3 \\ & \qquad x_2 \leq 2.5, \qquad x_1, x_2 \geq 0 \end{array}$$

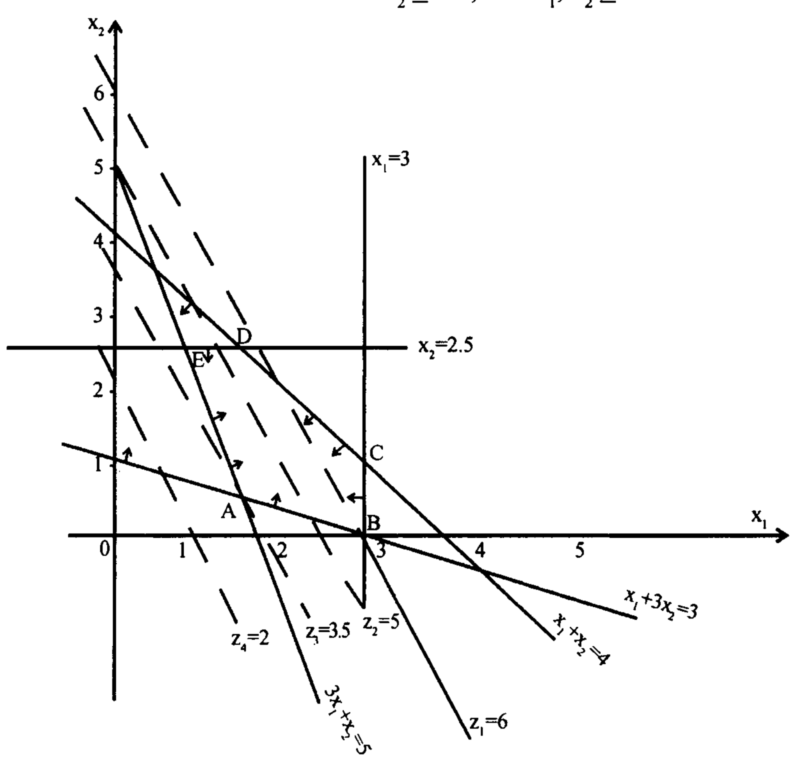


Figure 1.2

It is clear from the lines representing the linear form $2x_1 + x_2 = z$ for $z = 6$, $z = 5$ in Figure 1.2, that if we move parallel to these lines in the direction of decreasing z, it passes through the corner point A of the polygon ABCDE for $z = z_3$ and for $z < z_3$, it does not intersect the polygon at all. The minimum value of the linear form therefore is z_3 and is attained at A, which is the point of intersection of the lines $3x_1 + x_2 = 5$ and $x_1 + 3x_2 = 3$. The solution of the problem is then $x_1 = 3/2$, $x_2 = 1/2$ and min $z = 3.5$.

Example 4: There may be some linear programming problems for which there do not exist any point satisfying all the conditions or there may be problems where the solution is unbounded. We do not, however, expect any linear programming problem representing some practical situation which has no solution or an unbounded solution. But for theoretical consideration, we now consider the following problem.

$$\begin{aligned} \text{Maximize} \quad & 2x_1 + 3x_2 \\ \text{Subject to} \quad & x_1 + x_2 \geq 4 \\ & x_1 - x_2 \leq 2 \\ & x_1 \geq 0,\ x_2 \geq 0. \end{aligned}$$

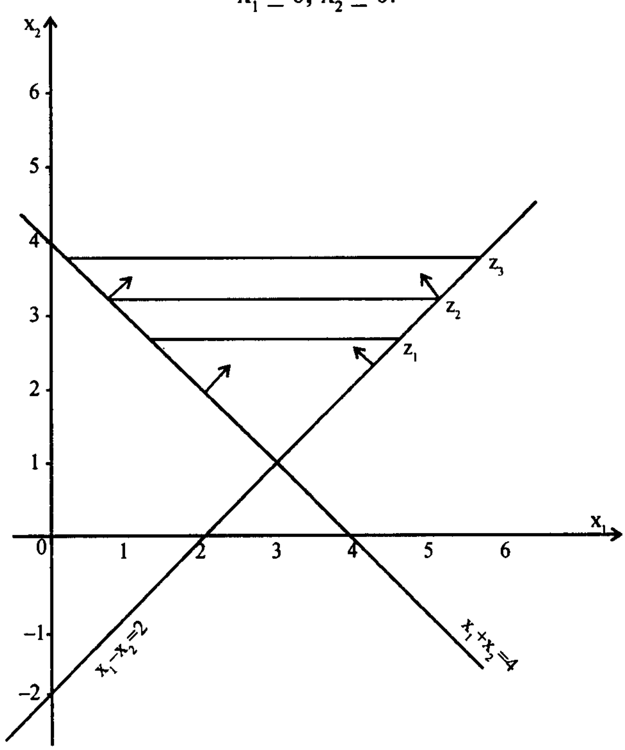


Figure 1.3

It is clear from Figure 1.3, that no matter how far the line $2x_1 + 3x_2 = z$ is moved parallel to itself in the direction of increasing z, the line will always have points satisfying the conditions. Hence z can be made arbitrarily large and the problem therefore has an unbounded solution.

1.5. Nonlinear Programming

Interest in nonlinear programming developed almost simultaneously with the growing interest in linear programming. It was soon recognized that many a practical problem cannot be represented by linear programming model. Therefore, attempts were made to develop methods of solutions for more general mathematical programs and many significant advances have been made. In 1951, Kuhn and Tucker [291] gave necessary and sufficient conditions for optimal solutions to programming problems which laid the foundations for a good deal of later work in nonlinear programming. Since then, many authors have developed methods of solutions for nonlinear programs of various nature and a large number of papers have appeared in the literature.

This Page Intentionally Left Blank

PART – 1
MATHEMATICAL FOUNDATIONS

In Part 1 of the book, we have reviewed some mathematical concepts that are needed to study mathematical programming and in particular the topics covered in the remainder of this book. For more details, see Bartle [27], Berge [52], Berge and Ghouila-Houri [53], Buck [62], Fenchel [157], Finkbeiner [161], Flemming [162], Gale [183], Halmos [218], Rudin [386] and Simmons [419].

This Page Intentionally Left Blank

Chapter 2

Basic Theory of Sets and Functions

2.1. Sets

A set A is a collection of objects of any kind which are called the elements or points of A. In general sets are denoted by capital Latin or Greek letters such as A, B, X or Ω, Λ, Γ.

The set of all real numbers is denoted by R.

If a is an element of the set A, we write $a \in A$ and we write $a \notin A$, if a is not an element of A.

A set is sometimes defined by listing its elements between the curly brackets. For example, the set consisting of the elements a, b, c may be written as $A = \{a, b, c\}$. But more often sets are defined by one or more properties that characterize its elements and we write $A = \{a \mid a \text{ satisfies property } P\}$. For example, the set of all nonnegative real numbers can be written as $A = \{x \mid x \in R, x \geq 0\}$.

The set that contains no element is called the empty set and is denoted by ϕ. A set is finite if the number of its elements is finite and infinite otherwise.

If A and B are two sets and all elements of A are elements of B, we write $A \subset B$, that is, A is contained in B or that B contains A and write $B \supset A$. A is then called a subset of B.

If $A \subset B$ and $B \subset A$, we say that A is equal to B and $A = B$.

Operations on Sets

If A and B are two sets, their union denoted by $A \cup B$ is defined to be the set of elements which belong to either A or B. That is,

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}.$$

The intersection of two sets A and B denoted by $A \cap B$ is defined to be the set of elements which belong to both A and B. That is,

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}.$$

If $A \cap B = \phi$, the sets A and B are said to be disjoint.

The difference of two sets A and B denoted by $A - B$ is the set of elements that belong to A but not to B, i.e.

$$A - B = \{x \mid x \in A, x \notin B\}.$$

The set of all elements do not belong to the set A is called the complement of A and is denoted by A^c.

The set operations of union and intersection are (i) commutative (ii) associative and (iii) distributive:

(i) $A \cup B = B \cup A$; $\quad A \cap B = B \cap A$.

(ii) $(A \cup B) \cup C = A \cup (B \cup C)$; $(A \cap B) \cap C = A \cap (B \cap C)$.

(iii) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Moreover, the operation of complementation has the following properties:

$$A \subset B \quad \text{implies} \quad A^c \supset B^c.$$

$$A \cap A^c = \phi, \qquad (A^c)^c = A$$

$$A - B = A \cap B^c, (A \cup B)^c = A^c \cap B^c,$$

$$(A \cap B)^c = A^c \cup B^c.$$

The operations of union and intersection for two sets extend easily to any finite and infinite number of sets. Let I be a finite or infinite set of integers. Then $A = (A_i \mid {}_i \in I)$ is called a family of the sets A_i.

The union of the A_1 is defined to be the set of elements which belong to at least one A_i and is denoted by $\bigcup\limits_{i \in I} A_i$.

The intersection of the A_i is defined to be the set of elements which belong to all the sets A_i and is denoted by $\bigcap\limits_{i \in I} A_i$.

The product of two sets A and B is defined to be the set of ordered pairs (x, y) of which $x \in A$ and $y \in B$ and is denoted by $A \times B$. That is,

$$A \times B = \{(x, y) \mid x \in A, y \in B\}.$$

The product of sets A_i, $i = 1, 2, .. n$ is defined to be the set of ordered n-triples $(x_1, x_2, .. x_n)$ such that $x_i \in A_i$, $i = 1, 2, .. n$. and is denoted by $A_1 \times A_2 .. \times A_n$. Thus,

$$A_1 \times A_2 .. \times A_n = \{(x_1, x_2, ... x_n \mid x_1 \in A_1, x_2 \in A_2, ... x_n \in A_n\}.$$

Closed and Open Intervals

Let a and b be two real numbers. The set of real numbers x satisfying $a \le x \le b$ is a closed interval and is denoted by [a, b]. The set of real numbers x satisfying $a < x < b$ is called an open interval and is denoted by (a, b).

Thus, $[a, b] = \{x \mid a \le x \le b\}$ is a closed interval

and $(a, b) = \{x \mid a < x < b\}$ an open interval

Similarly, the sets

$$[a, b) = \{x \mid a < x < b\}$$

$$\text{and} \quad (a, b] = \{x \mid a < x \le b\}$$

are known as left half-closed and right half-closed intervals, respectively.

Lower and Upper Bounds

Let S be a nonempty set of real numbers. If there is a number α such that $x \geq \alpha$ for all $x \in S$, then S is said to be bounded below and α is called a lover bound of S.

Similarly, if there is a number β such that $x \leq \beta$, for all $x \in S$, then S is said to be bounded above and β is called an upper bound of S.

If S is bounded above and below, then S is said to be bounded.

Greatest Lower Bound and Least Upper Bound

Let S b a nonempty set of real numbers. A lower bound $\bar{\alpha}$ is the greatest lower bound or the infimum of S (i.e. inf $\{x \mid x \in S\}$) if no number greater than $\bar{\alpha}$ is a lower bound of S.

Similarly, an upper bound $\bar{\beta}$ is the least upper bound or supremum of S (i.e. sup $\{x \mid x \in S\}$), if no number smaller than $\bar{\beta}$ is an upper bound of S.

Any nonempty set of real numbers which has a lower (upper) bound has a unique greatest lower (least upper) bound. This is one of the axioms of the real number systems.

Thus, if the set has no infimum, it has no lower bound and we then say that S is unbounded from below and write inf $S = -\infty$

Similarly, if S has no supremum, it has no upper bound and we then say that S is unbounded from above and write Sup $S = +\infty$

2.2. Vectors

An n-vector X is an ordered n-tuple of real number $x_1, x_2, .. x_n$. A vector is denoted by a capital letter. It is written as a column, i.e.

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

and is called a column vector or simply a vector. It is also represented in the form of a row vector

$$X^T = [x_1, x_2, ... x_n]$$

where T denotes transposition.

The number x_j is called the j th component or element of the vector X.

Special Vectors

Null vector: The null vector (zero vector) is a vector, all of whole components are zero and is denoted by **0**.

$$\mathbf{0} = [0, 0, \dots 0]^T.$$

Sum vector: The sum vector denoted by e is a vector for which all components are 1.

$$e = [1, 1, \dots 1]^T.$$

Unit Vectors: The ith unit vector, e_i is a vector whose ith component is 1 and all others are zeros. Thus

$$e_1 = [1, 0, 0, \dots 0]^T, e_2 = [0, 1, 0, \dots 0]^T, \dots e_n = [0, 0, 0, \dots 1]^T.$$

Addition of Vectors: The sum of two n-vectors X and Y is written as the n-vector X + Y, whose jth components is $x_j + y_j$.

$$X + Y = [x_1 + y_1, x_2 + y_2, \dots x_n + y_n]^T.$$

Multiplication by a scalar: The product of a vector X and a scalar α, denoted by αX is obtained by multiplying each element of X by α

$$\alpha X = [\alpha x_1, \alpha x_2, \dots \alpha x_n]^T.$$

Scalar Product: The scalar product or inner product of two n-vectors X and Y is defined by $X^T Y = \sum_{i=1}^{n} x_i y_i$. The two vectors are said to be orthogonal if $X^T Y = 0$.

The Norm or Length of a Vector: The norm (length) of an n-vector X, denoted by $||X||$ is defined by $||X|| = (X^T X)^{1/2}$

Distance between vectors: The distance between two n-vectors X, Y is defined by $d(X, Y) = ||X - Y||$.

The set of n-vectors of real numbers, denoted by R^n, with which we are mainly concerned in our study, forms the n-dimensional Euclidean Space (see chapter 3).

2.3. Topological Properties of R^n

We now consider some topological properties of R^n.

A topology in a set Ω is the family of open sets (defined below) in Ω if it is closed under operation of arbitrary unions and finite intersections and contains ϕ and Ω. Thus the family of open sets in R^n by its properties given below form a topology in R^n.

Neighbourhoods: Given a point $X_0 \in R^n$ and real number $\in > 0$, the set

$$N_\in (X_0) = \{X \mid X \in R^n \ ||X - X_0|| < \in\}$$

is called an $\in$ neighbourhood of X_0.

$N_\in (X_0)$ of often called an open ball $B_\in(X_0)$ (or simply $B(X_0)$ with centre X_0 and radius $\in$.

Interior Points and Open Sets

A point X is said to be an interior point of the set $S \subset R^n$, if there exists an $\in > 0$ such that $N_\in(X) \subset S$.

The set of all interior points of S is called the interior of S and is denoted by int. S. Obviously int $S \subset S$.

S is called open if S = int S, that is, if ever point of S is an interior point.

Points of Closure and Closed Sets

A point X is said to be a point of closure of the set $S \subset R^n$, if for each $\in > 0$

$$N_\in(X) \cap S \neq \phi$$

The set of points of closure of S, denoted by $\bar{S}$, is called the closure of S. Clearly $\bar{S} \supset S$.

The set S is said to be closed if $S = \bar{S}$, that is, if every point of closure of S is in S.

It should be noted that a point of closure of a set S may or may not be a point of the set. For example, if S is the set {1, ½, ¼,.....}, the point O is a point of closure of S but does not belong to S. However, every point in the set S is a point of closure of S.

Relatively Open (closed) Sets

Let S_1 and S_2 be two sets such that $S_1 \subset S_2 \subset R^n$. The S_1 is said to be open (closed) relative to S_2 if there is an open (closed) set Λ in R^n such that $S_1 = S_2 \cap \Lambda$.

The following properties of open and closed sets in R^n can be easily verified.

A: (i) Every Union of open sets is open.
(ii) Every finite intersection of open sets is open.
(iii) The empty set ϕ and R^n are open.

B: The complement of an open set is closed and the complement of a closed set is open.

C: (i) Every intersection of closed sets is closed.
(ii) Every finite union of closed sets is closed.
(iii) The empty set ϕ and R^n are closed.

Boundary Points

A point $X \in S \subset R^n$ is a boundary point of S if it is not an interior point of S, that is, if for each $\in > 0$, $N_\in$ (X) contains at least one point in S and one point not in S.

The set of all boundary points of S is called the boundary of S.

Bounded Sets

A set $S \subset R^n$ is said to be bounded if there exists a real number $\alpha > 0$ such that

$$\|X\| < \alpha, \qquad \text{for each } X \in S.$$

Compact Sets

A set $S \subset R^n$ is said to be compact if it is closed and bounded.

If S_1 and S_2 are sets in R^n which are respectively compact and closed, then their sum

$$S = S_1 + S_2 = \{x + y \mid x \in S_1, y \in S_2\}$$

is closed.

2.4. Sequences and Subsequences

A sequence in a set S is a function f from the set I of all positive integers into the set S. If $f(n) = x_n \in S$ for $n \in I$, we denote the sequence by $\{x_n\}$ or by x_1, x_2, x_3 ..... The elements x_1, x_2, x_3 .... however need not be distinct.

For any sequence $\{n_k\}$ of positive integers, such that $n_1 < n_2 < n_3$ ....., the sequence $\{x_{n_k}\}$ is called s subsequence of $\{x_n\}$.

If S is a set of real numbers, then the $\{x_n\}$ is a sequence of real numbers and if $S = R^n$, it is a sequence of vectors.

Limit Point

Let $\{X_n\}$ be a sequence in R^n. A point $\bar{X} \in R^n$ is said to be a limit point of the sequence if for any given $\in > 0$ there is a positive integer N such that

$$\|X_n - \bar{X}\| < \in, \qquad \text{for some } n \geq N.$$

A limit point is also called an accumulation point or a cluster point.

Limit of a Sequence

Let X_1, X_2, X_3, ... be a sequence of points in R^n. A point $\bar{X} \in R^n$ is said to be the limit of the sequence, if for any given $\in > 0$, there is a positive integer N such that

$$\|X_n - \bar{X}\| < \in, \qquad \text{for all } n \geq N.$$

and we say that $\{X_n\}$ converges to $\bar{X}$ or that $\bar{X}$ is limit of $\{X_n\}$ and write

$$X_n \to \bar{X}, \quad \text{as } n \to \infty, \text{ or}$$

$$\lim_{n\to\infty} X_n = \bar{X}$$

It can be shown that

(i) Every sequence $\{X_n\}$ in a compact set S in R^n has a limit point in S (Bolzano–Weierstrass) and then there is a convergent subsequence with a limit in S.

(ii) The limit of a convergent sequence is unique.

Cauchy Sequence

A sequence $\{X_n\}$ in R^n is said to be a Cauchy sequence if for any given $\in > 0$, there is a positive integer N such that for all m, $n \geq N$,

$$\|X_m - X_n\| < \in.$$

A sequence in R^n converges if and only if it is a Cauchy sequence.

2.5. Mappings and Functions

Mappings

Let X and Y be two sets. A mapping Γ from X into Y is a correspondence which associates with every element x of X a subset $\Gamma(x)$ of Y. The set $\Gamma(x)$ is called the image of x under the mapping Γ.

The set $X^* = \{x \mid x \in X, \Gamma(x) \neq \phi\}$ is called the domain (or set of definition) of Γ, and

$$Y^* = \bigcup_{x \in X} \Gamma(x)$$

is called the range (or set of values) of Γ.

Γ is also said to be defined on X^* and that Γ is a mapping ox X onto Y^*.

If B is a subset of Y, the set

$$\Gamma^{-1}(B) = \{x \in X \mid \Gamma(x) \in B\} \subset X$$

is called the inverse image B under the mapping Γ.

A mapping is called one-to-one if x_i, $x_2 \in X$ and $x_1 \neq x_2$ implies that $\Gamma(x_1) \neq \Gamma(x_2)$, and for every $y \in Y$, there is an $x \in X$ such that $y = \Gamma(x)$.

Functions

If the mapping Γ from a set X into a set Y is such that the image set $\Gamma(x)$ always consists of a single element, Γ is called a single valued mapping or a single valued function and is denoted by f.

If Y is an m-dimensional Euclidean space R^m, Γ is called a vector-valued function (or an m-vector function) f i.e. f(x) is a vector whose m components $f_1(x)$, $f_2(x)$,..$f_m(x)$ are real numbers.

If $Y = R^1$, the function is called a numerical function and if moreover $X = R^n$, then f is the real single valued function of n variables, often written as $f : R^n \rightarrow R^1$.

In our subsequent discussions, we will mainly be concerned with real single valued functions defined on R^n or on a subset of R^n.

2.6. Continuous Functions

Let S be a nonempty set in R^n. A function $f : S \to R^1$ is said to be continuous at $\bar{X} \in S$, if for any given $\epsilon > 0$, there exists a $\delta > 0$ such that

$$X \in S\ ,\ \| X - \bar{X} \| < \delta \text{ imply that } |f(x) - f(\bar{X})| < \epsilon$$

or equivalently if for each sequence $\{X_n\}$ in S,

$$\lim_{n\to\infty} f(X_n) = f(\bar{X}).$$

A function f is said to be continuous on S, if it is continuous at each point of S.

A vector valued function is said to be continuous at $\bar{X}$ if each of its components is continuous at $\bar{X}$.

Lower and Upper Semicontinuous Functions

Let $S \subset R^n$. A function $f : S \to R^1$ is said to be lower semicontinuous at $\bar{X} \in S$, if for each $\epsilon > 0$ there exists a $\delta > 0$ such that

$$X \in S \text{ and } \|X - \bar{X}\| < \delta \text{ imply that } f(X) - f(\bar{X}) > -\epsilon.$$

Similarly, a function $f : S \to R^1$ is said to be upper semicontinuous at $\bar{X} \in S$ if for each $\epsilon > 0$ there exists a $\delta > 0$ such that

$$X \in S \text{ and } \|X - \bar{X}\| < \delta \text{ imply that } f(X) - f(\bar{X}) < \epsilon.$$

f is said to be lower (upper) semicontinuous on S if it is lower (upper) semicontinuous at each point of S.

Note that f is continuous at $\bar{X} \in S$ if and only if it is both lower and upper semicontinuous at that point.

2.7. Infimum and Supremum of Functions

Bounded Functions

Let $S \subset R^n$. A function $f : S \to R^1$ is said to be bounded from below on S if there exists a number α such that

$$f(X) \geq \alpha, \qquad \text{for all } X \in S.$$

The number α is a lower bound of f on S.

f is said to be bounded from above on S if there exists a number β such that

$$f(X) \leq \beta, \qquad \text{for all } X \in S.$$

The number β is an upper bound of f on S.

The function f is said to be bounded on S if it is bounded from below and from above.

Infimum of Functions: Let $S \subset R^n$ and $f : S \to R^1$. If there exists a number $\bar{\alpha}$ such that

$$f(X) \geq \bar{\alpha}, \qquad \text{for all } X \in S$$

and for every $\in > 0$, there exists an $X \in S$ such that

$$f(X) < \bar{\alpha} + \in,$$

then $\bar{\alpha}$ is called the infimum of f on S and is denoted by $\bar{\alpha} = \inf \{f(x) \mid X \in S\} = \inf_{X \in S} f(x)$.

Supremum of Functions: Let $S \subset R^n$ and $f : S \to R^1$. If there exists a number $\bar{\beta}$ such that

$f(X) \leq \bar{\beta}$, for all $X \in S$ and for every $\in > 0$,

there exists an $X \in S$ such that

$$f(X) > \bar{\beta} - \in,$$

then $\bar{\beta}$ is called the supremum of f on S and is denoted by $\bar{\beta} = \sup \{f(X) \mid X \in S\} = \sup_{X \in S} f(X)$.

If R^1 is the complete set of real numbers, that is, if we admit the points $\pm \infty$, then every real single valued function has a supremum and infimum on the set S on which it is defined.

By convention, if $S = \phi$

$$\inf_{X \in S} f(X) = +\infty$$

$$\sup_{X \in S} f(X) = -\infty$$

and if $S \neq \phi$ and f is not bounded on S

$$\inf_{X \in S} f(X) = -\infty$$

$$\sup_{X \in S} f(X) = +\infty$$

2.8. Minima and Maxima of Functions

Let f be a real single valued function defined on the set S. If there is a point $X_0 \in S$ such that $f(X_0) \leq f(X)$ for all $X \in S$ then $f(X_0)$ is called the minimum of f on S and is written as

$$f(X_0) = \min_{X \in S} f(X)$$

and X_0 is then called a minimum point of f on S. Thus if the minimum of f exists, f attains the minimum at the infimum of f on S.

Similarly, if there exists a point $X^* \in S$ such that $f(X^*) \geq f(X)$ for all $X \in S$, then $f(X^*)$ is called the maximum of f on S and is written as

$$f(X^*) = \max_{X \in S} f(X)$$

X^* is called a maximum point of f on S. Clearly, if the maximum of f exists, it is attained at the supremum of f on S.

It should be noted that not every function $f : S \to R^1$ has a minimum or maximum.

Examples.

For $x \in R^1$,

(i) $\{f(x) = x \mid 0 < x < 1\}$ has no minimum or maximum.

(ii) $\{f(x) = x \mid 0 < x < 1\}$ has no minimum but achieves maximum at x = 1.

(iii) $\{f(x) = e^{-x} \mid x \in R^1\}$ has no minimum.

However, if the minimum (maximum) exists, it must be finite.

The following theorem gives sufficient conditions for existence of a minimum or a maximum of a function $f : S \to R^1$.

Theorem 2.1. Let S be a nonempty compact set in R^n and $f : S \to R^1$. If f is lower (upper) semicontinuous then f achieves a minimum (maximum) over S.

Since a continuous function is both lower and upper semicontinuous, it achieves both a minimum and a maximum over any compact set.

2.9. Differentiable Functions

Let $S \subset R^n$ be an open set and $f : S \to R^1$. Then f is said to be differentiable at $\bar{X} \in S$ if for all $X \in R^n$ such that $\bar{X} + X \in S$, we have

$$f(\bar{X}+X) = f(\bar{X}) + \theta(\bar{X})^T X + \beta(\bar{X}, X)\, \|X\|$$

where $\theta(\bar{X})$ is a bounded vector and $\beta(\bar{X}, X)$ is a numerical function which tends to 0 whenever X tends to 0.

f is said to be differentiable on S if it is differentiable at each point in S.

Partial Derivatives: Let $S \subset R^n$ be an open set and $f : S \to R^1$. Then f is said to have a partial derivative at $X \in S$ with respect to x_j, the jth component of X if

$$\lim_{h \to 0} \frac{f(X + he_j) - f(X)}{h}$$

where e_j is the jth unit vector, exists. This limit is called the partial derivative (or first partial derivative) with respect to x_j at X and is denoted by $\partial f / \partial x_j$.

The n-vector whose components are the n partial derivatives of f with respect to x_j, j = 1, 2,..n is called the gradient of f at X and is denoted by $\nabla f(X)$, that is,

$$\nabla f(X) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots \frac{\partial f}{\partial x_n}\right)^T.$$

The following theorem gives the relationship between continuity and differentiability of functions.

Theorem 2.2. Let S be an open set in R^n and $f : S \to R$.

(i) If f is differentiable at $\bar{X} \in S$, then f is continuous at $\bar{X}$ and $\nabla f(\bar{X})$ exists (but not conversely) and for $\bar{X} + X \in S$,

$$f(\bar{X} + X) = f(\bar{X}) + \nabla f(\bar{X})\, X + \beta(\bar{X}, X)\, \|X\|$$

where $\beta(\bar{X}, X) \to 0$ as $X \to 0$.

(ii) If f has continuous partial derivatives at $\bar{X}$ with respect to all the variables, then f is Differentiable at $\bar{X}$.

Whenever we say that f is differentiable (or that it has partial derivatives) on some set Λ (open or not), it will be tacitly understood that there is an open set θ containing Λ and f is defined on θ.

Vector-valued Functions

An m-dimensional vector function f defined on an open set S in R^n is said to be differentiable at $\bar{X} \in S$ (or on S) if each of its components $f_1, f_2, .. f_n$ is differentiable at $\bar{X}$ (or on S).

f is said to have partial derivatives at $\bar{X} \in S$ with respect to $x_1, x_2, .. x_n$ if each of its components has partial derivatives at $\bar{X}$ with respect to $x_1, x_2, .. x_n$.

The m by n matrix of the partial derivatives

$$\nabla f(X) = \begin{bmatrix} \frac{\partial f_1(\bar{X})}{\partial x_1} & \frac{\partial f_1(\bar{X})}{\partial x_2}, \ldots & \frac{\partial f_1(\bar{X})}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_m(\bar{X})}{\partial x_1} & \frac{\partial f_m(\bar{X})}{\partial x_2}, \ldots & \frac{\partial f_m(\bar{X})}{\partial x_n} \end{bmatrix}$$

is called the Jacobian (or Jacobian matrix) of f at $\bar{X}$

If an m-dimensional vector function f defined on an open set S in R^n and a function $g : R^m \to R^1$ are differentiable at $\bar{X} \in S$ and at $\bar{Y} = f(\bar{X})$ respectively, then the composite function $\phi : S \to R^1$ given by

$$\phi(X) = g[f(X)]$$

is differentiable at $\bar{X}$.

Furthermore, $\nabla\phi(\bar{X}) = \nabla g(\bar{Y})^T \nabla f(\bar{X})$

This is the well known chain rule of the differential calculus.

Twice-differentiable Functions

Let S be an open set in R^n and $f : S \to R^1$. Let $X \in S$.

Since each of the partial derivatives $\partial f(X)/\partial x_j$ with respect to x_j, j = 1, 2,..n is a function of X, we can obtain partial derivatives of each of these functions if they exist. These are defined as second partial derivatives of f and are denoted by $\partial^2 f(X)/\partial x_i \partial x_j$, i, j = 1, 2,..n. If the second partial derivatives exist and are continuous, then $= \dfrac{\partial^2 f(x)}{\partial x_i \partial x_j} = \dfrac{\partial^2 f(x)}{\partial x_j \partial x_i}$ for all i, j. The n x n symmetric matrix

$$H(X) = \left[\frac{\partial^2 f(X)}{\partial x_i \partial x_j}\right], \quad i, j = 1, 2, \ldots n$$

is called the Hessian matrix of f at X.

f is said to be twice differentiable at $\bar{X} \in S$ if in addition to the gradient vector there exist the Hessian matrix $H(\bar{X})$ of f at $\bar{X}$ and a numerical function $\beta(\bar{X}, X)$, such that

$$f(X) = f(\bar{X}) + \nabla f(\bar{X})^T (X - \bar{X}) + \tfrac{1}{2}(X - \bar{X})^T H(\bar{X})(X - \bar{X})$$

$$+ \|X - \bar{X}\|^2 \beta(\bar{X}, X). \text{ fpr eacj } X \in S.$$

where $\beta(\bar{X}, X) \to 0$ as $X \to \bar{X}$

A vector-valued function is twice differentiable if each of its components is twice differentiable.

Mean Value Theorem

Let S be an open convex set in R^n (See chapter 8) and $f : S \to R^1$ be differentiable. Then for every X_1 $X_2 \in S$.

$$f(X_2) = f(X_1) + \nabla f(X)^T (X_2 - X_1)$$

where $X = \lambda X_1 + (1 - \lambda) X_2$ for some $\lambda \in (0, 1)$.

Taylor's Theorem. Let S be an open convex set in Rn and $f : S \to R^1$ be twice differentiable. Then for every $X_1, X_2 \in S$, the second-order form of Taylor's theorem can be stated as

$$f(X_2) = f(X_1) + \nabla f(X_1)^T (X_2 - X_1) + \tfrac{1}{2}(X_2 - X_1)^T H(X)(X_2 - X_1)$$

where $X = \lambda X_1 + (1 - \lambda) X_2$ for some $\lambda \in (0, 1)$ and H(X) is the Hessian of f at X.

Implicit Function Theorem

Consider now the problem of solving a system of n equations (nonlinear) in n + m variables. We can see that if it is possible to express n of the variables as the function of the remaining m variables, the system of equations can be solved for the n variables as implicit functions of the remaining ones.

The following theorem gives the conditions for existence of such functions.

Theorem 2.3. Let f be an n-vector function defined on an open set $\Lambda \subset R^n \times R^m$ and let $(\bar{X}, \bar{Y}) \in \Lambda$.

Suppose that the following conditions hold:

(i) $f(\bar{X}, \bar{Y}) = 0$

(ii) f has continuous first partial derivatives at $(\bar{X}, \bar{Y})$ with respect to the components of X

(iii) $\nabla_x f(\bar{X}, \bar{Y})$ is nonsingular.

Then there exists a unique n-vector function g defined on an open set $U \subset R^m$, $\bar{Y} \in U$ such that

(a) $\bar{X} = g(\bar{Y})$

(b) $f[g(Y), Y] = 0$ for $Y \in U$.

(c) g has continuous first partial derivatives on U.

(d) $\nabla_x f(X, Y)$ is nonsingular in an open ball $B_\varepsilon(\bar{X}, \bar{Y})$

CHAPTER 3

Vector Spaces

3.1. Fields

A nonempty set F of scalars (real or complex numbers) on which the operations of addition and multiplication are defined, so that to every pair of scalars $\alpha, \beta \in F$ there corresponds scalars $\alpha + \beta$ and $\alpha\beta$ in F is called a field if the following conditions are satisfied for $\alpha, \beta, \gamma \in F$

(1) $\alpha + \beta = \beta + \alpha$, $\alpha\beta = \beta\alpha$. (commutativity)

(2) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$, $\alpha(\beta\gamma) = (\alpha\beta)\gamma$. (associativity)

(3) $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ (distributivity)

(4) For each scalar α, there exists a unique scalar 0 such that $\alpha + 0 = \alpha$. (additive identity)

(5) For each scalar α, there is a unique scalar $-\alpha$ such that $\alpha + (-\alpha) = 0$. (additive inverse)

(6) There exists a unique nonzero scalar 1 such that $\alpha.1 = \alpha$ for every scalar α (multiplicative identity)

(7) For every nonzero scalar α, there exists a unique scalar α^{-1} such that $\alpha\alpha^{-1} = 1$ (multiplicative inverse).

Thus, for example, the set R of all real numbers is a field and the same is true of the set C of all complex numbers and the set Q of all rational numbers. It can be seen that the set Z of all integers is not a field.

3.2. Vector Spaces

A vector space V over the field F is a set of vectors on which the operations of addition and multiplication by a scalar are defined, that is, for every pair of vectors $X, Y \in V$, there corresponds a vector $X + Y \in V$ and also $\alpha X \in V$, where α is a scalar such that the following conditions are satisfied:

For $X, Y, Z \in V$ and α, β scalars

1. $X + Y = Y + X$
2. $X + (Y + Z) = (X + Y) + Z$
3. For every vector X in V, there exists a unique vector 0 in V such that $X + 0 = X$

4. For every vector X in V, there is a unique vector – X such that X + (– X) = 0.
5. $\alpha(\beta X) = (\alpha\beta)X$.
6. 1 X = X, for every vector X.
7. $\alpha(X + Y) = \alpha X + \alpha Y$
8. $(\alpha + \beta)X = \alpha X + \beta X$.

If F is the field R of real numbers, V is called a real vector space and if F is the field C of complex numbers, V is a complex vector space.

It is easy to verify that the set R^n of all n-tuples of real numbers is a real vector space. It is also called n-dimensional real coordinate space.

3.3. Subspaces

A nonempty subset U of a vector space V is called a subspace of V or a linear manifold if it is closed under the operation of addition and scalar multiplication, that is,

$$\text{if } X, Y \in U, \text{ then } X + Y \in U$$

$$\text{if } X \in U, \alpha \in R, \text{ then } \alpha X \in U.$$

A subspace U of a vector space V is itself a vector space and as all vector spaces, it always contains the null vector 0.

The whole space V and the set consisting of the null vector alone are two special examples of subspaces.

Linear Variety

A subset L of a vector space V is called a linear variety if

$$\left.\begin{array}{r} X, Y \in L \\ \lambda \in R \end{array}\right\} \Rightarrow \lambda X + (1-\lambda)Y \in L.$$

A linear variety is also called an affine set.

It can be shown that if U is a vector subspace and $X_0 \in V$, the set $U + X_0$ is a linear variety and conversely, every linear variety is of the form $U + X_0$ for some vector subspace U and some vector X_0. Moreover, if a linear variety L is such that $L = U_1 + X_1$ and $L + U_2 + X_2$, then $U_1 = U_2$.

Thus, the subspace U such that $L = U + X_0$ is necessarily unique and is called the subspace parallel to the linear variety L.

Linear Combination

A vector X is a linear combination of the vectors $X_1, X_2, .. X_n$ in V if

$$X = \sum_{i=1}^{n} \alpha_i X_i$$

for some scalars α_i, i = 1, 2..n

It is clear that the zero vector can be expressed as a linear combination of any given set of vectors by taking all $\alpha_i = 0$.

If S is any nonempty set of vectors of a vector space V, then the set of all linear combinations of elements of S is a subspace of V. The subspace is then called the subspace spanned by (or generated by) S and is denoted by [S].

Linear combinations of vectors are so vital in the study of vector spaces that the term linear space is used as a synonym for vector space.

3.4. Linear Dependence

A finite set of vectors $\{X_1, X_2, .. X_n\}$ of a vector space V is said to be linearly dependent if there exist scalars α_i, i = 1, 2,..n not all zero, such that

$$\sum_{i=1}^{n} \alpha_i X_i = 0$$

Otherwise, the set $\{X_i\}$ is linearly independent, that is if $\sum_{i=1}^{n} \alpha_i X_i = 0$ implies that $\alpha_i = 0$, for each i, the set $\{X_i\}$ is linearly independent.

By convention, the empty set is said to be linearly independent.

Theorem 3.1. Any set of vectors containing the zero vector is linearly dependent.

Proof: Let $\{X_1, X_2, .. X_k\}$ be a set of vectors where X_k is the zero vector. Then the linear combination

$$0X_1 + 0X_2 + ... + 0X_{k-1} + \alpha_k X_k = 0$$

is satisfied for any α_k

Hence, $(X_1, X_2, .. X_k\}$ is linearly dependent.

Theorem 3.2. Any nonempty subset of a set of linearly independent vectors is itself linearly independent and any set of vectors containing a linearly dependent subset is itself linearly dependent.

Proof: Let us assume that the vectors $X_1, X_2, .. X_k$ are linearly independent and suppose that the vectors $X_2, X_3, .. X_k$ are not. This means that there exist α_i not all zero such that

$$\alpha_2 X_2 + \alpha_3 X_3 + .. + \alpha_k X_k = 0$$

Now this implies that

$$\alpha_1 X_1 + \alpha_1 X_1 + ... + \alpha_k X_k = 0$$

will always have a nontrivial solution because a nontrivial solution of the first equation with $\alpha_1 = 0$ will be a solution of the second. But this contradicts the assumption that the vectors $X_1, X_2, .. X_k$ are linearly independent.

Similarly, it can be shown that a set of vectors containing a linearly dependent subset is itself linearly dependent.

Theorem 3.3. If each of the (m + 1) vectors $Y_1, Y_2, .. Y_{m+1}$ of a vector space

V is a linear combination of the same set of m vectors $X_1, X_2, .. X_m$ of V, then the vectors Y_i $(i = 1, 2, .. m + 1)$ are linearly dependent.

Proof: Let

$$Y_i = \sum_{j=1}^{m} \alpha_{ij} X_j, \qquad i = 1, 2,.. (m+1)$$

The theorem will be proved by induction on m.

For m = 1,
$$Y_1 = \alpha_{11} X_1$$
$$Y_2 = \alpha_{21} X_1$$

If both α_{11} and α_{21} are zero, then $Y_1 = Y_2 = 0$ and Y_1, Y_2 are linearly dependent.

If not, let $\alpha_{11} \neq 0$, Then,

$$\alpha_{11} Y_2 - \alpha_{21} Y_1 = \alpha_{11}\alpha_{21} X_1 - \alpha_{21}\alpha_{11} X_1 = 0.$$

implying that Y_1, Y_2 are linearly dependent.

Let us now assume that the theorem is true for $m = k - 1$ and then we show that it is also true for $m = k$. By hypothesis, we have

$$Y_i = \sum_{j=1}^{k} \alpha_{ij} X_j, \qquad i = 1, 2, ... (k+1)$$

If all $\alpha_{ij} = 0$, then all Y_i are zero and hence they are linearly dependent.

Now assume that at least one α_{ij} is not equal to zero, say $\alpha_{11} \neq 0$.

Let us define

$$Z_i = Y_i - \frac{\alpha_{i1}}{\alpha_{11}} Y_1$$

$$= \sum_{j=2}^{k} \left(a_{ij} - \frac{\alpha_{i1}\alpha_{1j}}{\alpha_{11}} \right) X_j, \qquad i = 2,.. k+1$$

Since each of the k vectors Z_i is a linear combination of the same $(k - 1)$ vectors X_j $(j = 2, .. k)$, by the induction hypothesis, the Z_1 are linearly dependent.

Hence, there exist numbers $\beta_2, \beta_3, .. \beta_{k+1}$, not all zero, such that

$$0 = \sum_{i=2}^{k+1} \beta_i Z_i = \sum_{j=2}^{k+1} \beta_i \left(Y_i - \frac{\alpha_{i1}}{\alpha_{11}} Y_1 \right)$$

$$= \sum_{i=2}^{k+1} \beta_i Y_i = -\frac{1}{\alpha_{11}} \left(\sum_{i=2}^{k+1} \beta_i \alpha_{i1} \right) Y_1$$

which shows that the Y_i are linearly dependent.

Corollary 3.3.1: Any set of n + 1 vectors in R^n is linearly dependent.

Proof: Let e_i, $(i = 1, 2, .. n)$ be the i^{th} unit vector in R^n. Then, any vector $X \in R^n$ can be written as

$$X = \sum_{i=1}^{n} x_i e_i .$$

The corollary then follows from theorem 3.3.

Corollary 3.3.2. Any m vectors in R^n are linearly dependent if m > n.

Proof: Follows from corollary 3.3.1 and theorem 3.2.

Theorem 3.4. If S = $\{X_1, X_2, .. X_k)$ be a set of linearly independent vectors of a vector space V, then for any vector X, the set $S_1 = \{X_1, X_2, .. X_k, X\}$ is linearly independent if and only if X is not an element of the space spanned by S.

Proof: Suppose that X $\in$ [S]. Then, for some scalars α_i, i = 1, 2, ..k

$$X = \alpha_1 X_1 + \alpha_2 X_2 + .. + \alpha_k X_k$$

If $\alpha_i = 0$, for all i, then X is a zero vector and by theorem 3.1, the set S_1 is linearly dependent.

If $\alpha_i \neq 0$ for some i, the set S_1 is again linearly dependent.

Conversely, suppose that X $\in$ [S] and consider the equation

$$\beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + \beta X = 0$$

If $\beta \neq 0$, the vector X can be expressed as

$$X = \frac{-\beta_1}{\beta} X_1 + \frac{-\beta_2}{\beta} X_2 + ... + \frac{-\beta_k}{\beta} X_k$$

Thus, X is a linear combination of vectors of S and therefore X $\in$ [S], contradicting the assumption.

Hence, $\beta = 0$ and since $\{X_1, X_2, .. X_k\}$ is linearly independent, $\beta_i = 0$, for i = 1, 2, .. k. Then $S_1 = \{X_1, X_2, .. X_k X\}$ is linearly independent.

Theorem 3.5. The set of nonzero vectors $X_1, X_2, .. X_n$ is linearly dependent if and only if some X_k, $2 \leq k \leq n$ is a linear combination of the preceding vectors.

Proof: If for some k, X_k is a linear combination of $X_1, X_2, .. X_{k-1}$, then the set of vectors $X_1, X_2, .. X_k$ is linearly dependent and by theorem 3.2, the set of vectors $X_i, X_2, .. X_n$ is linearly dependent.

Conversely, suppose that the vectors $X_1, X_2, .. X_n$ are linearly dependent and let k be the first integer between 2 and n for which $X_1, X_2, .. X_k$ are linearly dependent. Then for suitable scalars α_i, (i = 1, 2,..k) not all zero.

$$\alpha_1 X_1 + \alpha_2 X_2 + ... + \alpha_k X_k = 0$$

If $\alpha_k = 0$, then $X_1, X_2, .. X_{k-1}$ are linearly dependent contradicting the definition of k. Hence $\alpha_k \neq 0$ and

$$X_k = \frac{-\alpha_1}{\alpha_k} X_1 + \frac{-\alpha_2}{\alpha_k} X_2 + ... + \frac{-\alpha_{k-1}}{\alpha_k} X_{k-1} .$$

as was to be proved.

3.5. Basis and Dimension

Basis: A set of maximum number of linearly independent vectors of a vector space V is called a basis of V. V is said to be finite-dimensional if it has a finite basis; otherwise V is infinite-dimensional.

We shall however be concerned with finite-dimensional vector spaces only.

Theorem 3.6: A set S of linearly independent vector $X_1, X_2, .. X_r$ of a vector space V is a basis of V if an only if every vector $Y \in V$ is a linear combination of the vectors $X_1, X_2, .. X_r$.

Proof: Suppose every $Y \in V$ is a linear combination of the vectors $X_1, X_2, .. X_r$. Then by theorem 3.3, Y, $X_1,.. X_r$ are linearly dependent. Consequently, S contains no larger set of linearly independent vectors. Therefore, r is the maximum number of linearly independent vectors in V and hence is a basis of V.

Conversely, suppose that X_i (i = 1, 2..r) constitute a basis of V. Then be definition r is the maximum number of linearly independent vectors. Therefore, for every $Y \in V$, Y, $X_i, X_2,.. X_r$ must be linearly dependent. Hence there exist scalars $\alpha_0, \alpha_1, .. \alpha_r$ not all zero such that

$$\alpha_0 Y + \alpha_1 X_1 + ... + \alpha_r X_r = 0.$$

Obviously $\alpha_0 \neq 0$, for otherwise, it will imply that $X_1, X_2,.. X_r$ are linearly dependent contradicting the assumption.

Hence Y can be expressed as

$$Y = -\frac{1}{\alpha_0}\sum_{i=1}^{r}\alpha_i X_i$$

which is the desired linear combination.

Theorem 3.7. The unit vectors e_i (i = 1,2,..n) in R^n form a basis.

Proof: The unit vectors e_i (i = 1,2,..n) in R^n are linearly independent, since

$$\sum_{i=1}^{n}\alpha_i e_i = 0 \text{ implies that } \alpha_i = 0 \text{ for } i = 1,2,..n$$

and every vector X can be written as

$$X = \sum_{i=1}^{n} x_i e_i$$

Hence e_i (i = 1, 2,..n) form a basis of R^n.

Theorem 3.8. The representation of a vector X of a vector space V as a linear combination of a basis $\{X_1, x_2,.. X_n\}$ is unique.

Proof: Suppose that the vector X can be represented in two different ways,

$$X = \sum_{i=1}^{n}\alpha_i X_i \quad \text{and} \quad X = \sum_{i=1}^{n}\beta_i X_i$$

By subtraction we obtain

$$\sum_{i=1}^{n}(\alpha_i - \beta_i)X_i = 0$$

and since X_i are linearly independent, $\alpha_i = \beta_i$ for i = 1, 2, .. n.

Theorem 3.9. Given a basis $X_1, X_2, .. X_k$ of a vector space V and the linear expression of a nonzero vector X of V expressed as a linear combination of the basis vectors,

$$X = \sum_{i=1}^{k} \alpha_i X_i, \tag{3.1}$$

a new basis is obtained by substituting X for X_i in the current basis if and only if α_i is not equal to zero.

Proof: Suppose that $\alpha_1 \neq 0$ and that the set of vectors $X, X_2,..X_k$ are linearly dependent, so that

$$\lambda X + \lambda_2 X_2 + ... + \lambda_k X_k = 0 \tag{3.2}$$

implies that the scalars $\lambda, \lambda_2, ... \lambda_k$ are not all zero.

It is clear that λ cannot be equal to zero, since by assumption $X_2, X_3, .. X_k$ are linearly independent.

Substituting, the expression for X from (3.1) in (3.2) we obtain

$$\lambda\alpha_1 X_1 + \sum_{i=2}^{k}(\lambda_i + \lambda\alpha_i)X_i = 0. \tag{3.3}$$

But $\lambda\alpha_1 \neq 0$. This contradicts the assumption $X_1, X_2, .. X_k$ are linearly independent. Hence, λ must be equal to zero and consequently $\lambda_i = 0$ for i = 2, 3, .. k, which implies that $X, X_2, ... X_k$ are linearly independent.

We now show that the linearly independent vectors $X, X_2, .. X_k$ form a basis. Since $X_1, X_2, .. X_k$ is a basis, any vector Y in V can be expressed as

$$Y = \sum_{i=1}^{k} \beta_i X_i \tag{3.4}$$

Since $\alpha_1 \neq 0$, from (3.1), we have

$$X_1 = \frac{1}{\alpha_i} X - \sum_{i=2}^{k} \frac{\alpha_i}{\alpha_1} X_i \tag{3.5}$$

Substituting (3.5) in (3.4), we obtain

$$Y = \frac{\beta_1}{\alpha_1} X + \sum_{i=2}^{k} \left(\beta_i - \frac{\alpha_i \beta_1}{\alpha_1}\right) X_i \tag{3.6}$$

The vector Y is thus expressed as a linear combination of vectors $X, X_2, .. X_k$ and hence $X, X_2, .. X_k$ form a basis of V.

If $\alpha_1 = 0$, we note from (3.1) that the vectors X, X_2, .. X_k are linearly dependent. Hence, the vectors, X, X_2,..X_k are linearly independent if and only if $\alpha_1 \neq 0$ and they form a basis.

Theorem 3.10. Every basis of a finite-dimensional vector space V has the same number of elements.

Proof: Let S = {X_1,.. X_n} and T = {Y_1, Y_2,.. Y_m} be bases of a finite dimensional vector space V. Since each one is a set of maximum number of linearly independent vectors, after vectors the set S_1 = {Y_1, X_1,.. x_n} is linearly dependent. By theorem 3.5 some X_i is a linear combination of the vectors which precede it and there exists a subset S'_1, of S

$$S'_1 = \{Y_1, X_1 .. X_{i-1}, X_{i+1}, .. X_n\}$$

which contains Y_1 as the first vector and is a basis of V.

Consider now the set

$$S_2 = \{Y_2, S'_1\}.$$

and again we note that S_2 is linearly dependent and some vector is a linear combination of the preceding vectors. This vector cannot be any of the Y's, since Y's are independent. Hence, there exists a subset S'_2 of S_2 containing Y_2 and Y_1 as the first two vectors and form a basis of V. The set S_3 = {Y_3, S'_2} is then linearly dependent and we apply the same argument. Continuing in this way, we see that all the X's will not be replaced before the Y's are exhausted, since otherwise the remaining Y's would have to be linear combinations of the ones that have already replaced the X's, which contradicts that Y's are linearly independent.

Hence, $m \leq n$.

Interchanging the roles of S and T in the replacement process, we obtain $n \leq m$.

Hence, m = n.

Dimension

The dimension of a finite-dimensional vector space V is the number of elements (vectors) in a basis of V.

The trivial space of empty vectors is said to have dimension zero and R^n is an n-dimensional vector space.

Theorem 3.11. Every set of (n + 1) vectors in an n-dimensional vector space V is linearly dependent.

Proof: Follows from corollary 3.3.1.

Theorem 3.12. Any linearly independent set of vectors in an n-dimensional vector space V can be extended to a basis.

Proof: Let {X_1, X_2,.. X_m} be a set of linearly independent vectors, which is not a basis of V. Hence, m < n and there exists a vector X_{m+1} such that $X_{m+1} \notin [X_1, X_2,.. X_m]$ and by theorem 3.4, the set {X_1, X_2,.. X_m, X_{m+1}} is a linearly independent set. If m+1 < n, we can repeat the argument. The process is continued until we obtain a set of n linearly independent vectors, which then is a basis of V.

Theorem 3.13. A set of n vectors in an n-dimensional vector space V is a basis of V if and only if it is linearly independent.

Proof: Let S = $\{X_1, X_2, .. X_n\}$ be a set of linearly independent vectors of V. If this set is not a basis, it can be extended to a basis by theorem 3.12. But a basis of V contains only n vectors and hence the set is a basis.

The only if statement follows from the definition of a basis.

3.6. Inner Product Spaces

So far we have considered vector spaces with reference to the concept of linearity but ignored the usual concepts of geometric measurement such as length, distance and angle. To generalize these concepts we are to introduce a metric in the vector space. By observing, these measurements in R^2, it can be seen that a distance function is a suitable metric to be introduced. This can be conveniently done by means of the inner product of a pair of vectors in the vector space.

An inner product in a real (complex) vector space is a real (complex) numerically valued function of the ordered pair of vectors X and Y defined by (X, Y) = $\sum_i x_i \bar{y}_i$, where $\bar{y}_i$ is the complex conjugate of y_i, such that

$$(X, Y) = \overline{(Y, X)} \tag{3.7}$$

$$(\alpha_1 X_1 + \alpha_2 X_2, Y) = \alpha_1 (X_1, Y) + \alpha_2 (X_2, Y) \tag{3.8}$$

$$(X, X) \geq 0, \text{ and } (X, X) = 0, \text{ if and only if } X = 0. \tag{3.9}$$

An inner product space is a vector space with an inner product.

The properties (3.7–3.9) imply that in a complex vector space, an inner product is Hermitian symmetric, conjugate bilinear and positive definite. In the case of a real vector space, an inner product is symmetric, bilinear and positive definite. Moreover, in either case, real or complex, $\|X\|$ satisfies the homogeneity property $\|\alpha X\| = |\alpha|. \|X\|$.

A real inner product space is called a Euclidean space and its complex analogue is called a unitary space. In other words, a real vector space in which a real inner product is defined is called a Euclidean space and a complex vector space in which a complex inner product is defined is called a unitary space.

R^n, the set of all real n-vectors is an example of n-dimensional Euclidean space.

Norm or Length

In an inner product space, the real nonnegative number

$$(XX)^{1/2} = \|X\|$$

is called the norm or length of the vector x. (3.10)

It has the following properties:

(a) $\|\alpha X\| = |\alpha|.\|X\|$

(b) $\|X\| \geq 0$ and $\|X\| = 0$, if and only if X = 0.

(c) $\|X + Y\| \leq \|X\| + \|Y\|$.

A vector space on which the norm $\|X\|$ is defined is called a normed vector space.

Orthogonality, Distance and Angle

Two vectors X and Y in an inner product space are said to be orthogonal if (X, Y) = 0. It is obvious that the vector 0 is orthogonal to every vector.

If X is a nonzero vector in an inner product space then $X/\|X\|$ is a vector of unit length and is called a normal vector.

A set of mutually orthogonal vectors, each of which is normal is called an orthonormal set, that is for X, Y in the set

$$\begin{aligned} (X, Y) &= 0, \text{ if } X \neq Y \\ (X, Y) &= 1, \text{ if } X = Y. \end{aligned} \tag{3.11}$$

Note that an orthonormal set is linearly independent.

Theorem 3.14 (Bessel Inequality)

Let $\{Y_1, Y_2, .. Y_n\}$ be a finite orthonormal set in an inner product space S. If X is any vector in S, then

$$\sum_{i=1}^{n} |(X, Y_i)|^2 \leq \|X\|^2.$$

Further, the vector $X' = X - \sum_{i=1}^{n} (X, Y_i) Y_i$ is orthogonal to each Y_j.

Proof:

$$0 \leq \| X' \|^2 = \left(X - \sum_{i=1}^{n} (X, Y_i) Y_i, X - \sum_{j=1}^{n} (X, Y_j) Y_j \right)$$

$$= (X, X) - \sum_{i=1}^{n} (X, Y_i)\overline{(X, Y_i)} - \sum_{j=1}^{n} (X, Y_j)\overline{(X, Y_j)}$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{n} (X, Y_i)\overline{(X, Y_j)}(Y_i, Y_j)$$

$$= \| X \|^2 \sum_{i=1}^{n} |(X, Y_i)|^2 - \sum_{i=1}^{n} |(X, Y_i)|^2 + \sum_{i=1}^{n} |(X, Y_i)|^2$$

$$= \| X \|^2 \sum_{i=1}^{n} |(X, Y_i)|^2 .$$

Now,

$$(X', Y_j) = \left(X - \sum_{i=1}^{n} (X, Y_i) Y_i, Y_j \right)$$

$$= (X, Y_j) - \sum_{i=1}^{n} (X, Y_i)(Y_i, Y_j)$$

$$= (X, Y_j) - (X, Y_j) = 0.$$

Theorem 3.15. (Schwarz Inequality)

If X and Y are vectors in an inner product space, then

$$|(X, Y)| \leq \|X\|. \|Y\|.$$

Proof: If Y = 0, the result is trivially true since then both sides vanish. Let us therefore assume that $Y \neq 0$. Then, the set consisting of the vector $Y/\|Y\|$ is orthonormal and hence by Bessel's inequality (theorem 3.14) we have

$$\left|\left(X, \frac{Y}{\|Y\|}\right)\right|^2 \leq \|X\|^2$$

which reduces to $|(X, Y)| \leq \|X\|. \|Y\|$.

Distance

The distance between two vectors X and Y in an inner product space is defined by

$$d(X, Y) = \|X - Y\| = (X - Y, X - Y)^{½}$$

which satisfies the following properties:

(a) $d(X, Y) = d(Y, X)$.

(b) $d(X, Y) \geq 0$; $d(X, Y) = 0$ if and only if $X = Y$ (3.12)

(c) $d(X, Y) \leq d(X, Z) + d(Z, Y)$.

$d(0, X)$ is then the distance of X from the origin and is called the length of X, which is equal to $\|X\|$ as stated in (3.10).

A nonempty set in which a distance function d is defined is said to be a metric space. The distance function d is called a metric. Thus, a normed vector space becomes a metric space, if the metric defined on it is the distance function d.

The properties (a) and (b) in (3.12) are obviously true and the triangular inequality (c) can be easily proved with the help of Schwarz's inequality.

Let U, V be two vectors in an inner product space, then

$$\begin{aligned}
\| U + V \|^2 &= (U + V, U + V) \\
&= (U, U) + (U, V) + (V, U) + (V, V). \\
&= \| U \|^2 + (U, V) + \overline{(U, V)} + \| V \|^2 \\
&= \| U \|^2 + 2R_e (U, V) + \| V \|^2 \\
&\leq \| U \|^2 + 2 \,|(U, V)| + \| V \|^2 \\
&\leq \| U \|^2 + 2 \| U \| \cdot \| V \| + \| V \|^2, \text{ by Schwarz's inequality} \\
&= (\| U \| + \| V \|)^2.
\end{aligned}$$

Setting $U = X - Z$ and $V = Z - Y$, we obtain

$$\|X - Y\| \leq \|X - Z\| + \|Z - Y\|$$

that is, $d(X, Y) \leq d(X, Z) + d(Z, Y)$.

Angle

Schwarz's inequality also enables us to define the angle between two vectors in a Euclidean space. We note that in the case of a Euclidean space, the expression

$$\frac{(X, Y)}{\| X \| \cdot \| Y \|}$$

is a real number between -1 and $+1$. Hence, it is the cosine of an angle θ in the range $0 \le \theta \le \pi$.

Therefore, the angle θ between the vectors X, Y in a Euclidean space is defined to be an angle in the range $0 \le \theta \le \pi$ such that

$$\text{Cos } \theta \frac{(X, Y)}{\| X \| \cdot \| Y \|}.$$

Two vectors X and Y are therefore orthogonal if $(X, Y) = 0$.

CHAPTER 4

Matrices and Determinants

4.1. Matrices

An m x n matrix is a rectangular array of m n numbers arranged in m rows and n columns and is written as

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots a_{1n} \\ a_{21} & a_{22} & \dots a_{2n} \\ a_{m1} & a_{m2} & \dots a_{mm} \end{bmatrix} = (a_{ij}) \tag{4.1}$$

The numbers a_{ij} (i = 1, 2..m; j = 1, 2..n) are called the elements of the matrix. The first subscript refers to the row and the second to the column.

It should be noted that a matrix has no numerical value. It is simply a convenient way of representing arrays of numbers.

Special Matrices

If m = n, the matrix is said to be a square matrix of order n.

If m = 1, the matrix has only one row and if n = 1, it has only one column. The matrices with one row or one column are also vectors. Thus the vector a_i,

$$a_i = (a_{i1}, a_{i2}, \dots a_{in})$$

is called the ith, row vector of A

and the vector a_j,

$$a_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

is called the jth column vector of A.

An m x n matrix may therefore be considered as made up of m row vectors a_i (i = 1, 2, ...m) or of n column vectors a_j (j = 1, 2, ...n).

If all the elements of a matrix are zero, the matrix is called a null or zero matrix and is denoted by **0**.

A square matrix, whose elements outside the main diagonal (the diagonal running from upper left to lower right) are all zero is called a diagonal matrix.

A triangular matrix is one whose elements above or below the main diagonal are all zero.

If in a square matrix of order n, all the rows (columns) are the n unit vectors e_i, (i = 1, 2..n), it is called an identity matrix (or unit matrix) of order n and is written as

$$I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = (\delta_{ij}) \quad (i, j = 1, 2, \dots n)$$

where δ_{ij} is the Kronecker delta defined by $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $i \neq j$.

The principal (main) diagonal elements of I_n are all unity and the remaining elements are zero.

The transpose of a matrix $A = (a_{ij})_{m \times n}$ is obtained from A by interchanging the rows and columns and is denoted by A^T.

Thus, $\qquad A^T = (a_{ji})_{n \times m}$

A square matrix A is called symmetric if it is equal to its transpose, that is, if $A = A^T$. It is a skew symmetric matrix if $A = -A^T$.

4.2. Relations and Operations

Equality: Two matrices A and B are equal if the corresponding elements are equal. Thus, A = B, if and only if $a_{ij} = b_{ij}$, for every ij.

Addition: The sum of two m x n matrices A and B is an m × n matrix C whose elements are,

$$c_{ij} = a_{ij} + b_{ij}, \qquad i = 1, 2..m, \quad j = 1, 2..n.$$

Thus, $C = (c_{ij}) = (a_{ij} + b_{ij}) = A + B$.

It should be noted that the equality and addition of matrices are defined only when the matrices have the same number of rows and the same number of columns. It is clear from the definition that the matrix addition is commutative and associative, i.e.

$$A + B = B + A.$$

$$(A + B) + C = A + (B + C).$$

Moreover, we have

$$(A + B)^T = A^T + B^T.$$

that is, the transpose of the sum of two matrices A, B is equal to the sum of A transpose and B transpose.

Multiplication by a Scalar: The product of a matrix A with a scalar λ is the matrix

$$\lambda A = (\lambda a_{ij})$$

Product of Two Matrices: The product of an $m \times p$ matrix A with an $p \times n$ matrix B is the matrix C, whose elements are given by

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}.$$

Thus, $$A\,B = C = (c_{ij}), \qquad i = 1, 2..m$$
$$j = 1, 2..n$$

The product A B is defined only if the number of columns of the premultiplier A is equal to the number of rows of the postmultiplier B. If A and B are square matrices of the same order, AB and BA are defined but in general AB $\neq$ BA. For example,

let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

then

$$AB = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

but

$$BA = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Moreover, if AB = 0, it does not imply that either A or B must be zero. It is always possible to find nonnull matrices A, B whose product AB is zero. For example,

$$\begin{bmatrix} 3 & -1 \\ -6 & 2 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Although, matrix multiplication is not in general commutative, it satisfies the associative and the distributive law with respect to addition and the homogeneity property with respect to scalar multiplication when the appropriate operations are defined.

$$(A\,B)\,C = A\,(B\,C)$$
$$A(B + C) = AB + AC$$
$$A\,(\lambda B) = \lambda\,(A\,B)$$

Transpose of the Product: It can be easily verified that the transpose of the product of two matrices is the product of the transposes in reverse order, that is $(AB)^T = B^T A^T$.

Orthogonal Matrix: A square matrix A is said to be an orthogonal matrix if and only if

$$A\,A^T = A^T A = I.$$

Idempotent Matrix: A matrix A is said to be an idempotent matrix if and only if $A^2 = A$.

An example other than the null and the unit matrix is

$$A = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$$

Submatrix: If in an $m \times n$ matrix A, all but r rows and s columns are deleted, the resulting $r \times s$ matrix is called a submatrix of A.

4.3. Partitioning of Matrices

A matrix A may be partitioned into submatrices by drawing lines parallel to the rows or columns

Thus,

$$A = \left[\begin{array}{cccc|cc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{24} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ \hdashline a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where A_{ij} are the submatrices.

One advantage of partitioning is that it simplifies multiplication of matrices if they are partitioned into conformable submatrices.

Suppose that A and B are two matrices, the number of columns of A being equal to the number of rows of B so that the product of AB is defined. Suppose A, B are partitioned as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} \qquad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

where the number of columns of A_{11} is equal to the number of rows of B_{11} and the number of columns of A_{12} is equal to the number of rows of B_{21}.

It can then be easily verified that

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \\ A_{31}B_{11} + A_{32}B_{21} & A_{31}B_{12} + A_{32}B_{22} \end{bmatrix}$$

Thus, the product of AB may be obtained by applying the multiplication rule to the appropriate submatrices of A and B.

The partitioning of matrices also enables in to write a system of linear equations AX = b, in a convenient way.

If the m x n matrix A is partitioned into n submatrices $a_1, a_2..a_n$, each a column vector with m components, and X is an n-component column vector, the system of linear equations may be written as

$$AX = [a_1 a_2..a_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_{j=1}^{n} a_j x_j = b.$$

4.4. Rank of a Matrix

Let $A = (a_{ij})$ be an m x n matrix which may be considered to be composed of m row vectors $a_i = (a_{i1}, a_{i2}, .. a_{in})$ or n column vectors $a_j = (a_{1j} .. a_{nj})^T$.

The maximum number of linearly independent row vectors (column vectors) is called the row rank (column rank) of the matrix A.

Theorem 4.1. The row rank of a matrix A is equal to its column rank.

Proof: Let r be the row rank of A and S be its column rank and suppose that $r < s$.

Without any loss of generality, we assume that the rows and the columns are so arranged that the first r rows and the first s columns of A are linearly independent. The submatrix of A with these r rows and s columns is denoted by

$$\hat{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ \vdots & \vdots & & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rs} \end{bmatrix} \tag{4.2}$$

Since r < s, there exists a non zero column vector Y (corollary 3.3.2) such that

$$\hat{a}_i Y = \sum_{j=1}^{s} a_{ij} y_j = 0, \qquad i = 1, 2,..r$$

$$\text{or} \quad \hat{A}Y = 0. \tag{4.3}$$

Also, since $a_1, a_2,..a_r$ are a row basis, it follows (theorem 3.6) that

$$a_k = \sum_{i=1}^{r} \lambda_{ik} a_i, \qquad k = 1, 2,..m \tag{4.4}$$

for some numbers λ_{ik}.

Hence, $$\hat{a}_k = \sum_{i=1}^{r} \lambda_{ik} \hat{a}_i \tag{4.5}$$

and therefore, $$\hat{a}_k Y = \sum_{i=1}^{r} \lambda_{ik} (\hat{a}_i Y) = 0, \text{ for all k} \tag{4.6}$$

which can also be written as,

$$\sum_{j=1}^{s} y_j a_j = 0 \tag{4.7}$$

This shows that the vectors a_j, $j \le s$ are linearly dependent, contradicting the assumption that they are linearly independent.

Hence, we must have $r \ge s$.

Starting with the assumption that $r > s$, we can use the same argument interchanging the role of rows and columns and get the reverse inequality $r \le s$.

Hence, $r = s$.

Definition. The rank of a matrix is defined to be the maximum number of linearly independent rows (or columns) of the matrix.

The rank of a matrix A is denoted by r(A).

A square matrix of order n is called nonsingular (or regular), if its rank is n, otherwise, it is called singular.

4.5. Determinants

The determinant of a square matrix $A = (a_{ij})$ of order n, denoted by $|A|$ (or det A) is the number defined by the sum

$$|A| = \sum_{\alpha} \delta(\alpha) a_{1,\alpha(1)} a_{2,\alpha(2)} \ldots a_{n,\alpha(n)}. \tag{4.8}$$

where α runs over all permutations of $\{1, 2,..n\}$, and $\delta(\alpha)$ is 1 if α is an even permutation and -1 if α is an odd permutation.

The determinant is thus a homogeneous polynomial in the a_{ij}; it has n! terms and each term includes one and only one element of each row and of each column of A.

The determinant of A is often denoted by

$$|A| = |a_{ij}| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \tag{4.9}$$

and is called an nth order determinant.

It can be easily verified that the determinant A may also be defined by keeping the second indices in the natural order and carrying out the permutations on the first indices in (4.8)

It should be noted that while a matrix has no numerical value, a determinant is a number.

4.6. Properties of Determinants

The following are some useful properties of determinants:

(1) If every element in a row or a column of a matrix A is zero, then the value of the determinant is zero.

(2) An interchange of two columns (or rows) of a determinant $|A|$, changes the sign of the determinant. If immediately follows that

(3) A determinant having two identical columns (rows) is zero.

(4) If A is a matrix of order n and λ a scalar, then $|\lambda A| = \lambda^n|A|$.

(5) $|A^T| = |A|$.

(6) The value of a determinant is unchanged by adding to a column (row) a linear combination of other columns (rows).

(7) For square matrices A, B of the same order $|A + B| \neq |A| + |B|$, in general.

(8) If A, B are square matrices of the same order then $|AB| = |A|.\ |B|$, that is, the determinant of the product is the product of the determinants. The multiplication of determinants is defined by the same rule as that of multiplication of matrices.

(9) If the columns (or the rows) of a determinant are linearly dependent, the determinant is zero.

4.7. Minors and Cofactors

The determinant of the submatrix obtained by deleting the ith row and the jth column of a square matrix A is called the minor $|M_{ij}|$ of the element a_{ij}.

The cofactor A_{ij} of the element a_{ij} is then defined by the determinant

$$A_{ij} = (-1)^{i+j}\,|M_{ij}|. \tag{4.10}$$

Expansion by Cofactors

It follows directly from (4.8) that for every i and j

$$|A| = \sum_j a_{ij}A_{ij} = \sum_i a_{ij}A_{ij} \tag{4.11}$$

These are the cofactor expansion of $|A|$ by the ith row and the jth column.

Further,

$$\sum_j a_{ij}A_{kj} = 0 = \sum_j a_{ji}A_{jk}, \qquad \text{if} \quad i \neq k \tag{4.12}$$

since these expressions are the row and column expansion of a determinant with two identical rows or columns.

Combining (4.11) and (4.12), we have

$$\sum_j a_{ij}A_{kj} = |A|.\delta_{ik} = \sum_j a_{ji}A_{jk} \tag{4.13}$$

where δ_{ik} is the Kronecker delta.

Adjoint Matrix: Let A_{ij} be the cofactor of a_{ij} in a square matrix A of order n and let $a^+_{ij} = A_{ji}$,

Then,

$$A^+ = (a^+_{ij}) = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix} \tag{4.14}$$

is called the adjoint matrix of A.

In other words, the adjoint matrix A^+ is the transpose of the matrix whose element (i, j) is the cofactor of a_{ij} in A.

(4.13) can then be written as

$$A.A^+ = A^+.A = |A|I_n \tag{4.15}$$

4.8. Determinants and Rank

Theorem 4.2. The rank of an m x n matrix A is r if and only if there exists at least one square submatrix of A of order r whose determinant is not zero and the determinant of every square submatrix of order greater than r is zero.

Proof: Necessity: Let the matrix A be of rank r. Hence, any (r + k) columns of A are linearly dependent and therefore by property 9, the determinant of every square submatrix of order (r + k) is zero.

We are now to show that there is at least one submatrix of A of order r whose determinant does not vanish. Suppose on the contrary, that all determinants of order r vanish. A being of rank r, there are r columns of A which are linearly independent and without loss of generality, let us assume that these are the first r columns of A. Now, suppose that the determinants of all the submatrices of order r in the first r columns and in particular the determinant of the submatrix formed from the first r rows vanish. Let A_{rj} be the cofactor of the element a_{rj} in that determinant.

We then have,

$$\sum_{j=1}^{r} a_{rj}A_{rj} = 0 \tag{4.16}$$

The same cofactors are obtained if a submatrix of order r is formed from the first r – 1 rows and any other row i, i = r + 1,...m.

Hence, $$\sum_{j=1}^{r} a_{rj}A_{rj} = 0, \quad i = r + 1,...m \tag{4.17}$$

Moreover, $$\sum_{j=1}^{r} a_{rj}A_{rj} = 0, \quad i = 1 + 2,...r - 1 \tag{4.18}$$

since the expression (4.18) is the expansion of the determinant with two identical rows.

It follows from (4.16), (4.17) and (4.18), that the first r columns of A are linearly dependent which contradicts the hypothesis.

Sufficiency: Now, suppose that every determinant of order greater than r is zero and that there exists at least one determinant of order r which does not vanish.

Let the matrix A is of rank s. s cannot be greater than r, because from the proof of the necessity, we note that there would exist at least one nonzero determinant of order s. Hence, $s \leq r$.

s cannot be less than r, because from the definition of rank, s columns of A are linearly independent and consequently every determinant of order r would be zero (Property 9) Hence, $s \geq r$.

The rank of A therefore, is r.

Nonsingular Matrices

A square matrix A is nonsingular if and only if its determinant is not zero ($|A| \neq 0$). It is called singular if $|A| = 0$

4.9. The Inverse Matrix

4.9.1. Definition and General Properties

Definition: If A be a square matrix of order n, then a matrix B, if it exists, such that

$$A\,B = B\,A = I_n,$$

is called the inverse of A.

Inverse of A is usually denoted by A^{-1}.

Theorem 4.2. The inverse of a matrix, if it exists is unique.

Proof: Suppose that B and C are two matrices, which are inverses of the matrix A, so that

$$AB = BA = I$$

and $$AC = CA = I$$

We then have, $$CAB = C(AB) = CI = C$$

$$CAB = (CA)B = IB = B$$

Hence, B = C.

Theorem 4.3. The inverse of a square matrix A exists if and only if A is nonsingular.

Proof: The condition is necessary.

Let B be the inverse of A, so that

$$AB = I.$$

Then, $|A|\ |B| = |I| = 1$

and hence, $|A| \neq 0$

The condition is sufficient:

Let $|A| \neq 0$.

It has already been shown in (4.15) that

$$AA^{+} = A^{+}A = |A|.I,$$

where A^{+} is the adjoint matrix of A

Consequently, $A^{+}/|A| = A^{-1}$.

If follows that the matrix $B = A^{-1}$ is also a nonsingular matrix of order n. Moreover, inverse of a singular matrix is not defined since

$$|AB| = |A|.|B| = 0 \times |B| \neq 1.$$

Properties of the Inverse

(a) If A is a nonsingular matrix, then,

(i) $(A^{-1})^{-1} = A$

(ii) $(A^{T})^{-1} = (A^{-1})^{T}$

(b) If A and B are nonsingular matrices, then,

$$(AB)^{-1} = B^{-1}A^{-1}$$

(c) If A is nonsingular and B is any other matrix, then,

$$AB = 0 \text{ implies } B = 0.$$

4.9.2. Inversion by Partitioning

Let A be a nonsingular matrix of order n partitioned as follows:

$$A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix},$$

where α is an $r \times r$ submatrix, β an $r \times s$ submatrix γ an $s \times r$ submatrix and δ an $s \times s$ submatrix $(r + s = n)$ such that δ is nonsingular and δ^{-1} is known.

A^{-1} exists and let it be partitioned in the same manner as A, so that

$$A^{-1} = \begin{bmatrix} \lambda & \mu \\ \rho & \nu \end{bmatrix}$$

Since $AA^{-1} = I$, we have

$$\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}\begin{bmatrix} \lambda & \mu \\ \rho & \nu \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & I_s \end{bmatrix}$$

and we obtain the equations

$$\alpha\lambda + \beta\rho = I_r$$
$$\alpha\mu + \beta\nu = 0$$
$$\gamma\lambda + \delta\rho = 0$$
$$\gamma\mu + \delta\nu = I_s$$

Solving these for λ, μ, ρ, ν we obtain

$$\lambda = (\alpha - \beta\delta^{-1}\gamma)^{-1}$$
$$\mu = -\lambda\beta\delta^{-1}$$
$$\rho = -\delta^{-1}\gamma\lambda$$
$$\nu = \delta^{-1} - \delta^{-1}\gamma\mu.$$

4.9.3. Product Form of the Inverse

Let $B = (b_1, b_2, .. b_n)$ be an $n \times n$ nonsingular matrix for which the inverse B^{-1} is known. Now another matrix Ba_r is formed by replacing the rth column of B by a vector 'a', i.e.

$$Ba = (b_1, b_2, ..b_{r-1}, a, b_{r+1} .. b_n)$$

and we wish to find the inverse of Ba.

Since B is nonsingular, the vectors b_j ($j = 1, 2, .. n$) are linearly independent and thus form a basis. The vector 'a' can then be expressed as

$$a = \sum_{j=1}^{n} y_j b_j = BY \tag{4.19}$$

A necessary and sufficient condition that Ba will be nonsingular is that $y_r \neq 0$. (See theorem 3.9). Assuming that this is true, we may write

$$b_r = -\frac{1}{y_r}\sum_{j\neq r} y_j b_j + \frac{1}{y_r} a$$

or $\qquad b_r = Ba\, V_r \qquad (4.20)$

where $\qquad V_r^T = \left(-\frac{y_1}{y_r}, ..., -\frac{y_{r-1}}{y_r}, \frac{1}{y_r}, -\frac{y_{r+1}}{y_r}, ..., -\frac{y_n}{y_r}\right)$

Defining the matrix $E = (e_1, e_2, .. e_{r-1}, V_r, e_{r+1}, .. e_n)$, where e_i is the ith unit vector, we have

$$B = Ba\, E$$

and hence, $\qquad Ba^{-1} = E\, B^{-1}. \qquad (4.21)$

From (4.19), we note that $Y = B^{-1} a$, from which Vr can be computed.

The preceding result may then be used for a step by step computation of the inverse B^{-1} of a nonsingular matrix

$$B = (b_1, b_2,.. b_n).$$

We start with the identity matrix

$$I_n = (e_1, e_2,... e_n)$$

and replace e_1 by b_1 to form the matrix B_1 and then replace e_2, the second column of B_1 to form the matrix B_2 and so on until e_n is replaced by b_n, thus forming the matrix $B_n = B$

By the previous result, we then have,

$$B_1^{-1} = E_1 I_n = E_1.$$

$$B_2^{-1} = E_2 B_1^{-1} = E_2 E_1.$$

$$\vdots$$

$$B^{-1} = B_n^{-1} = E_n B_{n-1}^{-1} = E_n E_{n-1},.. E_2 E_1. \tag{4.22}$$

where $$E_i = (e_1, e_2,..e_{i-1}, V_i, e_{i+1},..e_n) \tag{4.23}$$

$$V_i^T = \left(-\frac{y_{i1}}{y_{ii}}, -\frac{y_{i2}}{y_{ii}},.. -\frac{y_{i,i-1}}{y_{ii}}, \frac{1}{y_{ii}}, -\frac{y_{i,i+1}}{y_{ii}},..., -\frac{y_{in}}{y_{ii}} \right) \tag{4.24}$$

$$y_i^T = (y_{i1},..y_{in})$$

and $$y_i = B_{i-1}^{-1}\, b_i. = E_{i-1}, E_{i-2}...E_i b_i,\ B_0^{-1} = I_n \tag{4.25}$$

Since $B^{-1} = B_n^{-1}$ is expressed as the product of E matrices as in (4.22) it is called a product form of the inverse.

CHAPTER 5

Linear Transformations and Rank

5.1. Linear Transformations and Rank

A linear transformation T is a mapping of a vector space R^n into a vector space R^m, which has the following properties.

If $X_1, X_2 \in R^n$, then $T(X_1 + X_2) = T(X_1) + T(X_2)$;

If $X \in R^n$ and λ, a scalar, then $T(\lambda X) = \lambda T(X)$, (5.1)

Clearly, $T(0) = 0$, since $T(0) = T(0 + 0) = T(0) + T(0)$ and $T(X-Y) = T(X) - T(Y)$, since $T(X-Y) + T(Y) = T(X-Y+Y) = T(X)$.

It follows that the range of T r(J), $R(T) = \{TX \mid X \in R^n\}$ is a subspace of R^m and the null space of T, $N(T) = \{X \mid TX = 0\}$ is a subspace of the domain R^n and

$$\dim R(T) + \dim N(T) = n.$$

The dimension of the range space R(T) is called the rank of the transformation and the dimension of the null space N(T) is called the nullity of T.

Theorem 5.1. Every linear transformation of R^n into R^m may be completely determined by an $m \times n$ matrix A.

Proof: Let A be an $m \times n$ matrix. Then for any vector $X \in R^n$, the vector $Y = AX$ can be considered to be a vector in R^m. Moreover, the rules of matrix operations establish that the matrix A preserves the addition and the homogeneity property of linear transformations.

Thus, the $m \times n$ matrix A is a linear transformation of R^n into R^m.

Conversely, let $\{e_1, e_2, .. e_n\}$ and $\{f_1, f_2, .. f_m\}$ denote the usual orthonormal basis of R^n and R^m, respectively. A vector $X \in R^n$ and its transformation $Y = T(X)$ may then be expressed as

$$X = \sum_{j=1}^{n} x_j e_j \tag{5.2}$$

$$Y = \sum_{i=1}^{m} y_i f_i \tag{5.3}$$

Since $e_j \in R^n$, has a transformation $T(e_j) \in R^m$, it may be expressed as

$$T(e_j) = \sum_{i=1}^{m} a_{ij} f_i \ , \qquad j = 1, 2,..n. \tag{5.4}$$

for some unique $a_{1j}, a_{2j},.. a_{mj}$

We then obtain

$$Y = T(X) = T\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1}^{n} x_j T(e_j)$$

$$= \sum_{j=1}^{n} x_j \left(\sum_{j=1}^{m} a_{ij} f_i\right)$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{m} x_j a_{ij} f_i$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j\right) f_i \tag{5.5}$$

Comparing with (5.3), we have $y_i = \sum_{j=1}^{n} a_{ij} x_j$ and considering $A = (a_{ij})$ to be the $m \times n$ matrix, we have $Y = A X$.

The matrix A then represents the given linear transformation.

Moreover, it is clear that we may replace $\{e_1, e_2,.. e_n\}$ and $\{f_1, f_2,.. f_m\}$ by any other fixed ordered bases in R^n and R^m and set up similar one-to-one correspondence between linear transformations of R^n into R^m and $m \times n$ matrices.

Theorem 5.2. The rank of a linear transformation T is equal to the rank of the corresponding matrix A.

Proof: The linear transformation of a vector $X \in R^n$ is the vector $T(X) = Y$, which by theorem 5.1 is simply the matrix product Ax giving the image elements of R^m.

Thus, if the $m \times n$ matrix A corresponds to T, then

$$Y = AX = \sum_{j=1}^{n} x_j a_j$$

so that the dimension of the subspace R(T), called the rank of T, is equal to the maximum number of independent columns of A, which by definition is the rank of the matrix A.

A linear transformation T is said to be nonsingular if the corresponding matrix A is nonsingular. The transformation is said to be orthogonal if A is an orthogonal matrix.

5.2. Product of Linear Transformations

Let T_1 be a linear transformation which maps R^n into R^k and T_2, a linear transformation which maps R^k into R^m. Then it is easy to verify that the product $T_3 = T_2T_1$ of the two linear transformations T_1, T_2 defined by

$$T_3(X) = T_2[T_1(X)]$$

is also a linear transformation which maps R^n into R^m.

It should however, be noted that multiplication of transformations is not commutative.

The corresponding matrix product is easily obtained. If the matrix A $(k \times n)$ corresponds to T_1 and the matrix B$(m \times k)$ corresponds to T_2, then their product $C = BA$ is an $m \times n$ matrix and corresponds to the linear transformation T_3.

If for $X \in R^n$, $\quad AX = Y, \quad Y \in R^k$

and $\quad BY = Z, \quad Z \in R^m$

Then $\quad Z = BY = B(AX) = (BA)X.$

Thus, $\quad C = BA$ maps R^n into R^m.

From theorem 5.2, we note that the rank of the transformation C is the rank of the product BA of the two matrices. This leads to the following important result.

Theorem 5.3. If A and B are two matrices then the rank of the product AB

$$r(AB) \le \min[r(A), r(B)].$$

Proof: The matrices A, B represent linear transforms and since $(AB)X = A(BX)$, it follows that R(AB) is contained in R(A), so that

$$r(AB) \le r(A).$$

Similarly, $(B^TA^T)Y = B^T(A^TY)$

which implies that $r(B^TA^T) \le r(B^T)$

Now since the rank of a matrix is equal to the rank of its transpose

$$r(AB) \le \min[r(A), r(B)].$$

Theorem 5.4. The rank of the product of a matrix A with a nonsingular matrix is equal to the rank of the matrix A.

Proof: Let a matrix A be post-multiplied by a nonsingular matrix B.

Then by theorem 5.3.

$$r(AB) \le r(A).$$

Again, since $A = AB.B^{-1}$

$$r(A) \le r(AB)$$

Hence, $\quad r(AB) = r(A).$

If A is premultiplied by a nonsingular matrix, the proof can be established in exactly the same way.

5.3. Elementary Transformations

Certain simple operations which are applied to a system of linear equations to find an efficient solution procedure (Gaussian elimination), are also found useful in finding the rank of a matrix. These operations when performed on the rows and columns of a matrix reduce it to a matrix of the same rank from which the rank of the matrix can be found very easily.

These operations are called elementary transformation. They are:

(1) Interchange of two rows (columns).

(2) Multiplication of a row (column) by a nonzero scalar.

(3) Addition of one row (column) to another row (column).

5.3.1. Elementary Matrices

A matrix obtained from a unit matrix, by subjecting it to any of the elementary transformation is called an elementary matrix.

It can be easily verified that

(i) every elementary matrix is nonsingular, and

(ii) each of the elementary row (column) transformation of a matrix may be considered as the result of the premultiplication (postmultiplication) with the corresponding elementary matrix.

Examples

(1) Exchange of the first and second rows.

Let $$A = \begin{bmatrix} 3 & 4 \\ 2 & 1 \\ 6 & 5 \end{bmatrix}$$

Consider the elementary matrix

$$E = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

then $$EA = \begin{bmatrix} 2 & 1 \\ 3 & 4 \\ 6 & 5 \end{bmatrix}$$

(2) Multiply the second row by 3,

Considering, $$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$ we have

$$EA = \begin{bmatrix} 3 & 4 \\ 6 & 3 \\ 6 & 5 \end{bmatrix}$$

(3) Add four times the second row to the third row.

Taking, $$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix},$$ we have

$$EA = \begin{bmatrix} 3 & 4 \\ 2 & 1 \\ 14 & 9 \end{bmatrix}$$

Analogous elementary transformations can be performed on the columns of a matrix by postmultiplying the matrix by the corresponding elementary matrices obtained by the desired transformations on the columns of a unit matrix.

Since elementary matrices are nonsingular, it is obvious that the product of a matrix with an elementary matrix does not change the rank of the matrix.

5.4. Echelon Matrices and Rank

A matrix having the following structure is called an echelon matrix:

1. The first k rows are nonzero; the other rows are zero.
2. In each nonzero row, the first nonzero element starting from the left is 1.
3. If c_i denotes the column in which the element unity occurs, which is the first unity in row i, then

$$c_1 < c_2 < .... < c_k.$$

Example,

$$\begin{bmatrix} 0 & 1 & a_{13} & a_{14} & a_{15} & a_{16} \\ 0 & 0 & 1 & a_{24} & a_{25} & a_{26} \\ 0 & 0 & 0 & 0 & 1 & a_{36} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Theorem 5.5. Every matrix A can be reduced to an echelon matrix by a series of elementary transformations.

Proof: The proof presented here is constructive, i.e. it describes how the reduction is actually carried out to obtain the echelon matrix.

The following steps are followed:

1. Move the columns of A having all elements zero in the beginning of the matrix.
2. If the first column having at least one nonzero element does not occur in the first row, exchange the first row with any other row having a nonzero element. Let this element be α. Divide the first row by α so that the element in the first row is now 1 (i.e., $a_1c_1 = 1$).
   If a_ic_1, i = 2,...m, the elements in other rows of this column are nonzero, reduce them to zero by subtracting a_ic_1 times the first row from them.

3. We now move to the next column which has at least one nonzero element in one of the rows 2, 3,...m and repeat step 2 to obtain $a_2c_2 = 1$.
4. Continue in this way until an echelon matrix is obtained.

Theorem 5.6. The rank of an echelon matrix H is the number of nonzero rows in H.

Proof: Let k be the number of rows in H, which have at least one nonzero element. Obviously, the rank of the matrix H cannot be greater than k. To establish the linear independence of the k nonzero rows h_i (i = 1) we are to show that the linear relation between them

$$\sum_i \lambda_i h_i = 0$$

implies that all $\lambda_i = 0$

Since row h_1 has an element unity in column c_1 and all other elements in this column are zero, $\lambda_i = 0$, Further we have,

$$\lambda_1 a_1c_2 + \lambda_2 = 0 \qquad \Rightarrow \lambda_2 = 0$$
$$\lambda_1 a_1c_3 + \lambda_2 a_2c_3 = 0 \qquad \Rightarrow \lambda_3 = 0$$

and so on

and thus all $\lambda_i = 0$

Hence, r(H) = k

It is thus observed that any matrix A can be reduced to an echelon matrix H by premultiplying it by a suitable elementary matrix E

i.e. EA = H.

and since E is nonsingular, rank of A is immediately known from the echelon matrix H. By theroem 5.6, the rank of H is simply the number of nonzero rows in H.

It can also be shown that by postmultiplying the echelon matrix H by a suitable elementary matrix E_1, it can further be reduced to the matrix

$$H\,E_1 = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}$$

Since H = EA, we have

$$E\,A\,E_1 = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix},$$

where k is the rank of A.

If A is an n x n nonsingular matrix, $E\,A\,E_1 = I_n$

Moreover, since every nonsingular matrix is a product of elementary matrices, it can be said that for every matrix A of rank k there exist nonsingular matrices P and Q such that

$$P\,A\,Q = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}$$

The matrix B = PAQ is called equivalent to A.

CHAPTER 6

Quadratic Forms and Eigenvalue Problems

6.1. Quadratic Forms

A homogeneous expression of the second degree

$$Q = \sum_{i,j=1}^{n} a_{ij} x_i x_j . \tag{6.1}$$

in the n variables $x_1, x_2, .. x_n$, where a_{ij} are real constants is called a quadratic form.

In matrix notation it can be written as

$$Q = X^T AX$$

where $X^T = (x_1, x_2 .. x_n)$, and

$$A = \begin{bmatrix} a_{11} & a_{12} & .. & a_{1n} \\ a_{21} & a_{22} & .. & a_{2n} \\ a_{x1} & a_{x2} & .. & a_{nn} \end{bmatrix}$$

Since the coefficient of $x_i x_j$ in the above expression is $(a_{ij} + a_{ji})$ for $i \neq j$, we can define

$$b_{ij} = b_{ji} = \frac{a_{ij} + a_{ji}}{2} \quad \text{for all i, j} \tag{6.2}$$

and the quadratic form Q can then be expressed as

$$Q = X^T BX,$$

where $B = [b_{ij}]$ is a symmetric matrix

It should be noted that this redefinition of the coefficients does not change the value of Q for any X. Hence, without loss of generality, we always assume that the matrix A associated with the quadratic form $X^T AX$ is symmetric and the quadratic form then is a symmetric quadratic form.

The symmetric matrix A is said to be the matrix of the quadratic form and the rank of the quadratic form is defined to be the rank of A

Example: The matrix of the quadratic form

$$Q = 3x_1^2 + 4x_1x_2 + 6x_2x_1 + x_2^2$$

is $A = \begin{bmatrix} 3 & 4 \\ 6 & 1 \end{bmatrix}$, which is not symmetric.

If we now define

$$b_{12} = b_{21} = \frac{6+4}{2} = 5, \; b_{11} = 3, b_{22} = 1.$$

we note that, $Q = X^TAX = X^TBX$,
where the matrix $B = (b_{ij})$ is symmetric.

6.2. Definite Quadratic Forms

A quadratic form X^TAX is said to be positive definite if $X^TAX > 0$ for every X except $X = 0$. It is said to be positive semidefinite if $X^TAX \geq 0$ for every X and there exists atleast one $X \neq 0$, for which $X^TAX = 0$.

Negative definite and negative semidefinite quadratic forms are defined by reversing the inequality signs in the above definitions.

A quadratic form X^TAX is said to be indefinite if it is positive for some vectors X and negative for others.

A symmetric matrix A is said to be positive definite, positive semidefinite, negative definite, etc., according as the quadratic form associated with it is so.

Examples.

(1) The quadratic form $Q = 3x_1^2 + 2x_2^2 + x_3^2$ is positive definite as it is positive for every $X = [x_1, x_2, x_3]^T$ except when $x_1 = x_2 = x_3 = 0$.

(2) $Q = (x_1 - x_2)^2 + 2x_3^2$ is positive semidefinite as it is always nonnegative and is zero if $x_1 = x_2, x_3 = 0$.

(3) $Q = -x_1^2 - 3x_2^2$ is negative definite and

(4) $Q = x_1^2 - 3x_2^2$ is indefinite since it is positive when $x_1 = 2, x_2 = 1$ and negative when $x_1 = 1, x_2 = 1$.

Theorem 6.1. If A is a positive semidefinite matrix, then $X^TAX = 0$ implies that $AX = 0$.

Proof: For any Y and arbitrary λ, we have

$$\begin{aligned} 0 &\leq (Y + \lambda X)^T A (Y + \lambda X) \\ &= Y^TAY + 2\lambda Y^TAX + \lambda^2 X^TAX \\ &= Y^TAY + 2\lambda Y^TAX, \quad \because X^TAX = 0. \end{aligned} \tag{6.3}$$

Now, the above expression holds only if the coefficient of λ vanishes, i.e. $Y^TAX = 0$ and since (6.3) must be true for every Y, it follows that $AX = 0$.

Theorem 6.2. Let A be a symmetric matrix of order n and B be any n x m $(n > m)$ matrix.

(i) If the matrix A is positive semidefinite, then the symmetric matrix B^TAB is also positive semidefinite.

(ii) If A is positive definite and $r(B) = m$, then B^TAB is also positive definite.

Proof: (i) Let $Y = BX$, then the quadratic form corresponding to the symmetric matrix B^TAB, $X^TB^TABX = Y^TAY \geq 0$ for all Y and hence for all X and it assumes the value zero for nonzero values of X satisfying $BX = 0$, (i.e. $Y = 0$)

(ii) Since $r(B) = m$, the m columns of B are linearly independent. Then $X \neq 0$ implies $Y \neq 0$ and since A is positive definite, $Y^TAY > 0$ for all $Y \neq 0$, Hence, $X^TB^TABX > 0$ for all $X \neq 0$.

Corollary 6.1. The symmetric matrix B^TB, called the gram matrix of B, is positive definite or positive semidefinite according as the rank of B is equal to m or less than m.

Proof: This follows from theorem 6.2 with A replaced by the unit matrix I.

Theorem 6.3. If A is an $n \times n$ symmetric matrix of rank r, then there exists a nonsingular matrix P such that

$$P^TAP = \begin{bmatrix} D_r & 0 \\ 0 & 0 \end{bmatrix}$$

where D_r is a nonsingular r-rowed diagonal matrix.

Proof: The result is obtained by premultiplying and postmultiplying A with pairs of elementary matrices each of which is the transpose of the other. Thus, there exists a nonsingular matrix P, the product of the postmultiplied elementary matrices such that P^TAP is a diagonal matrix with r nonzero diagonal elements.

Corollary 6.2. Let $Q = X^TAX$ be a quadratic form where A is $n \times n$ symmetric matrix of rank r. Then there exists a nonsingular linear transformation $X = PY$ such that

$$X^TAX = Y^TP'APY = Y^TDY$$

where D is a diagonal matrix, so that

$$Y^TDY = d_1y_1^2 + d_2y_2^2 + .. + d_ry_r^2$$

where d_j, $1 \leq j \leq r$ are nonzero numbers.

Theorem 6.4. Let A be an n x n symmetric matrix of rank r. Then there exists a nonsingular matrix P such that P^TAP is a diagonal matrix with 1 in the first s diagonal positions, – 1 in the next r – s diagonal positions and zeros elsewhere.

Proof: By theorem 6.3, there exists a nonsingular matrix L such that L^TAL is a diagonal matrix with r nonzero diagonal elements.

Let $L^TAL = \text{diag}\,[b_1, b_2, ..b_r, 0, 0. ... 0]$

and suppose that s of the nonzero elements are positive so that (r - s) elements are negative. Without loss of generality, let us assume that $b_1, b_2, ..b_s$ are positive and $b_{s+1}, b_{s+2}, ..b_r$ are negative.

Now, there exist numbers $\beta_1, \beta_2,..\beta_r$ such that

$$\beta_1^2 = b_1, \beta_2^2 = b_2, ...\beta_s^2 = b_s; \beta_{s+1}^2 = -b_{s+1}, \beta_{s+2}^2 = -b_{s+2} \cdots \beta_r^2 = -b_r$$

Let M = diag $\left[\beta_1^{-1},...\beta_r^{-1},1,1,...1\right]$.

Taking P = LM, we obtain

$$P^TAP = M^TL^TALM$$

$$= \text{diag}\left[\beta_1^{-1},...\beta_r^{-1},1,...1\right].\text{diag}\left[b_1,b_2,...b_r,0,0..0\right].\ \text{diag}\left[\beta_1^{-1},...\beta_r^{-1},1,...1\right]$$

$$= \text{diag}\left[1,1,...1-1,..-1,0,0..0\right].$$

so that 1 and –1 appear s and (r – s) times, respectively.

Corollary 6.3. If $Q = X^TAX$ is a quadratic form of rank r then there exists a nonsingular linear transformation X = PY, which transforms X^TAX to

$$Y^TP^TAPY = y_1^2 + y_2^2 + ... + y_3^2 - y_{s+1}^2 - ... - y_r^2$$

The above expression is called the cannonical form of the quadratic form X^TAX.

The number of positive terms s and the difference s – (r – s) = 2s – r between the number of positive and negative terms are respectively called the index and signature of the quadratic form.

Corollary 6.4: Let $Q = X^TAX$ be a quadratic form in n variables of rank r and index s. Then the quadratic form Q is

(a) positive definite, if r = n, s = n
(b) negative definite, if r = n, s = 0
(c) positive semidefinite, if r < n, s = r
(d) negative semidefinite, if r < n, s = 0.

In every other case the quadratic form is indefinite.

Theorem 6.5. A set of necessary and sufficient conditions for the quadratic form $Q = X^TAX$ to be positive definite is

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad ... \quad \begin{vmatrix} a_{11} & ... & a_{1n} \\ a_{n1} & ... & a_{nn} \end{vmatrix} > 0.$$

In other words, a quadratic form X^TAX is positive definite if and only if the leading principal minors of A are all positive.

6.3. Characteristic Vectors and Characteristic Values

A nonzero vector X such that

$$AX = \lambda X, \text{ for some scalar } \lambda \tag{6.4}$$

where $A = (a_{ij})$ is a given matrix of order n is called a characteristic vector of A. The scalar λ is called the characteristic value of A associated with the characteristic vector X.

Now, a nonzero solution X of (6.4) exists if and only if (A – λI) is singular, that is, if and only if

$$|A - \lambda I| = 0. \tag{6.5}$$

Obviously, the determinant is a polynomial of degree n in λ and can be written as

$$f(\lambda) = |A - \lambda I| = (-1)^n\lambda^n + b_1\lambda^{n-1} + \ldots\ldots + b_{n-1}\lambda + b_n \qquad (6.6)$$

where the b's are sums of products of a_{ij}. $f(\lambda)$ is called the characteristic polynomial of the matrix A.

The equation (6.5) is called the characteristic equation and its roots λ_i (i = 1,2, .. n) are called the characteristic values, eigenvalues or latent roots of the matrix A. The characteristic vectors are also called eigenvectors of A.

In general, the eigenvalues of a matrix A need not be real numbers—they may be complex. It can however be shown that if the matrix A is symmetric, the eigenvalues are real.

Theorem 6.6. If A is a real symmetric matrix, then all eigenvalues and eigenvectors of A are real.

Proof: If possible, let the eigenvalue λ of the real symmetric matrix be complex. Then all the components of an eigenvector corresponding to λ cannot be real and we have

$$A\,X = \lambda X \qquad (6.7)$$

$$A\,\bar{X} = \bar{\lambda}\,\bar{X} \qquad (6.8)$$

where '–' denotes the complex conjugate.

Multiplying (6.7) by $\bar{X}^T$ and (6.8) by X^T, we have

$$\bar{X}^T AX = \lambda\bar{X}^T X \qquad (6.9)$$

$$X^T A\bar{X} = \bar{\lambda} X^T\bar{X} \qquad (6.10)$$

Since $\bar{X}^T X = X^T\bar{X}$ and $\bar{X}^T AX = X^T A\bar{X}$, subtracting (6.10) from (6.9) we get

$$(\lambda - \bar{\lambda})X^T\bar{X} = 0 \qquad (6.11)$$

Further, since $X^T\bar{X}$ is real and positive ($\because$ $X \neq 0$) we have $\lambda = \bar{\lambda}$ and hence λ is real.

Now, since the coefficients of the homogeneous linear equations $(A - \lambda I)X = 0$ are real, the components of the vector X are also real.

It can also be noted that the eigenvectors corresponding to different eigen values are orthogonal.

Theorem 6.7. (Cayley Hamilton theorem)

Every square matrix satisfies its characteristic equation, that is, if A is a matrix with characteristic equation

$$(-1)^n\lambda^n + b_1\lambda^{n-1} + \ldots\ldots + b_{n-1}\lambda + b_n = 0$$

$$\text{then } (-1)^n A^n + b_1 A^{n-1} + \ldots\ldots + b_{n-1}A + b_n I = 0$$

Proof: Consider the adjoint matrix of $(A - \lambda I)$. Then $C = \text{adj}\,(A - \lambda I)$ is a matrix, whose elements are polynomials in λ of degree $(n - 1)$

Thus, $C = C_0 + C_1\lambda + .. + C_{n-2}\lambda^{n-2} + C_{n-1}\lambda^{n-1}$.

where the C_r are matrices whose elements are polynomials in the a_{ij}.

Now, by (4.15),

$$(A - \lambda I)C = |A - \lambda I| \times I.$$
$$= ((-1)\lambda^n + b_1\lambda^{n-1} + .. + b_n)I.$$

or

$$(A - \lambda I)(C_0 + C_1\lambda + .. + C_{n-2}\lambda^{n-2} + C_{n-1}\lambda^{n-1})$$
$$= ((-1)^n\lambda^n + b_1\lambda^{n-1} + .. + b_n)I.$$

Equating the coefficients of λ, we have.

$$-C_{n-1} = (-1)^n I$$
$$-C_{n-2} + AC_{n-1} = b_1 I$$
$$-C_{n-3} + AC_{n-2} = b_2 I$$
$$\cdots\cdots$$
$$-C_0 + AC_1 = b_{n-1} I$$
$$AC_0 = b_n I$$

Premultiplying the above successively with

$$A^n, A^{n-1}, A^{n-2} \ldots. A, I$$

and adding we obtain

$$(-1)^n A^n + b_1 A^{n-1} + b_2 A^{n-2} \ldots + b_{n-1} A + b_n = 0.$$

Theorem 6.8. If $Q = X^TAX$ is a quadratic form of rank $r \leq n$, then there exists an orthogonal transformation

$$X = PY$$

which transforms Q to the form

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + .. + \lambda_r y_r^2$$

where λ_i's are the r nonzero eigenvalues of A.

Proof: For any real symmetric matrix A with rank r, there exists an orthogonal matrix P whose columns are an orthonormal set of eigenvectors of A, such that

$$P^TAP = P^{-1}AP = D$$

Where D is a diagonal matrix whose diagonal elements are the eigenvalues of A, that is,

$$D = \text{diag}(\lambda_1, \lambda_2 .. \lambda_r, 0, 0 .. 0).$$

Let $X = PY$, then

$$Q = X^TAX = Y^TP^TAPX = Y^TDY$$
$$= \lambda_1 y_1^2 + \lambda_2 y_2^2 + .. + \lambda_r y_r^2$$

Corollary 6.5. The quadratic form $Q = X^TAX$ is

(a) Positive (negative) definite if and only if all eigenvalues of A are positive (negative).

(b) Positive (negative) semidefinite if and only if all eigenvalues of A are nonnegative (nonpositive) and at least one of the eigenvalues is zero.

(c) Indefinite if and only if the matrix A has both positive and negative eigenvalues.

Example. Consider the quadratic form

$$Q = 7x_1^2 + 10x_2^2 + 7x_3^2 - 4x_1x_2 + 2x_1x_3 - 4x_2x_3$$

The symmetric matrix A is then

$$A = \begin{bmatrix} 7 & -2 & 1 \\ -2 & 10 & -2 \\ 1 & -2 & 7 \end{bmatrix}$$

The eigenvalues are the roots of the characteristic equation

$$|A - \lambda| = \begin{bmatrix} \lambda-7 & 2 & -1 \\ 2 & \lambda-10 & 2 \\ -1 & 2 & \lambda-7 \end{bmatrix} = 0$$

or $\qquad \lambda^3 - 24\lambda^2 + 180\lambda - 432 = 0$

or $\qquad (\lambda - 6)^2 (\lambda - 12) = 0$

The eigenvalues are then 6, 6, 12 which are all positive. Therefore, the quadratic form is positive definite.

Theorem 6.9. Let $Q_1 = X^TAX$ and $Q_2 = X^TBX$ be two real quadratic forms in n variables and let Q_2 be positive definite. Then there exists a real nonsingular transformation $X = PY$ which transforms the quadratic forms to

$$\lambda_1 y_1^2 + .. + \lambda_n y_n^2$$

and $\qquad y_1^2 + .. + y_n^2$

respectively, where $\lambda_1, \lambda_2, .. \lambda_n$ are the roots of

$$|A - \lambda B| = 0$$

CHAPTER 7

Systems of Linear Equations and Linear Inequalities

7.1. Linear Equations

Consider a system of m linear equations in n unknowns $x_1, x_2, \ldots x_n$

$$
\begin{aligned}
a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n &= b_1 \\
a_{21} x_1 + a_{22} x_2 + \ldots + a_{2n} x_n &= b_2 \\
&\vdots \\
a_{m1} x_1 + a_{m2} x_2 + \ldots + a_{mn} x_n &= b_m
\end{aligned} \tag{7.1}
$$

where a_{ij}, b_i (i = 1,2..m, j = 1,2..n) are known constants.

The system of equations (7.1) can also be written in matrix notation as

$$AX = b \tag{7.2}$$

where the m x n matrix $A = (a_{ij})$ is called the coefficient matrix or the matrix of the system and the matrix $A_b = (A, b)$ is called the augmented matrix of the system. Necessarily, $r(A) \leq r(A_b)$.

Any vector X which satisfies (7.2) is called a solution to the system. The system is called consistent if it has at least one solution, otherwise it is said to be inconsistent. If b = 0, the system of equations is said to be homogeneous and if $b \neq 0$, it is called nonhomogeneous. A homogeneous system of equations always has a solution X = 0, which is called a trivial solution.

Two system of linear equations are said to be equivalent if they have the same set of solutions.

7.2. Existence Theorems for Systems of Linear Equations

Theorem 7.1. Any system of m homogenous linear equations in n unknowns always has a nonzero solution if m < n.

Proof: Let AX = 0 be the system of linear equation where A is an m x n matrix and X is an n component vector, m < n.

Let the m-vector a_j be the jth column vector of a, so that $a_j = [a_{1j}, a_{2j}, \ldots a_{mj}]^T$, for j = 1,2,..n. a_j are then a set of n vectors in m dimensional space where n > m.

Hence by corollary 3.3.1, they are dependent and therefore there exist x_j's not all zero such that

$$\sum_{j=1}^{n} a_j x_j = 0.$$

Theorem 7.2. If A is an m × n matrix of rank r, the set S of all solutions of the homogeneous system AX = 0 is an (n – r) dimensional subspace of R^n.

Proof: Since S, the set of all solutions of the system AX = 0, is closed under addition and scalar multiplication, S is a subspace of R^n. We shall now show that the dimension of this subspace is (n – r).

Since A is of rank r, there exist r linearly independent column vectors of A. Suppose (reordering if necessary) that these are the first r column vectors $a_1, a_2, \ldots a_r$ so that for j > r, we have

$$a_j = \sum_{k=1}^{r} \alpha_{jk} a_k, \qquad j = r+1, \ldots n \tag{7.3}$$

Let us define

$$\bar{X}_j^T = (-\alpha_{j1}, -\alpha_{j2}, \ldots -\alpha_{jr}, 0, 0, \ldots 0, 1, 0, \ldots 0), \; j = r+1, \ldots n. \tag{7.4}$$

where the component 1 is the j th component of $\bar{X}_j$

It follows from (7.3) that each $\bar{X}_j$, (j = r + 1,...n) is a solution of AX = 0. Moreover, it follows that $\bar{X}_j$ are linearly independent.

It now remains to show that every vector in S is a linear combination of the $\bar{X}_j$.

Let $X = (x_j)$ be an arbitrary solution of AX = 0.

Then $$X^* = X - \sum_{j=r+1}^{n} x_j \bar{X}_j \tag{7.5}$$

is also a solution and we note from (7.4) that

$$x_j^* = 0, \qquad \text{for } j = r+1, \ldots n. \tag{7.6}$$

Hence, $$\sum_{j=1}^{r} a_j x_j^* = 0$$

and since a_j, j = 1,2...r are linearly independent,

$$x_j^* = 0, \qquad \text{for } j = 1, 2..r \tag{7.7}$$

Thus, $$x_j^* = 0, \qquad \text{for } j = 1,2..n \tag{7.8}$$

and from (7.5), we then have

$$X = \sum_{j=r+1}^{n} x_j \bar{X}_j$$

as was to be shown.

Theorem 7.3. If r (A) = m, the system of m equations in n unknowns AX = b (which implies $m \leq n$) always has

(i) a unique solution if m = n

(ii) an infinite number of solutions if m < n.

Proof: (i). If r(A) = m = n, the n column vectors a_1, a_2,..a_n being linearly independent form a basis of R^n. Hence, any vector b can be uniquely expressed as a linear combination of the basis vectors, that is,

$$a_1 x_1 + a_2 x_2 +..+ a_n x_n = b \tag{7.9}$$

$X = (x_1, x_2,..x_n)^T$ is therefore the unique solution of the system.

In fact, when r(A) = m = n, the matrix a is nonsingular and A^{-1} exists. A unique solution of the system AX = b can then be obtained by premultiplying both sides of the system by A^{-1} so that

$$X = A^{-1} b. \tag{7.10}$$

We thus have an explicit solution of the system of equations through the use of the inverse of the coefficient matrix. If A^{-1} is not known, the well known Cramer's rule may be applied to obtain the solution.

(ii) If m < n, we can assign arbitrary values to the unknowns associated with the (n–m) vectors of A not forming the basis and the values of the remaining unknowns can be obtained uniquely as in case (i). Since the (n–m) unknowns are assigned arbitrary values, the system has infinitely many solutions.

Theorem 7.4. In a system of m linear equations in n unknowns AX = b, if

(i) $r(A) = r(A_b) = n$, the system has a unique solution,

(ii) $r(A) = r(A_b) < n$, the system has infinite number of solutions,

(iii) $r(A) < r(A_b)$, the system has no solution.

Proof: If $r(A_b) = r < m$, r rows of A_b are linearly independent. Let us assume that these are the first r rows. Every row of A_b can then be expressed as a linear combination of these r linearly independent rows of A_b. The (m–r) last equations can therefore be dropped from the system without any effect on the solutions and hence they are redundant.

The system is then reduced to an equivalent system

$$\bar{A}X = \bar{b} \tag{7.11}$$

where $\bar{A}$ is a r × n matrix and $r(\bar{A}, \bar{b}) = r$

Moreover, $r(A) = r(\bar{A})$

If $r(\bar{A}) = r$, the theorem 7.3 can be applied to the system (7.11) to prove cases (i) and (ii).

Case (iii)

If $r(A) = r(\bar{A}) < r = r(\bar{A}, \bar{b})$ then every set of r linearly independent columns of $(\bar{A}, \bar{b})$ must contain the column $\bar{b}$. Hence, $\bar{b}$ is linearly independent of the columns of $\bar{A}$.

Thus, there are no x_j such that

$$\sum_{j=1}^{n} x_j a_j = b$$

Hence the system AX = b has no solution.

Theorem 7.5. A necessary and sufficient condition for a system of m homogeneous linear equations in n unknowns ($m \leq n$) AX = 0 to have a nonzero solution is that r (A) < n.

Proof: If r (A) = n, A is nonsingular and

$$X = A^{-1} 0 = 0.$$

Thus there is a unique solution and it is trivial. Now let r(A) < n. Since b = 0, $r(A) = r(A_b) < n$ and the theorem then follows from case (ii) of theorem 7.4.

Theorem 7.6. If X_0 is a solution of the nonhomogeneous system of equations AX = b, then for all solutions X of the homogeneous system AX = 0,

$$X_0 + X \tag{7.12}$$

is a solution of AX = b and all solutions of the nonhomogeneous system can be expressed in the form (7.12)

Proof: Let X_0 be a solution of the nonhomogeneous system AX = b and X be a solution of the homogeneous system AX = 0. We then have

$$A (X_0 + X) = AX_0 + AX = b + 0 = b.$$

which shows that $X_0 + X$ is a solution of AX = b.

Now let Y be an arbitrary solution of AX = b, so that we have

$$A (Y - X_0) = AY - AX_0 = b - b = 0$$

which implies that $U = Y - X_0$ is a solution of the homogeneous system and we have

$$Y = X_0 + U.$$

7.3. Basic Solutions and Degeneracy

Consider a system of m linear equations in n > m unknowns AX = b and suppose that $r(A) = r(A_b) = r < n$. We have noted earlier that if r < m, (m–r) equations of the system are redundant and can therefore be ignored. Without loss of generality therefore, we assume that $r(A) = r(A_b) = m$.

Let B be a submatrix of A formed from m linearly independent columns of A. These columns then constitute a basis of the set of column vectors of A and the submatrix B is often called a basis matrix of A.

The system of equations can then be written as

$$AX = BX_B + NX_N = b \tag{7.13}$$

where B is an m x m nonsingular matrix, N is an m × (n–m) matrix and $X^T = [X_B^T, X_N^T]$, X_B, X_N being the vectors of variables associated with the columns of B and N, respectively.

All the solutions of the linear system (7.13) can then be obtained by assigning arbitrary values to X_N. The particular solution by setting $X_N = 0$, is given by

$$X_B = B^{-1} b, X_N = 0 \tag{7.14}$$

and $X^T = [(X_B = B^{-1}b)^T, X_N^T = 0]$ is called the basic solution associated with B.

The components of X_B are called the basic variables and the remaining (n–m) components of X are known as nonbasic variables

Since there are atmost $\binom{n}{m}$ sets of m linearly independent vectors from the n column vectors of A,

$$\binom{n}{m} = \frac{n!}{m!(n-m)!} \tag{7.15}$$

is the maximum number of basic solutions.

Degeneracy: A basic solution to AX = b is called degenerate if one or more of the basic variables have a zero value. If all basic variables have a nonzero value, the solution is called nondegenerate.

Theorem 7.7. A necessary and sufficient condition for the existence and nondegeneracy of all possible basic solutions of the system of m equations in n unknowns AX = b, where r (A) = m is that every set of m columns from the augmented matrix $A_n = (A, b)$ is linearly independent.

Proof: (Necessity). Let us suppose that all $\binom{n}{m}$ basic solutions exist and none is degenerate. In that case, every set of m columns from A must be linearly independent and for any set of m columns say, a_j, (j = 1,2, .. m) of A, we have

$$\sum_{j=1}^{m} a_j x_j = b$$

$$\text{and} \quad x_j \neq 0, \qquad \text{for } j = 1,2,..m. \tag{7.16}$$

Since none of $x_j = 0$, any column a_j, (j = 1,2, .. m) can be replaced by b and the new set of vectors also form a basis (see theorem 3.9).

Hence, b and any (m–1) columns from A are linearly independent.

(Sufficiency): Let us now suppose that any set of m columns from A_b are linearly independent. This implies that all basic solutions exist. Then a_1, a_2, .. a_m are linearly independent and b can be expressed as a linear combination of a_1, a_2,..a_m, that is,

$$\sum_{j=1}^{m} a_j x_j = b \tag{7.17}$$

Now by assumptions, the vectors

$$a_1, a_2, .. a_m \tag{7.18}$$

and also $$b, a_2, .. a_m \tag{7.19}$$

are linearly independent.

This means that if a_1 in (7.18) is replaced by b, the set (7.19) also forms a basis and therefore the coefficient x_1 of a_1 in (7.17) cannot be zero. Similarly, since a_1, b, a_3, ..a_m are linearly independent, the coefficient x_2 of a_2 cannot be zero. Thus, b can replace any of the m columns of every basis matrix from A and none of the x_j can be zero for any basic solution. Hence all basic solutions exist and are nondegenerate.

7.4. Theorems of the Alternative

We have been considering the conditions which ensure the existence of a solution of a system of linear equations. It is sometimes however, useful to know the conditions under which the systems of equations or inequalities do not have a solution. The following theorems give positive criteria for determining when a system of linear equations or inequalities has no solution.

Theorem 7.8. (Gale [183])

Let A be an m x n matrix and b be an m vector. Then exactly one of the following two systems has a solution:

System 1. $AX = b$, for some $X \in R^n$

System 2. $A^TY = 0$

$b^TY = \alpha$, for some $Y \in R^m$

where α is any nonzero number.

Proof: Suppose that System 1 has a solution, that is, there is an $X \in R^n$ such that $AX = b$. Then for every Y, $X^T AY = b^TY$

If Y is a solution of System 2, then the l.h.s. of the above equation is zero whereas the r.h.s. is equal to $\alpha \neq 0$. Hence system 2 cannot have a solution.

Now, suppose that System 1 has no solution. Let r be the rank of A and without loss of generality, let us assume that the first r columns $a_1, a_2,...a_r$ of A are linearly independent. Then these vectors together with b are linearly independent otherwise b would be a linear combination of the a_j's giving a solution of System 1.

It then follows that there exists a vector Y such that (see theorem 7.3)

$$a_j^T Y = 0, \qquad j = 1, 2,..r$$

$$b^TY = \alpha$$

for every α.

Moreover since A is of rank r, we have

$$a_k = \sum_{j=1}^{r} \lambda_j a_j, \qquad k = r + 1,...n.$$

and then $$a_k^T Y = \sum_{j=1}^{r} \lambda_j a_j^T Y = 0, \qquad k = r + 1,...n.$$

Hence, $a_j^T Y = 0,$ for all $j = 1, 2,...n.$

Thus, $A^T Y = 0,$ and $b^T Y = \alpha$

which shows that System 2 has a solution.

Theorem 7.9. (Farkas theorem [155]).

Let A be an m x n matrix and b be an m vector. Then exactly one of the following two systems has a solution.

System 1: $AX = b, X \geq 0, \qquad X \in R^n$

System 2: $A^T Y \geq 0, b^T Y < 0, \qquad Y \in R^m$

Proof: [183] It is easy to see that both Systems 1 and 2 cannot hold simultaneously for if X and Y are their solutions, we have

$X^T A^T Y = b^T Y$ from System 1

and $Y^T AX \geq 0$ from System 2, since $X \geq 0$

But this contradicts that $b^T Y < 0$.

If $Ax = b$ has no solution, then by theorem 7.8 taking $\alpha < 0$, we find that there exists $Y \in R^m$ such that $A^T Y = 0$. $b^T Y = \alpha < 0$, Hence Y is a solution of System 2.

We then suppose that $Ax = b$ has a solution but no nonnegative solution and show by induction on n, the number of columns of A that in that case System 2 has a solution.

If $n = 1$, $Ax = b$ reduces to

$$a_1 x_1 = b$$

and by hypothesis it has a solution $x_1 < 0$.

Then $Y = -b$ is a solution of System 2, since

$$b^T Y = -b^2 < 0$$

and $$a_1^T Y = \frac{b^T Y}{x_1} > 0$$

Let us now assume that the theorem is true when the number of columns of A is $(n - 1)$ and show that it is also true when the number of columns is n.

Since by hypothesis, the system of equations

$$\sum_{j=1}^{n} a_j x_j = b, \tag{7.20}$$

has no nonnegative solution, it follows that the system

$$\sum_{j=1}^{n-1} a_j x_j = b \tag{7.21}$$

also has no nonnegative solution since a nonnegative solution of (7.21) with $x_n = 0$ would satisfy (7.20)

Hence by the inductive hypothesis, there exists a vector Y_1 such that

$$\begin{aligned} a_j^T Y_1 &\geq 0, \qquad j = 1, 2, \ldots (n-1) \\ b^T Y_1 &< 0 \end{aligned} \tag{7.22}$$

If also $a_n^T Y_1 \geq 0$, then Y_1 satisfies System 2 and the theorem is proved.

If $a_n^T Y_1 < 0$, let us set

$$\bar{a}_j = a_n\left(a_j^T Y_1\right) - \left(a_n^T Y_1\right) a_j, \quad \text{fro } j = 1, 2, \ldots (n-1) \tag{7.23}$$

$$\bar{b} = a_n\left(b^T Y_1\right) - \left(a_n^T Y_1\right) b \tag{7.24}$$

Now, the system of equations

$$\sum_{j=1}^{n-1} \bar{a}_j \bar{x}_j = \bar{b} \tag{7.25}$$

cannot have a nonnegative solution for if $\bar{x}_j \geq 0$, substituting (7.28) and (7.24) in (7.25), we get

$$a_n \left[\frac{1}{-a_n^T Y_1} \left\{ \sum_{j=1}^{n-1} \left(a_j^T Y_1\right) \bar{x}_j - b^T Y_1 \right\} \right] + \sum_{j=1}^{n-1} a_j \bar{x}_j = b \tag{7.26}$$

which shows that the system 1 has a nonnegative solution contrary to our assumption.

Then by the inductive hypothesis applied to (7.25), we note that there exists a vector $\bar{Y}$ such that

$$\bar{a}_j^T \bar{Y} \geq 0, \quad j = 1, 2, \ldots (n-1) \text{ and } \bar{b}^T \bar{Y} < 0 \tag{7.27}$$

Note let

$$Y = \left(a_n^T \bar{Y}\right) Y_1 - \left(a_n^T Y_1\right) \bar{Y}. \tag{7.28}$$

From (7.23) (7.24) and (7.27), we then have

$$a_j^T Y = \left(a_n^T \bar{Y}\right) a_j^T Y_1 - \left(a_n^T Y_1\right) a_j^T \bar{Y}.$$

$$= \left[a_n^T \left(a_j^T Y_1\right) - \left(a_n^T Y_1\right) a_j^T\right] \bar{Y}$$

$$= a_j^T \bar{Y} \geq 0 \quad \text{for } j = 1, 2, \ldots (n-1)$$

$$a_n^T Y = 0$$

and $$b^T Y = \left[a_n^T \left(b^T Y_1\right) - \left(a_n^T Y_1\right) b^T\right] \bar{Y}$$

$$= \bar{b}\bar{Y} < 0.$$

and thus Y satisfies system 2 and the theorem is proved.

Theorem 7.10. (Gale [183]).

Let A be an m × n matrix and b be an m-vector. Then exactly one of the following two systems has a solution.

System 1. $AX \geq b, X \in R^n$

System 2. $A^T Y = 0, b^T Y = 1, Y \geq 0, Y \in R^m$

Proof: It can be easily verified that both System 1 and System 2 cannot have solutions simultaneously.

Suppose that the system of equations

$$A^T Y = 0, b^T Y = 1$$

has no nonnegative solution, that is,

$$\begin{pmatrix} A^T \\ b^T \end{pmatrix} Y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

has no nonnegative solution.

Then by theorem 7.9, there exists a vector $X \in R^n$ and a number η such that

$$(A \;\; b) \begin{pmatrix} X \\ \eta \end{pmatrix} \geq 0$$

$$(0 \;\; 1) \begin{pmatrix} X \\ \eta \end{pmatrix} < 0$$

that is, there exists and $X \in R^n$ and a number η which satisfy the system

$$AX + b\eta > 0$$

$$\eta < 0$$

which shows that, $-X/\eta$ is a solution of System 1.

Theorem 7.11. (Gale [183]) Let A be an m × n matrix and b be an m-vector. Then exactly one of the following two systems has a solution.

System 1. $AX \leq b, X \geq 0, X \in R^n$

System 2. $A^TY \geq 0, b^TY < 0, Y \geq 0, Y \in R^m$

Proof: As above, it is easy to see that both System 1 and System 2 cannot have solutions simultaneously.

Now, suppose that System 1 has no solution and since System 1 is equivalent to the system

1': $AX + IZ = b.$

$X, Z \geq 0, X \in R^n, Z \in R^m$

where I is a unit matrix of order m, System 1' has no solution.

By theorem 7.9, then, there exists $Y \in R^m$ such that

$$A^TY \geq 0$$
$$IY \geq 0$$
$$b^TY < 0.$$

Thus there exists $Y \geq 0$ which satisfies System 2.

Theorem 7.12. Let A be an m × n matrix B be an p × n matrix and C be an n-vector. Then exactly one of the following two systems has a solution.

System 1: $AX \geq 0, BX = 0, C^TX < 0, X \in R^n$

System 2: $A^TY + B^TZ = C, Y \geq 0, Y \in R^m, Z \in R^p.$

Proof: Writing $Z = Z_1 - Z_2, Z_1 \geq 0, Z_2 \geq 0$, System 2 of the theorem becomes

$$\left[A^T, B^T, -B^T\right]\begin{bmatrix} Y \\ Z_1 \\ Z_2 \end{bmatrix} = C, \quad Y, Z_1 Z_2 \geq 0$$

which is in the form System 1 of the theorem 7.9 and the result then follows.

Theorem 7.13. (Gordan's theorem [209]).

Let A be an m x n matrix. Then exactly one of the following two systems has a solution.

System 1: $AX = 0, X \geq 0, X \neq 0, X \in R^n$

System 2: $A^TY > 0, Y \in R^m$

Proof: Both System 1 and System 2 cannot have solutions simultaneously. For if X and Y are their solutions we have

from System 1, $Y^TAX = 0$

and from System 2, $X^TA^TY > 0$

which is a contradiction.

Now, suppose that system 1 has no solution. This is equivalent to the statement that the equations

$$AX = 0$$

$$e^T X = 1$$

where e is a vector of ones have no nonnegative solution.

Then by theorem 7.9 there exists a vector Y and a number η such that

$$\begin{bmatrix} A^T & e \end{bmatrix} \begin{bmatrix} Y \\ \eta \end{bmatrix} \geq 0, \quad \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} Y \\ \eta \end{bmatrix} < 0$$

and thus $A^T Y > 0$ and hence System 2 has a solution.

Theorem 7.14. (Stiemke's theorem [443])

Let A be an $m \times n$ matrix. Then exactly one, of the following two systems has a solution.

System 1: $\quad AX = 0, X > 0, \quad X \in R^n$

System 2: $\quad A^T Y \geq 0, Y \neq 0 \quad Y \in R^m$

Proof: Follows from theorem 7.13.

Theorem 7.15. (Motzkin's theorem [356]).

Let A, B and C be $m \times n$, $p \times n$ and $q \times n$ matrices respectively. Then exactly one of the following two systems has a solution.

System 1. $\quad AX > 0, BX \geq 0, CX = 0, \quad X \in R^n$

System 2. $\quad A^T Y_1 + B^T Y_2 + C^T Y_3 = 0$

$\quad Y_1 \geq 0, Y_1 \neq 0, Y_2 \geq 0, \quad Y_1 \in R^m, Y_2 \in R^p, Y_3 \in R^q$

Proof: Let System 1 has a solution and if possible let System 2 has also a solution. If X and Y_1, Y_2, Y_3 be their solutions then since $Y_1 \geq 0$, $Y_1 \neq 0$, $Y_2 \geq 0$, we would have from System 1,

$$X^T A^T Y_1 > 0, X^T B^T Y_2 \geq 0, X^T C^T Y_3 = 0$$

and therefore

$$X^T(A^T Y_1 + B^T Y_2 + C^T Y_3) > 0$$

which contradicts the first equality of System 2. Hence if System 1 has a solution, System 2 cannot have a solution.

Now suppose that System 1 has no solution. Then there is no solution of the system

$$AX \geq e\eta$$
$$BX \geq 0$$
$$CX = 0$$

where e is a vector of ones and $\eta \in R^1$, $\eta > 0$

That is, the system

$$\begin{pmatrix} A & -e \\ B & 0 \end{pmatrix} \begin{pmatrix} X \\ \eta \end{pmatrix} \geq 0, \quad \begin{pmatrix} C & 0 \end{pmatrix} \begin{pmatrix} X \\ \eta \end{pmatrix} = 0 \quad (0, -1) \begin{pmatrix} X \\ \eta \end{pmatrix} < 0$$

has no solution. Hence by theorem 7, 12, there exists a solution of the system

$$\begin{pmatrix} A^T & B^T \\ -e^T & 0 \end{pmatrix}\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} + \begin{pmatrix} C^T \\ 0 \end{pmatrix} Y_3 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

$$Y_1 \geq 0, Y_2 \geq 0$$

or
$$\begin{aligned} A^T Y_1 + B^T Y_2 + C^T Y_3 &= 0 \\ -e^T Y_1 \qquad\qquad &= -1 \\ Y_1 \geq 0, Y_2 \geq 0 & \end{aligned}$$

has a solution. Since $e^T Y_1 = 1$, $Y_1 \geq 0$, $Y_1 \neq 0$. Thus system 2 has a solution.

Theorem 7.16. (Tucker's theorem [473]).

Let G, H and K be given matrices of order $p \times n$, $q \times n$ and $r \times n$ respectively with G nonvacuous. Then exactly one of the following two systems has a solution.

System 1. $GX \geq 0, GX \neq 0, HX \geq 0, KX = 0, X \in R^r$

System 2. $G^T Y_2 + H^T Y_3 + K^T Y_4 = 0,$

$Y_2 > 0, Y_3 \geq 0, Y_2 \in R^p, Y_3 \in R^q, Y_4 \in R^n$

The proof is similar to the proof of theorem 7.15 (Motzkin's theorem)

We now consider an important property of the linear homogeneous inequlity system where the matrix of the system is skew-symmetric.

Theorem 7.17. the system of inequalities

$$\begin{aligned} KX &\geq 0 \\ X &\geq 0 \end{aligned} \tag{7.29}$$

where K is an $n \times n$ skew-symmetric matrix (i.e. $K^T = -K$) has at least one solution $\bar{X}$ such that

$$K\bar{X} + \bar{X} > 0$$

Proof: We first show that there always exists a solution $\bar{X}^i$ to (7.29) such that for K_i, the ith row of matrix K,

$$K_i \bar{X}^i + \bar{X}^i_i > 0 \tag{7.30}$$

Let us take $A = -K$, and

$b = -e_i = (0,0,..0, -1,0,....0)^T$

and apply theorem (7.11).

Then either

$$\begin{aligned} -KX &\leq -e_i \\ X &\geq 0 \end{aligned} \tag{7.31}$$

has a solution.

or,
$$\begin{aligned} -K^T Z &\geq 0 \\ -e^T_i Z &< 0 \\ Z &\geq 0 \end{aligned} \tag{7.32}$$

has a solution.

If $\bar{X}^i \geq 0$ is a solution to (7.31), we have

$$-K\bar{X}^i \leq -e_i$$

and thus

$$K\bar{X}^i \geq 0 \tag{7.32}$$

$$K_i\bar{X}^i \geq 1 > 0$$

and if $\bar{X}^i \geq 0$ satisfies (7.32), we have

$$-K^T\bar{X}^i \geq 0$$

$$-\bar{X}_i^i < 0$$

i.e.,

$$K\bar{X}^i \geq 0, \quad \text{since } K^T = -K \tag{7.34}$$

$$\bar{X}_i^i > 0$$

The result holds true for every row of matrix K, if the vector b is taken as the vectors $-e_1, -e_2, \ldots, -e_n$ successively. The vector

$$\bar{X} = \sum_{i=1}^{n} \bar{X}^i$$

is then a solution of (7.29) such that

$$K_i\bar{X} + \bar{X}_i \geq K_i\bar{X}^i + \bar{X}_i^i > 0, \qquad i = 1, 2,..n.$$

CHAPTER 8

Convex Sets and Convex Cones

8.1. Introduction and Preliminary Definitions

The concept of convexity, which we now introduce is of great importance in the study of optimization problems. Before we define convex sets and drive their properties, we first give some definitions.

Line: Let $X_1, X_2 \in R^n$, $X_1 \neq X_2$. The line passing through X_1 and X_2 is defined to be the set,

$$\{X \mid X = \lambda X_1 + (1 - \lambda)X_2, \lambda \in R^1\}.$$

Line Segment: The line segment, joining points $X_1, X_2 \in R^n$ is defined to be the set

$$\{X \mid X = \lambda X_1 + (1 - \lambda)X_2, 0 \leq \lambda \leq 1\} : \text{Closed}$$

$$\{X \mid X = \lambda X_1 + (1 - \lambda)X_2, 0 < \lambda < 1\} : \text{Open}$$

Half Line: The set $D = \{X \mid X = \lambda d, \lambda \geq 0\}$ where d is a nonzero vector in R^n, is called the half-line or ray starting from the origin.

Hyperplane: The set $H = \{X \mid a^TX = \alpha\}$ is said to be a hyperplane in R^n, where a is a nonzero vector in R^n and α is a scalar. It passes through the origin if and only if $\alpha = 0$. The nonzero vector a is usually referred to as the normal to the hyperplane and if the vector a is of unit length, it is called unit normal. Two hyperplanes are said to be parallel if they have the same unit normal.

Half-Spaces: A hyperplane H determines two closed half-spaces

$$H_1 = \{X \mid a^TX \leq \alpha\}, H_2 = \{X \mid a^TX \geq \alpha\}$$

and two open half-spaces

$$H_3 = \{X \mid a^TX < \alpha\}, H_4 = \{X \mid a^TX > \alpha\}.$$

The hyperplane $H = \{X \mid a^TX = \alpha\}$, is called the generating hyperplane of the half-spaces.

8.2. Convex Sets and their Properties

A set S in R^n is said to be convex if the line segment joining any two points in the set also belongs to the set. In other words, a set $S \in R^n$ is convex if for any two points $X_1, X_2 \in S$ and $\lambda \in R^1$,

$$X = \lambda X_1 + (1 - \lambda)X_2 \in S, \text{ for each } \lambda, 0 \leq \lambda \leq 1 \tag{8.1}$$

The expression $\lambda X_1 + (1 - \lambda)X_2$, $0 \leq \lambda \leq 1$ is referred to as a convex combination of X_1, X_2.

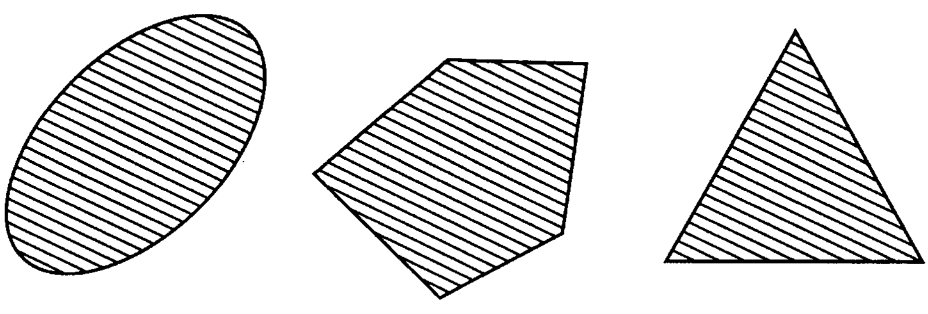
Linear subspaces, triangles and spheres are some simple examples of convex sets. In particular the empty set ϕ, sets with a single point only and R^n are convex.

Extreme points: A point X in a convex set S is called an extreme point or a vertex of S, if there exist no two distinct points $X_1, X_2 \in S$ such that

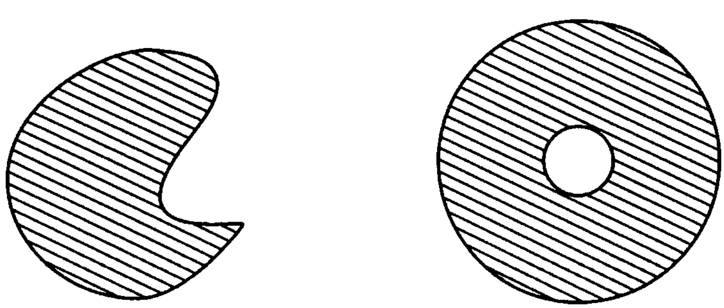
$$X = \lambda X_1 + (1 - \lambda)X_2, \text{ for } 0 < \lambda < 1 \tag{8.2}$$

Extreme Directions: Let S be a closed convex set in R^n. A nonzero vector d in R^n is called a direction of S, if for each $X \in S$, the ray $\{X + \lambda d : \lambda \geq 0\}$ emanating from X also belongs to S. Two directions d_1 and d_2 of S are said to be distinct if $d_1 \neq d_2$ for any $\alpha > 0$.

A direction d of S is called an extreme direction of S if d cannot be expressed as a positive linear combination of two distinct directions of S, that is, if $d = \lambda_1 d_1 + \lambda_2 d_2$, for $\lambda_1, \lambda_2 > 0$ then $d_1 = \alpha d_2$ for some $\alpha > 0$.



Convex



Nonconvex

Figure 8.1. convex and nonconvex sets.

Extending the concept of convex combination, we now generalize its definition.

Convex Combination: A vector X in R^n is said to be a convex combination of the vectors $X_1, X_2, .. X_m \in R^n$ if there exist real numbers $\lambda_1, \lambda_2, .. \lambda_m$ such that

$$X = \sum_{i=1}^{m} \lambda_i X_i, \quad \lambda i \geq 0, \quad i = 1,2,..m, \ \sum_{i=1}^{m} \lambda_i = 1 \tag{8.3}$$

Theorem 8.1: A necessary and sufficient condition for a set S in R^n to be convex is that every convex combination of any m points in S belongs to S.

In other words, a set S in R^n is convex if and only if for any m,

$$X_i \in S, \ i = 1,2,..m.$$

$$\lambda_i \geq 0, \ i = 1,2,..m.$$

$$\sum_{i=1}^{m} \lambda_i = 1$$

implies that $\sum_{i=1}^{m} \lambda_i X_i \in S.$

Proof: Suppose that every convex combination of m points in S belongs to S. This implies that every convex combination of two points in S belongs to S which by definition implies that the set S is convex. Hence the condition is sufficient.

The necessity of the condition will be proved by the method of induction. Suppose that S is convex. For m = 1, the condition is trivially true. For m = 2, the condition holds by definition. Assume now that the condition holds for m = k and we will show that it also holds for m = k + 1.

Let $X = \sum_{i=1}^{k+1} \lambda_i X_i, \ X_i \in S, \ \lambda_i \geq 0, \ i = 1..k+1, \ \sum_{i=1}^{k+1} \lambda_i = 1$

If $\lambda_{k+1} = 0$, than $X = \sum_{i=1}^{k} \lambda_i X_i \in S$, by assumption.

If $\lambda_{k+1} = 1, \quad X = X_{k+1} \in S,$

If $0 < \lambda_{k+1} < 1$, we have

$$X = (1 - \lambda_{k+1}) \sum_{i=1}^{k} \frac{\lambda_i}{1 - \lambda_{k+1}} X_i + \lambda_{k+1} . X_{k+1}$$

$$X = (1 - \lambda_{k+1}) \bar{X} + \lambda_{k+1} X_{k+1}.$$

Since by assumption, the condition holds for k points, $\bar{X}$ is a point of S and consequently X is also a point of S.

This completes the proof.

Theorem 8.2: The intersection of a finite or infinite number of convex sets is convex.

Proof: Let I be a finite or infinite set of indices and S_i, $i \in I$ are convex sets.

$$\text{Let } X_1, X_2 \in \bigcap_{i \in I} S_i.$$

Since X_1, X_2 belong to each S_i, and S_i are convex, $X = \lambda X_1 + (1 - \lambda)X_2$, $0 \leq \lambda \leq 1$, belongs to each S_i and therefore belongs to

$$\bigcap_{i \in I} S_i$$

Hence $\bigcap_{i \in I} S_i$ is convex.

Theorem 8.3: Let S be a convex set in R^n. Then the product αS, where α is a real number, is a convex set.

Proof: The product αS is defined by

$$\alpha S = \{Z \mid Z = \alpha X, \alpha \in R^1, X \in S\}.$$

Let $Z_1, Z_2 \in \alpha S$, then $Z_1 = \alpha X_1$, $Z_2 = \alpha X_2$, $X_1, X_2 \in S$

Then for $0 \leq \lambda \leq 1$, we have,

$\lambda Z_1 + (1 - \lambda)Z_2 = \alpha [\lambda X_1 + (1 - \lambda)X_2] \in \alpha S$, since S is convex. Hence αS is convex.

Theorem 8.4: The sum $S = S_1 + S_2$ of two convex sets S_1, S_2 in R^n is a convex set.

Proof: The sum $S = S_1 + S_2$ is defined by,

$$S = S_1 + S_2 = \{Z \mid Z = X + Y, X \in S_1, Y \in S_2\}$$

Now, let $Z_1, Z_2 \in S$, then $Z_1 = X_1 + Y_1$, $Z_2 = X_2 + Y_2$, $X_1, X_2, \in S_1$; $Y_1, Y_2 \in S_2$.

For $0 \leq \lambda \leq 1$, we then have,

$$\begin{aligned}\lambda Z_1 + (1 - \lambda)Z_2 &= \lambda (X_1 + Y_1) + (1 - \lambda)(X_2 + Y_2)\\ &= [\lambda X_1 + (1 - \lambda)X_2] + [\lambda Y_1 + (1 - \lambda)Y_2]\\ &\in S_1 + S_2 = S\end{aligned}$$

Hence $S = S_1 + S_2$ is convex.

Similarly, it can be shown that the set $S_1 - S_2$ is convex.

8.3. Convex Hulls

Let S be an arbitrary set in R^n. The convex hull of S, denoted by [S], is the intersection of all convex sets in R^n containing S.

Since the intersection of convex sets is convex, [S] is convex and is the smallest convex set containing S. Obviously, if S is convex, then [S] = S.

Theorem 8.5: The convex hull [S] of a set S in R^n is the set of all convex combinations of points of S.

Proof: Let K be the set of all convex combinations of points of S. Then

$$K = \left\{X \mid X = \sum_{i=1}^{m} \alpha_i X_i,\ X_i \in S,\ \alpha_i \in R^1,\ \alpha_i \geq 0,\ i = 1, 2, ..m, \sum_{i=1}^{m} \alpha_i = 1\right\}$$

where m is an arbitrary positive integer.

Let $X_1, X_2 \in K$, then

$$X_1 = \sum_{i=1}^{m} \alpha_i X_i^1, \quad X_i^1 \in S, \quad \alpha_i \geq 0, \ \sum_{i=1}^{m} \alpha_i = 1$$

$$X_2 = \sum_{i=1}^{m} \beta_i X_i^2, \quad X_i^2 \in S, \quad \beta_i \geq 0, \ \sum_{i=1}^{m} \beta_i = 1$$

For $0 \leq \lambda \leq 1$, we then have

$$\lambda X_1 + (1-\lambda) X_2 = \sum_{i=1}^{m} \lambda \alpha_i X_i^1 + \sum_{i=1}^{m} (1-\lambda) \beta_i X_i^2$$

where $\lambda\alpha_i \geq 0, \quad (1-\lambda)\beta i \geq 0, \ i = 1, 2, \ldots m$

and $\sum_{i=1}^{m} \lambda \alpha_i + \sum_{i=1}^{m} (1-\lambda) \beta_i = 1$

Hence $\lambda X_1 + (1-\lambda)X_2 \in K$ and K is a convex set.

Obviously, $S \subset K$ and since K is convex $[S] \subset K$. Now, since [S] is a convex set containing S, we have by theorem 8.1 that it must also contain all convex combinations of points of S. Hence $[S] \supset K$. Thus $[S] = K$.

The theorem, therefore shows that a point in the convex hull of a set in R^n, can be expressed as a convex combination of a finite number of points in the set. The following theorem shows that it is not really necessary to form convex combinations involving more than (n + 1) points in S.

Theorem 8.6. (Caratheodory's Theorem [64])

Let S be an arbitrary set in R^n. Then every point of the convex hull [S], can be expressed as a convex combination of atmost (n + 1) points of S.

Proof: Let $X \in [S]$, then

$$X = \sum_{i=1}^{m} \lambda_i X_i, \text{ where } X_i \in S, \ \lambda_i > 0, \ i = 1, 2, .. m \sum_{i=1}^{m} \lambda_i = 1 \tag{8.4}$$

If $m \leq n + 1$, the theorem is true.

Now, suppose that $m > n + 1$. Since $m - 1 > n$, the vectors $X_2 - X_1$, $X_3 - X_1$, .. $X_m - X_1$ are linearly dependent. Hence there exists scalars μ_i (i = 2, ... m) not all zero, such that

$$\sum_{i=2}^{m} \mu_i (X_i - X_1) = 0$$

Let $\mu_1 = -\sum_{i=2}^{m} \mu_i$ and it follows that

$$\sum_{i=1}^{m} \mu_i X_i = 0, \qquad \sum_{i=1}^{m} \mu_i = 0 \text{ and not all } \mu_i = 0 \tag{8.5}$$

which means that for at least one i, $\mu_i > 0$.

Let θ be a real number such that,

$$\lambda_i - \theta\mu_i \geq 0,\ i = 1,2 \,..\, m \text{ and}$$
$$\lambda_i - \theta\mu_i = 0 \text{ for some i, say } i_0. \tag{8.6}$$

This can be achieved if we chose θ as

$$\theta = \text{Min}\left\{\frac{\lambda_i}{\mu_i}, \mu_i > 0\right\} = \frac{\lambda_{i0}}{\mu_{i0}}, \quad \text{for some } i = i_0.$$

and then $\quad \theta > 0. \qquad (8.7)$

If $\mu_i \leq 0$, $\lambda_i - \theta\,\mu_i > 0$ and if $\mu_i > 0$, then

$$\frac{\lambda_i}{\mu_i} \geq \frac{\lambda_{i0}}{\mu_{i0}} = \theta \text{ and hence } \lambda_i - \theta\mu_i \geq 0, \text{ for } i = 1, 2,..m. \tag{8.8}$$

By (8.5), we then have

$$X = \sum_{i=1}^{m} \lambda_i X_i = \sum_{i=1}^{m} \lambda_i X_i - \theta \sum_{i=1}^{m} \mu_i X_i = \sum_{i=1}^{m} (\lambda_i - \theta\mu_i) X_i$$

where $\quad \lambda_i - \theta\mu_i \geq 0, \quad i = 1, 2,..m$

$$\sum_{i=1}^{m} (\lambda_i - \theta\mu_i) = 1 \qquad \text{and further} \tag{8.9}$$

$$\lambda_i - \theta\mu_i = 0 \qquad \text{for } i = i_0, \quad i_0 \in \{1, 2,..m\}.$$

It implies that X is expressed as a convex combination of atmost (m – 1) points of S. If m – 1 > n + 1, the above argument is applied again to express X as a convex combination of (m – 2) points of S. The process is repeated until $X \in [S]$ is expressed as a convex combination of (n + 1) points of S.

8.4. Separation and Support of Convex Sets

The notions of supporting hyperplanes and separation of disjoint convex sets are very important for a wide range of optimization problems. It can be easily visualized that if we have two disjoint convex sets then there is a hyperplane (called a separating hyperplane) such that one set lies entirely on one side of the hyperplane and the second set on the other side. Also, extending the geometric concept of tangency, the generalized tangency to a convex set S is expressed by a hyperplane H (called a supporting hyperplane) where S is contained in one of the half-spaces of H and the boundary of S has a point in common with H.

Separation of a Convex Set and a Point

Theorem 8.7: Let S be a nonempty closed convex set in R^n and Y be a point not in S. Then there exists a hyperplane H, called a separating hyperplane which contains Y such that S is contained in one of the open half-spaces produced by H.

For the proof of the theorem, we need the following important result from calculus.

Lemma 8.1: Let S be a closed convex set in R^n and $Y \notin S$. Then there exists a unique point $X_0 \in S$, with minimum distance from Y, that is,

$$\|X_0 - Y\| \leq \|X - Y\|, \qquad \text{for all } X \in S. \tag{8.10}$$

Proof of the theorem: Let S be a nonempty closed convex set and $Y \notin S$. Hence by lemma 8.1, there exists a unique point $W \in S$, closest to Y, so that

$$\|W - Y\| \leq \|X - Y\|, \qquad \text{for all } X \in S. \tag{8.11}$$

Since for any $X \in S$, the point $Z = \lambda X + (1 - \lambda)W$, $0 \leq \lambda \leq 1$ is also in S, we have,

$$\|W - Y\| \leq \|Z - Y\|$$

or $$\|W - Y\| \leq \|\lambda X + (1 - \lambda) W - Y\|,$$

or $$\|(W - Y) + \lambda (X - W)\| \geq \|W - Y\|.$$

We then have

$$\lambda^2(X - W)^T(X - W) + (W - Y)^T(W - Y) + 2\lambda (W - Y)^T(X - W) \geq (W - Y)^T(W - Y)$$

or $$\lambda^2(X - W)^T(X - W) + 2\lambda (W - Y)^T(X - W) \geq 0 \tag{8.12}$$

Taking $\lambda > 0$ and dividing by λ, we have

$$\lambda (X - W)^T(X - W) + 2 (W - Y)^T(X - W) \geq 0$$

Now, let $\lambda \to 0$ and then in the limit

$$(W - Y)^T(X - W) \geq 0 \tag{8.13}$$

Since Y is a given point not in S and W is a unique point in S, $C = W - Y$ is a constant vector

and thus $$C^T(X - W) \geq 0 \text{ or } C^TX \geq C^TW \tag{8.14}$$

Moreover, $$C^T(W - Y) = C^TC > 0 \text{ or } C^TW > C^TY \tag{8.15}$$

Hence, $$C^TX \geq C^TW > C^TY = \alpha \text{ (say)}$$

Thus for any point $X \in S$, $C^TX > \alpha$.,

which means that the convex set S lies wholly in the open half space $C^TX > \alpha$, produced by the hyperplane $C^TX = \alpha$.

The geometrical interpretation of the theorem in R^2 or R^3 is simple. The hyperplane $C^TX = \alpha$ is the plane in R^3 or the line in R^2 through $Y \notin S$ perpendicular to the line joining Y and $W \in S$, representing the shortest distance from Y to S (Fig 8.2)

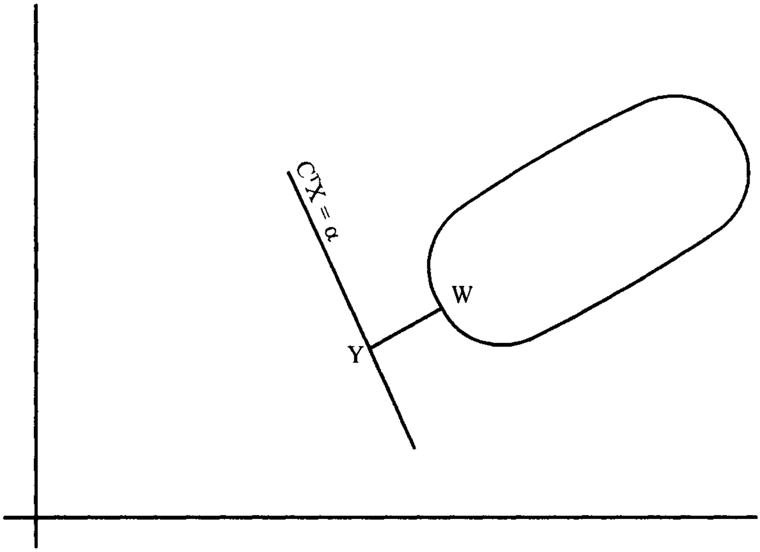


Figure 8.2

Let us now imagine Y to approach W through a sequence of points along the line joining the points Y and W. In the limit when Y = W, the hyperplane $C^TX = \alpha$ meets S at W and for all $X \in S$, $C^TX \geq \alpha$.

It is clear that W is a boundary point of S. For otherwise, there is a neighbourhood which is entirely in S and W being in the hyxerplane $C^TX = \alpha$, there will be points in S lying on either side of the hyperplane. This will contradict the theorem proved above.

Supporting Hyperplane

Let S be a nonempty set in R^n and let W be a boundary point of S. A hyperplane $H = \{X \mid C^TX = \alpha\}$ is called a supporting hyperplane of S at W if $C^TW = \alpha$ and S is contained either in the halfspace $C^TX \leq \alpha$ or in the halfspace $C^TX \geq \alpha$.

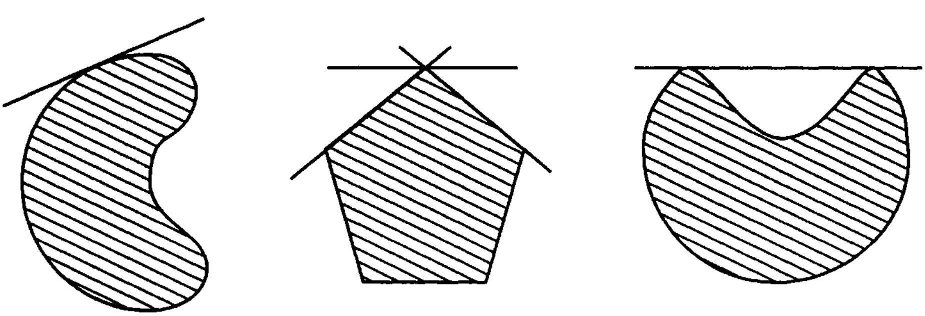


Figure 8.3. Supporting hyperplanes.

From the discussion in theorem 8.7, it can be seen that the hyperplane which meets S at its boundary point W is a supporting hyperplane of S. W however is not an arbitrary boundary point. It is determined by the choice of Y. The following theorem shows that there exists a supporting hyperplane of a convex set at each of its boundary point.

Theorem 8.8. Let S be a closed convex set in R^n and let W be a boundary point of S. Then there exists at least one supporting hyperplane of S at W.

Proof: Let W be any boundary point of S. Let Y_k be a point not in S but in the ϵ_k ($\epsilon_k = 1/k$ say) neighbourhood of W. By theorem 8.7, there is a boundary point W_k of S whose distance from Y_k is minimal and there exists a supporting hyperplane $C^T_kX = \alpha_k$ at W_k.

Now, $$\|W_k - W\| = \|W_k - Y_k + Y_k - W\|$$
$$\leq \|W_k - Y_k\| + \|Y_k - W\| \tag{8.16}$$

Since $\|Y_k - W\| < \epsilon_k$, $\|Y_k - W\| \to 0$ as $k \to \infty$

Further, since $\|W_k - Y_k\|$ is the shortest distance between Y_k and S
$$\|W_k - Y_k\| \leq \|Y_k - W\| \to 0 \text{ as } k \to \infty$$

Hence, $$\|W_k - W\| \to 0 \text{ as } k \to \infty$$

or $$W_k \to W.$$

Now, dividing both sides of the supporting hyperplane $C^T_kX = \alpha_k$, at W_k, by the length of C_k, we get

$$\bar{C}_k^T X = \bar{\alpha}_k, \text{ where } \bar{C}_k = \frac{C_k}{\|C_k\|} \text{ and } \bar{\alpha}_k = \frac{\alpha_k}{\|C_k\|}, \tag{8.17}$$

so that $\|\bar{C}_k\| = 1$.

Since $\{\bar{C}_k\}$ is bounded, it has a convergent subsequence with limit C, whose norm is equal to one.

Again by Schwarz inequality, since W_k is on the hyperplane,

$$|\bar{\alpha}_k| = |\bar{C}_k^T W_k| \leq \|\bar{C}_k\| \; \|W_k\| = \|W_k\| \tag{8.18}$$

Hence $\{\bar{\alpha}_k\}$ from a bounded infinite sequence and has a limit point α.

Hence $\bar{C}_k \to C$, $\bar{\alpha}_k \to \alpha$ as $\bar{W}_k \to W$. $\quad$ (8.19)

Thus there exists a supporting hyperplane

$$C^TX = \alpha \text{ at } W \in S \text{ and for all } X \in S,$$
$$C^TX \geq \alpha. \tag{8.20}$$

Since this is true for every point in the neighbourhood of W and not in S, the result follows.

It should be noted that the theorem holds even if the convex set S is not closed.

Since $\bar{S}$, the closure of S is convex and W is a boundary point of $\bar{S}$, the theorem holds for $\bar{S}$ and then it certainly holds for S. Thus without loss of generality, we may assume that S is closed.

Corollary 8.1: If S is a closed and bounded convex set, the linear function C^TX assumes its maximum or minimum value on the boundary of S.

Proof: Excercise.

Separation of two sets

A hyperplane $H = \{X \mid C^TX = \alpha, X \in R^n\}$ is said to separate two nonempty sets S_1 and S_2 in R^n if

$$C^TX \geq \alpha, \text{ for each } X \in S_1$$
$$\text{and} \quad C^TX \leq \alpha, \text{ for each } X \in S_2$$

The hyperplane H is said to strictly separate S_1 and S_2 if $C^TX > \alpha$ for each $X \in S_1$ and $C^TX < \alpha$ for each $X \in S_2$

The hyperplane H is then called a separating hyperplane.

We now prove some results when the sets S_1 and S_2 are convex.

Lemma 8.2. Let S be a nonempty convex set in R^n, not containing the origin 0. Then there exists a hyperplane $\{X \mid C^TX = 0\}$ $c \pm 0$ which separates S and the origin 0.

Proof: (i) If the origin is an exterior point of $\bar{S}$, the closure of S, then by theorem 8.7, there exists a vector $C \neq 0$ such that $C^TX > 0$ for all $X \in S$.

Thus, the hyperplane $\{X \mid C^TX = 0\}$ separates S and the origin.

(ii) If the origin is a boundary point of $\bar{S}$, then the lemma follows from theorem 8.8,

Theorem 8.9 (Separation theorem). Let S_1 and S_2 be two nonempty disjoint convex sets in R^n. Then there exists a hyperplane $H = \{X \mid C^TX = \alpha\}$, which separates S_1 and S_2, that is there exists a nonzero vector C in R^n such that

$$\inf \{C^TX \mid X \in S_1\} \geq \text{Sup } \{C^TX \mid X \in S_2\}.$$

Proof: The set $S = S_1 - S_2 = \{X_1 - X_2, X_1 \in S_1, X_2 \in S_2\}$ is convex by theorem 8.4 and since $S_1 \cap S_2 = \phi$, $0 \notin S$.

Hence by Lemma 8.2, there exists a nonzero vector C such that

$$C^TX \geq 0 \text{ , for all } X \in S.$$

Thus for all $X_1 \in S_1$ and $X_2 \in S_2$

$$C^T(X_1 - X_2) \geq 0$$
$$\text{or} \quad C^TX_1 \geq C^TX_2.$$

which implies that

$$\inf \{C^TX \mid X \in S_1\} \geq \text{Sup } \{C^TX \mid X \in S_2\}.$$

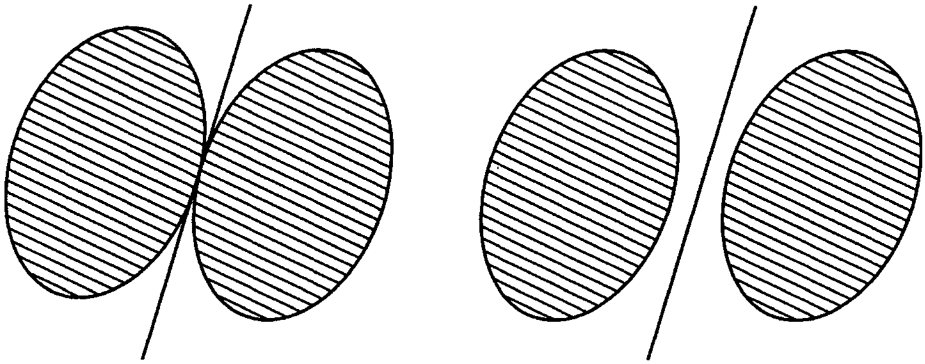


Figure 8.4. Separation of two convex sets.

Lemma 8.3. Let S be a nonempty closed convex set in R^n, not containing the origin 0. Then there exists a hyperplane $H = \{X \mid C^TX = \alpha, X \in R^n\}$, $\alpha > 0$ which strictly separates S and the origin.

Proof: Since S is closed and does not contain the origin 0, 0 is an exterior point. Hence by theorem 8.7, there exists a vector $C \neq 0$ such that

$$C^TX > 0 \text{ , for all } X \in S$$

Hence there is a positive number α such that

$$C^TX > \alpha > 0$$

The hyperplane $H = \{X \mid C^TX = \alpha\}$, then strictly separates S and the origin 0.

Theorem 8.10. Let S_1 and S_2 be two nonempty disjoint convex sets in R^n and suppose that S_1 is compact and S_2 is closed. Then there exists a hyperplane $H = \{X|C^TX = \alpha\}$, $C \neq 0$ that strictly separates S_1 and S_2.

Proof: Since the set $S = S_1 - S_2$ is convex by theorem 8.4 and since S_1 is compact and S_2 is closed it follows that S is a nonempty closed convex set. Then by Lemma 8.3. there exists a hyperplane $H = \{X \mid C^TX = \alpha\}$, $\alpha > 0$, which strictly separates S_1 and S_2.

Theorem 8.11. A nonempty closed convex set S in R^n, bounded from below (or from above) has atleast one extreme point.

Proof: The theorem will be proved by induction on the number of dimension n.

Since S is closed and bounded from below, for n = 1,

$S = \{X \mid X \geq a, a \in R^1\}$. Obviously, a is an extreme point of S and the theorem is true for n = 1.

Let us now assume that the theorem is true for n = m. We shall then prove that it is true for n = m + 1. Let $S \subset R^{m+1}$. Since S is closed and bounded from below, there exists a boundary point W of S and by theorem 8.8. there is a supporting hyperplane $H = \{X \mid C^TX = \alpha\}$ of S at W. Let T* be the intersection of S and H. T* is then a nonempty, closed convex set bounded from below and is of dimension m. Therefore by hypothesis, $T^* = S \cap H$ has an extreme point. We contend that the extreme point of T* is also an extreme point of S. To prove our contention, let us suppose that $\bar{X}$ is an extreme point of T* which is not an extreme point of S.

Then $\qquad \bar{X} = \lambda X_1 + (1 - \lambda)X_2,\ X_1, X_2 \in S,\ 0 < \lambda < 1 \qquad (8.21)$

Since by (8.20), $\qquad C^T X_1 \geq \alpha,\ C^T X_2 \geq \alpha$

and since $\qquad C^T \bar{X} = \alpha$, we have

$$\alpha = C^T \bar{X} = \lambda C^T X_1 + (1 - \lambda) C^T X_2 \qquad (8.22)$$

and this will hold if and only if

$$C^T X_1 = \alpha,\ C^T X_2 = \alpha$$

that is, if and only if $X_1, X_2 \in T^*$ which contradicts that $\bar{X}$ is an extreme point of T*. Hence the extreme point of T* is also an extreme point of S.

Thus, if the theorem holds for $n = m$, it holds for $n = m + 1$ and we have seen that it holds for $n = 1$. Hence by induction, the theorem is true for all $n \geq 1$.

Corollary 8.2: Every supporting hyperplane of a nonempty compact convex set S contains atleast one extreme point of S.

Theorem 8.12. Every point of a nonempty closed bounded convex set S in R^n is a convex combination of its extreme points.

Proof: The proof will again be given by induction on the number of dimension n.

If $n = 1$, S is a closed and bounded interval [a, b] whose extreme points are a and b and every point of which is a convex combination of a and b, hence the theorem holds for $n = 1$.

Suppose that the theorem holds for $n = m$ and we shall prove that it also holds for $n = m + 1$. Let $S \subset R^{m+1}$ and X_0 be any point in S. Now, (a) X_0 may be a boundary point of S or (b) an interior point.

Case (a): Let X_0 be a boundary point of S and $H = \{X \mid C^T X = \alpha\}$ be a supporting hyperplane at X_0. The set $S \cap H$ is then a nonempty, closed, bounded convex set of dimension m. By hypothesis therefore, every point of $S \cap H$ is a convex combination of its extreme points. But as shown in the proof of theorem 8.11, an extreme point of $S \cap H$ is also an extreme point of S. Hence X_0 is a convex combination of the extreme points of S.

Case (b): Let X_0 be an interior point of S. Then any line through X_0 intersects S in a line segment with boundary points X_1 and X_2 and therefore X_0 can be expressed as a convex combination of X_1 and X_2.

Now, since X_1, X_2 are also boundary points of S, by case (a), they are convex combinations of the extreme points of S. Hence X_0 is also a convex combination of the extreme points of S.

This completes the proof.

8.5. Convex Polytopes and Polyhedra

A set in R^n which is the intersection of a finite number of closed half-spaces in R^n is called a polytope.

From the convexity of the half-spaces, it follows that a polytope is a closed convex set.

If a polytope is bounded, it is called a convex polyhedron.

Now, if the set of the intersection of a finite number of closed half-spaces is bounded, it is equal to the set of linear convex combinations of a finite number of points, known as the convex hull of these points.

Therefore, a convex polyhedron may also be defined as the convex hull of a finite number of points.

Theorem 8.13. The set of extreme points of a convex polyhedron is a subset of the set of its spanning points.

Proof: Let $K = \{X_1, X_2, .. X_m\}$ be the set of points spanning the convex polyhedron S and V be the set of its extreme points.

Suppose that $X \in V$ but $X \notin K$. Since $X \in S$, it is a convex combination of points in K and thus

$$X = \sum_{i=1}^{m} \mu_i X_i, \quad \mu_i \geq 0, \quad i = 1,2,..m, \sum_{i=1}^{m} \mu_i = 1$$

If $X \notin K$, at least two μ_i say μ_1, and μ_2 are nonzero and we may write

$$X = \mu_i X_i + (1-\mu_1)\bar{X}, \quad 0 < \mu_1 < 1, \quad \bar{X} = \sum_{i=1}^{m} \frac{\mu_i}{1-\mu_1} X_i \in S$$

which contradicts that X is an extreme point

Hence $X \in K$. Thus $V \subset K$

Theorem 8.14. Let T be a polytope in R^n defined by the intersection of the hyperspaces

$$a_i x \leq b_i, \quad i = 1, 2, \dots m; \quad m \geq n \tag{8.23}$$

A point $\bar{x} \in T$ is an extreme point of T if and only if $\bar{x}$ is a solution of the system.

$$a_i x = b_i, \quad i \in \{1, 2, \dots m\}. \tag{8.24}$$

where a_i is a row vector and the matrix of the system is a nonsingular matrix of order n.

Proof: Let the matrix of the system (8.24) be written in the matrix form as

$$\bar{A}X = \bar{b}$$

Let X ba a solution of (8.25) and let us suppose that there exist two distinct point X_1 and X_2 in T such that

$$X = \lambda X_1 + (1-\lambda)X_2, \quad 0 < \lambda < 1 \tag{8.26}$$

Since $\bar{A}X_1 \leq \bar{b}$, $\bar{A}X_2 \leq \bar{b}$, we have

$$\bar{b} = \bar{A}X = \lambda\bar{A}X_1(1-\lambda)\bar{A}X_2 \leq \bar{b}$$

Hence $\bar{A}X = \bar{A}X_1 = \bar{A}X_2 = \bar{b}$

abd subce $\bar{A}$ is nonsingular,

$$X = X_1 = X_2$$

which contradicts that X_1 and X_2 are distinct. Hence X is an extreme point of T.

Conversely, let $\bar{X}$ be an extreme point of T. $\bar{X}$ therefore satisfies the system

$$a_iX \leq b_i, \quad i = 2, 2, \dots m.$$

Suppose that

$$a_i\bar{X} < b_i, \text{ for all i, } i = 1, 2, \dots m$$

We can then always find a small nonzero vector of $\bar{X}$ such that

$$a_i(\bar{X} + d\bar{X}) < b_i \text{ and } a_i(\bar{X} - d\bar{X}) < b_i, \quad i = 1, 2, \dots m \tag{8.27}$$

Then $\bar{X} = \frac{1}{2}[(\bar{X} + d\bar{X}) + (\bar{X} - d\bar{X})]$, is a convex combination of two distinct points of T and hence $\bar{X}$ is not an extreme point.

Thus, there exists a nonempty set $I \subset \{1, 2, \dots m\}$, such that

$$a_i\bar{X} = b_i, \quad i \in I \tag{8.28}$$

$$a_i\bar{X} < b_i, \quad i \notin I \tag{8.29}$$

Writing (8.28) in matrix form as $A^I\bar{X} = b^I$, we now show that the columns of A^I are linearly independent. Suppose they are not. In that case, there exists a nonzero vector y such that

$$a_iY = 0, \quad Y \neq 0, i \notin I. \tag{8.30}$$

We can then find a small nonzero scalar $d\bar{X}$ such that

$$a_i(\bar{X} + Yd\bar{X}) < a_i(\bar{X} - Yd\bar{X}) < b_i, \quad i \notin I. \tag{8.31}$$

Thus from (8.28), (8.30) and (8.31), we have

$$a_i(\bar{X} + d\bar{X}) \leq b_i, \quad i = 1, 2, \dots m.$$

Then, $\bar{X} = \frac{1}{2}\left[(\bar{X} + Yd\bar{X}) + (\bar{X} - Yd\bar{X})\right]$, is a convex combination of two distinct points of T and hence $\bar{X}$ is not an extreme point of T, which contradicts the hypothesis.

The columns of A^I are therefore linearly independent; A^I contains a nonsingular submatrix $\bar{A}$ of order n. $\bar{X}$ is thus a solution of the system $\bar{A}X = \bar{b}$.

Corollary 8.3. The number of extreme points of a polytope is finite.

Proof: It follows if we note that the maximum possible number of nonsingular matrices from A is finite.

Theorem 8.15. Let S be a nonempty convex polytope in R^n, given by $S = \{X \mid AX = b, X \geq 0\}$, where A is an $m \times n$ matrix with rank m. Then, the set of extreme points of S is not empty and has a finite number of points. Further, the set of extreme directions of S is empty if and only if S is bounded (i.e. S is a convex polyhedron). If S is not bounded, then S has a finite number of extreme directions.

Let $X_1, X_2, \ldots X_k$ be the extreme points and the vectors $d_1, d_2, \ldots d_l$ be the extreme directions of S. Then $X \in S$ if and only if X can be written as

$$X = \sum_{j=1}^{k} \lambda_j X_j + \sum_{j=1}^{l} \mu_j d_j$$

$$\sum_{j=1}^{k} \lambda_j = 1$$

$$\lambda_j \geq 0, \qquad j = 1, 2, \ldots k$$
$$\mu_j \geq 0, \qquad j = 1, 2, \ldots l.$$

Proof: For a proof see [35].

Simplex

Let $X_1, X_2, \ldots X_{m+1}$ be a finite number of points in R^n. If $X_2 - X_1, X_3 - X_1, \ldots X_{m+1} - X_1$ are linearly independent then the convex hull of $X_1, X_2, \ldots X_{m+1}$, that is, the set of all convex combinations of $X_1, X_2, \ldots X_{m+1}$ is called an m-dimensional simplex (m-simplex) in R^n with vertices $X_1, X_2, \ldots X_{m+1}$. Since the maximum number of linearly independent vectors in R^n is n, there could be no simplex in R^n with more than (n + 1) vertices. Thus a 0-simplex is a point, a 1-simplex is a line segment, a 2-simplex is a triangle, a 3-simplex is a tetrahedron and for m > n, there is no m-simplex.

The Faces of Polytopes

Let T be a polytope in R^n

$$T = \{X \mid a_i X \leq b_i, i = 1,2..m\}$$

and let V_J denote the linear variety

$$a_i X = b_i, i \in J \subset \{1, 2, \ldots m\}.$$

We call F_J, the face associated with J, the intersection of the linear variety V_J with T.

If $V_J = \phi$, X is a vertex and we say that $\{X\}$ is a face of order 0, if V_J has dimension 1, the face F_J is an edge or a face of order 1 of T. If V_J has dimension k, we say that F_J is a face of order k of T.

Thus an edge of a convex polyhedron S is the line segment joining any two extreme points (vertices) of S, if it is the intersection of S with a supporting hyperplane of S.

The two extreme points of S are said to be adjacent extreme points, if the line segment joining them is an edge of S.

8.6. Convex Cones

In this section we introduce the concept of convex cones and briefly discuss their properties. The material in preceding sections is sufficient for understanding the subsequent chapters. This section may therefore be skipped without loss of continuity.

Cones: A subset C of R^n is called a cone with vertex zero, if

$$X \in C \Rightarrow \lambda X \in C, \text{ for all } \lambda \geq 0$$

The sum $C_1 + C_2$ of two cones defined by

$$C_1 + C_2 = \{X \mid X = X_1 + X_2,\ X_1 \in C_1, X_2 \in C_2\}$$

is also a cone.

Convex Cones: A cone is a convex cone if it is a convex set.

Equivalently, a convex cone C is a set of points such that

$$X \in C \Rightarrow \lambda X \in C, \text{ for all } \lambda \geq 0.$$

$$X_1 \in C, X_2 \in C \Rightarrow X_1 + X_2 \in C.$$

Linear subspaces, the nonnegative orthant $X \geq 0$, a half line $\{X \mid X = \lambda b, \lambda \geq 0\}$, a closed half space $\{X \mid a^T X \leq 0\}$ are examples of convex cones.

It immediately follows that if C_1 and C_2 are convex cones then

(a) their sum $C_1 + C_2$ is a convex cone.

(b) their intersection $C_1 \cap C_2$ is a convex cone. and thus

(c) the set of positive linear combination of a finite number of points a_j, j = 1, 2... k, that is, the sum of a finite number of half lines is a convex cone C:

$$C = \left\{ Y \mid Y = \sum_{j=1}^{k} a_j x_j, \quad x_j \geq 0 \right\}$$

Equivalently, C is a cone if for some matrix A

$$C = \{Y \mid Y = AX, X \geq 0\}.$$

(d) The intersection of a finite number of closed half-spaces is a convex cone H:

$$H = \{X \mid b_i X \leq 0, i = 1,2..p\}$$

or in matrix notation,

$$H = \{X \mid BX \leq 0\}$$

Polar Cones: The polar cone C^+ of a cone C is defined by the set

$$C^+ = \{a \mid a^T X \leq 0, \text{ for all } X \in C\}.$$

Clearly, C^+ is a closed convex cone.

For any cones C_1, C_2 the following properties are simple consequences of the definition of polar cones

(i) $C_1 \subset C_2 \Rightarrow C_1^+ \supset C_2^+$

(ii) $C_1 \subset C_1^{++}$, where C_1^{++} is the polar cone of C_1^+

(iii) $(C_1 + C_2)^+ = C_1^+ \cap C_2^+$

Theorem 8.16. If C is a convex cone defined by

$$C = \{X \mid X = A\,U,\ U \geq 0\},$$

then $$C^+ = \{Y \mid A^T Y \leq 0\}$$

Proof: By definition, the polar cone C^+ of C is

$$C^+ = \{Y \mid Y^T X \leq 0,\ X \leq C\}.$$
$$= \{Y \mid Y^T A\ U \leq 0,\ U \geq 0\}$$

Now, $Y^T AU \leq 0$, for all $U \geq 0$ if and only if $A^T Y \leq 0$. Hence.

$$C^+ = \{Y \mid A^T Y \leq 0\}$$

as was to be shown.

Theorem 8.17. If C is a closed convex cone, then $C = C^{++}$

Proof: By property (ii) of polar cones $C \subset C^{++}$. Now, let $X \in C^{++}$ and suppose that $X \notin C$. Then by theorem 8.7 there exists a nonzero vector a and a scalar α such that such that

$$a^T Y \leq \alpha, \text{ for all } Y \in C \text{ and } a^T X > \alpha.$$

Since $Y \in C$, $Y = 0$ is in C and hence by the first inequality $\alpha \geq 0$ and therefore $a^T X > 0$.

This implies that $a \in C^+$, For if not, then $a^T\ \bar{Y} > 0$, for some $\bar{Y} \in C$ and then taking λ arbitrarily large $a^T(\lambda\ \bar{Y})$ can be made arbitrarily large which contradicts that $a^T Y \leq \alpha$ for all $Y \in C$. Hence $a \in C^+$. Since $X \in C^{++}$, $a^T X \leq 0$, but this contradicts that $a^T X > 0$. Therefore $X \in C$ and the proof is complete.

Theorem 8.17 can be used to prove the well-known Farkas' theorem.

Theorem 8.18. (Farkas' theorem)

Let A be an $m \times n$ matrix. Then exactly one of the following two systems has a solution

System 1: $AX = b,\ X \geq 0,\ X \in R^n$

System 2: $A^T Y \geq 0,\ b^T Y < 0,\ Y \in R^m$

An equivalent statement of Farkas' theorem can be given as follows.

System 1: $AX = b,\ X \geq 0$, is consistent if and only if.

System 2: $A^T Y \leq 0,\ b^T Y \leq 0$, has a solution.

Proof: Consider the cones

$$C = \{AX \mid X \geq 0\}$$
$$C^+ = \{Y \mid A^T Y \leq 0\}, \text{ by theorem 8.16}$$

$$C^{++} = \{U \mid Y^T U \leq 0, Y \in C^+\}, \text{ by definition}$$

By theorem 8.17, $b \in C^{++}$, if and only if $b \in C$.

Now, $b \in C^{++}$ implies that whenever $Y \in C^+$, then $b^T Y \leq 0$ which means that $A^T Y \leq 0$ implies that $b^T Y \leq 0$

And $b \in C$ implies that $AX = b, X \geq 0$.

Thus system 1 is consistent if and only if System 2 has a solution.

Polyhedral Convex Cones

The convex cone spanned by a finite set of vectors $a_1, a_2,..a_r$ is called a polydral convex cone and is denoted by

$$C_p = \left\{X \mid X = \sum_{k=1}^{r} \lambda_k a_k, \quad \lambda_k \leq 0, \ k = 1,2,..r\right\}$$

In matrix notation, it can be written as

$$C_p = \{X \mid X = A\lambda, \lambda \geq 0\}$$

where A is the $n \times r$ matrix whose column vectors are a_k, $k = 1,2..r$ and λ is $r \times 1$.

Thus, C_p is the sum of finite number of half-lines.

A polyhedral convex cone may also be defined as the intersection of a finite number of closed half-spaces whose generating hyperplanes pass through the origin.

A polyhedral convex cone is thus the set of solutions to some finite system of homogeneous linear inequalities.

CHAPTER 9

Convex and Concave Functions

In this chapter we introduce convex and concave functions defined on convex sets in R^n and give some of their basic properties and obtain some fundamental theorems involving these functions. We will see later that these theorems are very important in deriving optimality conditions for nonlinear programming problems and developing suitable computational schemes.

In the following definitions and theorems, the functions are numerical functions, that is real single valued functions, defined on a convex set S in R^n.

9.1. Definitions and Basic Properties

A function f defined on a convex set S in R^n, is said to be a convex function on S, if

$$f[\lambda X_1 + (1-\lambda)X_2] \leq \lambda f(X_1) + (1-\lambda) f(X_2)$$

for each $X_1, X_2 \in S$ and each λ, $0 \leq \lambda \leq 1$.

The function f is said to be strictly convex on S if the above inequality is strict for $X_1 \neq X_2$, and $0 < \lambda < 1$.

A function f is said to be concave (strictly concave) if $-f$ is convex (strictly convex).

It is clear that a linear function is convex as well as concave but neither strictly convex nor strictly concave.

Alternatively, a function f defined on a convex set S in R^n is convex (concave) if linear interpolation between the values of the function never underestimates (overestimates) the actual value at the interpolated point.

A geometrical interpretation of convex and concave functions is given in Figure 9.1.

Let x_1 and x_2 be two points in R^1 and consider the point $x = \lambda x_1 + (1-\lambda)x_2$, $\lambda \in (0,1)$, $\lambda f(x_1) + (1-\lambda) f(x_2)$ then gives the linear interpolation between the values of $f(x)$ at x_1 and x_2 while $f[\lambda x_1 + (1-\lambda)X_2]$ gives the value of $f(x)$ at the interpolated point $\lambda x_1 + (1-\lambda)x_2$.

So for a convex function f, we have

$$f[\lambda x_1 + (1-\lambda)x_2] \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

and for a concave function f

$$f[\lambda x_1 + (1-\lambda)x_2] \geq \lambda f(x_1) + (1-\lambda)f(x_2).$$

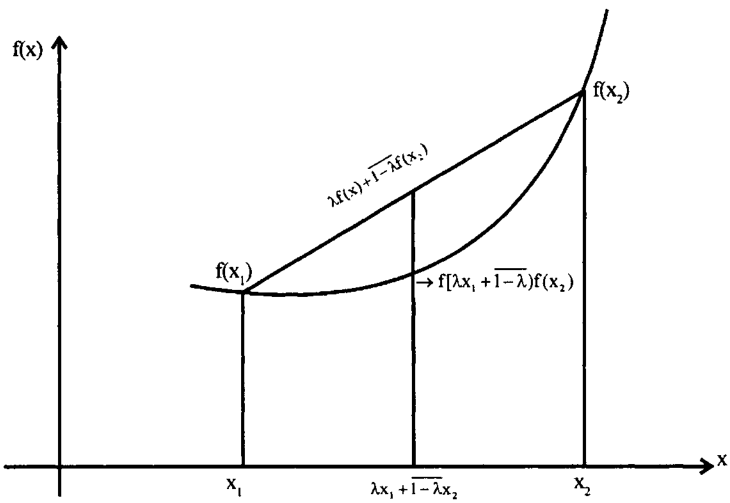


Figure 9.1(a): Convex Function.

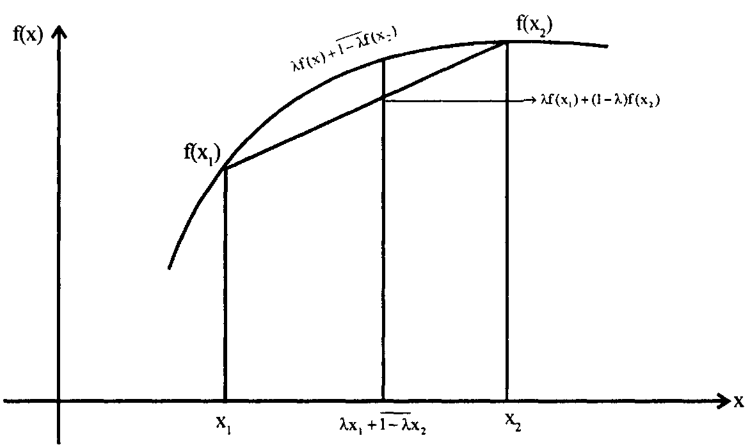


Figure 9.1(b): Concave Function.

Theorem 9.1: A necessary and sufficient condition that a function f defined on a convex set S in R^n is convex is that

$$f\left(\sum_{i=1}^{m} \lambda_i X_i\right) \leq \sum_{i=1}^{m} \lambda_i f(x_i), \quad \text{for all integers m,}$$

where $$X_i \in S,\ \lambda_i \geq 0,\ \ i = 1,2,..m, \sum_{i=1}^{m} \lambda_i = 1$$

Proof: Suppose that the condition is satisfied. For m = 2, clearly f is convex and this is sufficient to show that f is convex for all integers m.

Conversely, suppose that f is a convex function. The proof will be given by the method of induction. Clearly, the condition is satisfied for m = 1 and m = 2. Suppose that it is satisfied for m = r and we prove that it is also satisfied for m = r+1

Now, $$\sum_{i=1}^{r+1} \lambda_i X_i = \sum_{i=1}^{r} \lambda_i X_i + \lambda_{r+1} X_{r+1}$$

$$= \sum_{i=1}^{r} \lambda_i Y + \lambda_{r+1} X_{r+1}$$

where $$Y = \sum_{i=1}^{r} \mu_i X_i,\quad \mu_i \frac{\lambda_i}{\sum_{i=1}^{r} \lambda_i},\quad i = 1,2,..r$$

and then $Y \in S$

Therefore,

$$f\left(\sum_{i=1}^{r+1} \lambda_i X_i\right) \leq \left(\sum_{i=1}^{r} \lambda_i\right) f(Y) + \lambda_{r+1} f(X_{r+1})$$

$$= \left(\sum_{i=1}^{r} \lambda_i\right) f\left(\sum_{i=1}^{r} \mu_i X_i\right) + \lambda_{r+1} f(X_{r+1})$$

$$\leq \left(\sum_{1}^{r} \lambda_i\right)\left(\sum_{i=1}^{r} \mu_i f(X_i)\right) + \lambda_{r+1} f(X_{r+1})$$

$$= \left(\sum_{i=1}^{r} \lambda_i\right) \frac{\sum_{i=1}^{r} \lambda_i f(X_i)}{\sum_{i=1}^{r} \lambda_i} + \lambda_{r+1} f(X_{r+1})$$

$$= \sum_{i=1}^{r} \lambda_i f(X_i) + \lambda_{r+1} f(X_{r+1})$$

$$= \sum_{i=1}^{r+1} \lambda_i f(X_i)$$

Hence, the theorem is proved by induction.

Theorem 9.2: A necessary and sufficient condition that a function f defined on a convex set S in R^n is convex is that the one dimensional function, ϕ defined by $\phi(\lambda) = f[\lambda X_1 + (1-\lambda)X_2]$, $X_1, X_2 \in S$ is convex on [0,1]

Proof: Let ϕ be a convex function on [0,1].

Then for $\lambda \in [0,1]$,

$$f[\lambda X_1 + (1-\lambda)X_2] = \phi(\lambda) = \phi[\lambda 1 + (1-\lambda)0].$$
$$\leq \lambda \phi(1) + (1-\lambda)\phi(0).$$
$$= \lambda f(X_1) + (1-\lambda) f(X_2)$$

This holds for all $X_1, X_2 \in S$ and hence f is convex on S.

Conversely, suppose that f is convex on S.

Then for $\lambda_1, \lambda_2 \in [0,1]$

$$\phi[\alpha\lambda_1 + (1-\alpha)\lambda_2] = f[\{\alpha\lambda_1 + (1-\alpha)\lambda_2\} X_1 + \{1-(\alpha\lambda_1 + \overline{1-\alpha}\,\lambda_2)\}X_2]$$
$$X_1, X_2 \in S,\ 0 \leq \alpha \leq 1.$$
$$= f[\alpha\{X_2 + \lambda_1(X_1 - X_2)\} + (1-\alpha)\{X_2 + \lambda_2(X_1 - X_2)\}]$$
$$= f[\alpha\{\lambda_1 X_1 + (1-\lambda_1)X_2\} + (1-\alpha)\{\lambda_2 X_1 + (1-\lambda_2)X_2\}]$$
$$\leq \alpha f[\lambda_1 X_1 + (1-\lambda)X_2] + (1-\alpha) f[\lambda_2 X_1 + (1-\lambda_2)X_2]$$
$$= \alpha \phi(\lambda_1) + (1-\alpha)\phi(\lambda_2)$$

Hence ϕ is convex on [0,1]

Theorem 9.3: If f_h, h = 1, 2..m are convex functions on a convex set S in R^n and $\lambda_h \geq 0$, h = 1,2..m, then the function f defined by

$$f(X) = \sum_{h=1}^{m} \lambda_h f_h(X)$$

is also a convex function on S and strictly convex if at least one of the functions f_h is strictly convex.

Proof: Let f_h, h = 1, 2..m be convex functions on a convex set in R^n and let

$$f(X) = \sum_{h=1}^{m} \lambda_h f_h(X), \quad \lambda_h \geq 0, \quad h = 1,2..m.$$

Let $0 \leq \alpha \leq 1$, then

$$f[\alpha X_1 + (1-\alpha)X_2] = \sum_{h=1}^{m} \lambda_h f_h[\alpha X_1 + (1-\alpha)X_2]$$

$$\leq \sum_{h=1}^{m} \lambda_h[\alpha f_h(X_1) + (1-\alpha)f_h(X_2)] \quad \because f_h(X) \text{ is convex.}$$

$$= \alpha \sum_{h=1}^{m} \lambda_h f_h(X_1) + (1-\alpha)\sum_{h=1}^{m} \lambda_h f_h(X_2)$$

$$= \alpha f(X_1) + (1-\alpha) f(X_2).$$

Hence f is convex.

If at least one of the functions f_h is strictly convex, then strict inequality will follow in the above expression and in that case f is strictly convex.

Theorem 9.4. A function f defined on a convex set S in R^n is convex if and only if its epigraph defined by the set

$$S_{epi} = \{(X, \xi) \mid X \in S, \xi \in R^1, f(X) \leq \xi\}$$

is a convex set in R^{n+1}.

Proof: Suppose that f is convex and let (X_1, ξ_1) and $(X_2, \xi_2) \in S_{epi}$. Then for $0 \leq \lambda \leq 1$,

$$f[\lambda X_1 + (1-\lambda)X_2] \leq \lambda f(X_1) + (1-\lambda)f(X_2)$$
$$\leq \lambda\xi_1 + (1-\xi)\xi_2.$$

Hence $[\lambda X_1 + (1-\lambda)X_2, \lambda\xi_1 + (1-\lambda)\xi_2] \in S_{epi}$ and S_{epi} is is a convex set in R^{n+1}.

Conversely, assume that S_{epi} is convex. Let $X_1, X_2, \in S$, then $[X_1, f(X_1)]$ and $[X_2, f(X_2)]$ belong to S_{epi} Since S_{epi} is convex, we must have

$$[\lambda X_1 + (1-\lambda)X_2, \lambda f(X_1) + (1-\lambda)f(X_2)] \in S_{epi}, \text{ for } 0 \leq \lambda \leq 1$$
$$\text{or } f[\lambda X_1 + (1-\lambda)X_2] \leq \lambda f(X_1) + (1-\lambda)f(X_2), \text{ for } 0 \leq \lambda \leq 1$$

and hence f is convex on S.

Theorem 9.5. If (f_i, $i \in I$) is a family of convex functions on a convex set S in R^n, then the function f defined by $f(X) = \text{Sup}_{i\in I} f_i(X)$ is convex on S.

Proof: Since f_i, $i \in I$ are convex functions on S, then by theorem 9.4, their epigraphs

$$S_{epi} = \{(X, \xi) \mid X \in S, \xi \in R^1, f_i(X) \leq \xi\}, i \in I$$

are convex sets in R^{n+1}.

Hence their intersection

$$\bigcap_{i\in I} S_{epi} = \{(X, \xi) \mid X \in S, \xi \in R^1, f_i(X) \leq \xi, \text{ for all } i \in I\}$$

$$= \left\{(X,\xi) | X \in S,\ \xi \in R^1, \quad \text{Sup}_{i\in I} f_i(X) \leq \xi\right\}$$

is also a convex set in R^{n+1}. Thus the epigraph of f is convex.

By theorem 9.4, then f is a convex function on S.

Theorem 9.6. If f is a convex function defined on a convex set S in R^n, then the set

$$S_\alpha = \{X \mid X \in S, f(X) \leq \alpha\},$$

called a level set, is convex for each real number α.

Proof: Let f be a convex function on S and let $X_1, X_2 \in S_\alpha$. Then $X_1, X_2, \in S$, $f(X_1) \leq \alpha$, $f(X_2) \leq \alpha$.

By convexity of S, $X = \lambda X_1 + (1-\lambda)X_2 \in S$, for $0 \leq \lambda \leq 1$ and since f is convex,

$$\begin{aligned} f(X) &= f[\lambda X_1 + (1-\lambda)X_2] \\ &\le \lambda\, f(X_1) + (1-\lambda)\, f(X_2) \\ &\le \lambda\alpha + (1-\lambda)\,\alpha = \alpha \end{aligned}$$

Hence $X \in S_\alpha$ and therefore S_α is convex.

Local and Global Minima of Functions

A point $\bar{X} \in S \subset R^n$ is called a local minimum point of the function f in S if there exists an $\in$-neighbourhood $N_\in(\bar{X})$ around $\bar{X}$ such that $f(X) \ge f(\bar{X})$, for each $X \in S \cap N_\in(\bar{X})$. If a point $\bar{X} \in S$ is such that $f(X) \ge f(\bar{X})$, for each $X \in S$, then $\bar{X}$ is called a global minimum point of f.

A local and global maximum point of a function can be defined similarly.

Theorem 9.7: Any local minimum of a convex function attained on a convex set is a global minimum. The set of all these minima is convex

Proof: Let $\bar{X}$ minimize the convex function f on a convex set S in some $\in$-neighbourhood $N_\in(\bar{X})$ around $\bar{X}$ such that

$$f(\bar{X}) \le f(X)\text{ , for each } X \in S \cap N_\in(\bar{X})$$

that is, $\bar{X}$ is a local minimum point

Suppose that $\bar{X}$ is not a global minimum point so that there exists a point $X_0 \in S$, but not in $N \in (\bar{X})$ such that

$$f(X_0) < f(\bar{X})$$

Further, suppose that $X_1 \in S$ is a point in the neighbourhood of $N_\in(\bar{X})$ so that

$$X_1 = \lambda\, X_0 + (1-\lambda)\ \bar{X}\text{ , } 0 < \lambda < 1.$$

Now,

$$\begin{aligned} f(X_1) &= f[\lambda X_0 + (1-\lambda)\ \bar{X}] \\ &\le \lambda f(X_0) + (1-\lambda)\ f(\bar{X}) \\ &= f(\bar{X}) + \lambda\ [f(X_0) - f(\bar{X})] \\ &< f(\bar{X}), \qquad \text{since } f(X_0) < f(\bar{X}) \end{aligned}$$

which contradicts the assumption that $\bar{X}$ is a local minimum point. Hence f attains a global minimum at $\bar{X}$.

Let $\bar{S}$ be the set of points at which the convex function f attains its minimum.

Let $X_1, X_2 \in \bar{S}$ and $f(X_1) = f(X_2) = m$

Now, for $0 \le \alpha \le 1$, $f[\alpha X_1 + (1-\alpha)X_2] \le \alpha\ f(X_1) + (1-\alpha)f(X_2) = \alpha m + (1-\alpha)m = m$

which means f attains its minimum also at $\alpha X_1 + (1-\alpha)X_2$, $0 \le \alpha \le 1$ and therefore belongs to $\bar{S}$. Hence $\bar{S}$ is convex.

Theorem 9.8. A strictly convex function on a convex set S in R^n attains its global minimum at an unique point in S.

Proof: Let f be a strictly convex function on a convex set S in R^n and attains its global minimum at $X_1, X_2, \in S$, $X_1 \neq X_2$ so that $f(X_1) = f(X_2) = m$

Let $X = \lambda X_1 + (1-\lambda)X_2$, $0 < \lambda < 1$.

Then

$$\begin{aligned} f(X) &= f[\lambda X_1 + (1-\lambda)X_2] \\ &< \lambda f(X_1) + (1-\lambda) f(X_2) \\ &= \lambda m + (1-\lambda) m \\ &= m. \end{aligned}$$

The above inequality contradicts that m is the global minimum.

This completes the proof.

Theorem 9.9. A quadratic form $Q(X) = X^TBX$ is convex for all $X \in R^n$, if and only if the symmetric matrix B is positive semi-definite.

Proof: Let $X_1, X_2 \in R^n$. Then for $0 \le \lambda \le 1$,

$$\lambda Q(X_1) + (1-\lambda) Q(X_2) - Q[\lambda X_1 + (1-\lambda)X_2]$$

$$= \lambda\left(X_1^TBX_1\right) + (1-\lambda)\left(X_2^TBX_2\right) - [\lambda X_1 + (1-\lambda)X_2]^T B[\lambda X_1 + (1-\lambda)X_2]$$

$$= \lambda\left(X_1^TBX_1\right) + (1-\lambda)\left(X^TBX_2\right) - \lambda^2 X_1^TBX_1 - (1-\lambda)^2 X_2^TBX_2 - 2\lambda(1-\lambda)X_1^TBX_2]$$

$$= \lambda(1-\lambda)X_1^TBX_1 + \lambda(1-\lambda)X_2^TBX_2 - 2\lambda(1-\lambda)X_1^TBX_2$$

$$= \lambda(1-\lambda)\left[X_1^TBX_1 + X_2^TBX_2 - 2X_1^TBX_2\right]$$

$$= \lambda(1-\lambda)(X_1-X_2)^TB(X_1-X_2)$$

≥ 0, for all λ, $0 \le \lambda \le 1$, if and only if the matrix B is positive semidefinite.

This proves the theorem.

Corollary 9.1. A positive definite quadratic form is a strictly convex function over all of R^n

Continuity of Convex Functions

It should be noted that a convex function f defined on a convex set S in R^n is not necessarily continuous everywhere. For example, the convex function

$$\begin{aligned} f(x) &= 2 \text{ for } |x| = 1 \\ &= x^2 \text{ for } |x| < 1 \end{aligned}$$

defined on $S = \{x \mid -1 \le x \le 1\}$ is not continuous at the boundary points of S.

It can however, be shown that if f is a convex function on an open convex set S in R^n, then it is continuous on S (Theorem 9.10)

Theorem 9.10. If f is a convex function on an open convex set S in R^n then f is continuous on S

Proof: See Flemming [162]

Since the interior of any set S in R^n is open, it follows that a convex function defined on a convex set S in R^n is continuous on the interior of S.

9.2. Differentiable Convex Functions

Theorem 9.11. If f is a differentiable function on an open convex set S in R^n, then f is convex on S if and only if for each $X_1, X_2 \in S$,

$$f(X_2) - f(X_1) \geq \nabla f(X_1)^T (X_2 - X_1).$$

Proof: Let f be a differentiable convex function on S. For $X_1, X_2 \in S$ and $0 < \lambda \leq 1$, we then have

$$(1-\lambda) f(X_1) + \lambda f(X_2) \geq f[(1-\lambda)X_1 + \lambda X_2]$$

or $$\lambda[f(X_2) - f(X_1)] \geq f[X_1 + \lambda(X_2 - X_1)] - f(X_1).$$

$$= f(X_1) + \lambda \nabla f(X_1)^T (X_2 - X_1)$$

$$+ \lambda \|X_2 - X_1\| \alpha [X_1; \lambda (X_2 - X_1)] - f(X_1)$$

where $\alpha [X_1; \lambda (X_2 - X_1)] \to 0$ as $\lambda \to 0$.

or $$f(X_2) - f(X_1) \geq \nabla f(X_1)^T (X_2 - X_1) + \|X_2 - X_1\| \alpha [X_1; \lambda (X_2 - X_1)]$$

Now, taking the limit as $\lambda \to 0$, we obtain

$$f(X_2) - f(X_1) \geq \nabla f(X_1)^T (X_2 - X_1).$$

Conversely, suppose that

$$f(X_2) - f(X_1) \geq \nabla f(X_1)^T (X_2 - X_1), \text{ for } X_1, X_2 \in S.$$

Since S is convex, $\lambda X_1, + (1-\lambda) X_2 \in S$, for $0 \leq \lambda \leq 1$.

and we then have

$$f(X_2) - f[\lambda X_1 + (1-\lambda)X_2] \geq \lambda \nabla f [\lambda X_1 + (1-\lambda)X_2]^T (X_2 - X_1)$$

$$f(X_1) - f[\lambda X_1 + (1-\lambda)X_2] \geq -(1-\lambda) \nabla f [\lambda X_1 + (1-\lambda)X_2]^T (X_2 - X_1)$$

Multiplying the first inequality by $(1-\lambda)$ and the second one by λ and adding we get

$$\lambda f(X_1) + (1-\lambda) f(X_2) \geq f[\lambda X_1 + (1-\lambda)X_2]$$

Hence f is convex.

Theorem 9.12. If f is a differentiable function on an open convex set S in R^n, then f is strictly convex if and only if for each $X_1, X_2 \in S$, $X_1 \neq X_2$,

$$f(X_2) - f(X_1) > \nabla f(X_1)^T (X_2 - X_1)$$

Proof: It immediately follows from theorem 9.11 by changing all inequalities to strict inequalities for $0 < \lambda < 1$.

Theorem 9.13: If f is a differentiable function on an open convex set S in R^n, then f is convex if and only if for each $X_1, X_2 \in S$

$$[\nabla f(X_2) - \nabla f(X_1)]^T (X_2 - X_1) \geq 0.$$

Proof: Let f be convex on S. Then for any $X_1, X_2 \in S$,
we have by theorem 9.11

$$f(X_2) - f(X_1) \geq \nabla f(X_1)^T (X_2 - X_1)$$
$$f(X_1) - f(X_2) \geq \nabla f(X_2)^T (X_1 - X_2)$$

Adding the above two inequalities, we get

$$[\nabla f(X_2) - \nabla f(X_1)]^T (X_2 - X_1) \geq 0.$$

To prove the converse, let $X_1, X_2 \in S$. Since S is convex, $\lambda X_1 + (1-\lambda)X_2 \in S$ for $0 \leq \lambda \leq 1$.

Now, by the mean value theorem we have for some $\lambda \in (0, 1)$

$$f(X_2) - f(X_1) = \nabla f\,[\lambda X_1 + (1-\lambda)X_2]^T (X_2 - X_1)$$

But by assumption,

$$[\nabla f[\lambda X_1 + (1-\lambda)X_2] - \nabla f(X_1)]^T [\{\lambda X_1 + (1-\lambda)X_2\} - X_1] \geq 0$$

or $\quad (1-\lambda)[\nabla f[\lambda X_1 + (1-\lambda)X_2] - \nabla f(X_1)]^T (X_2 - X_1) \geq 0$

or $\quad \nabla f\,[\lambda X_1 + (1-\lambda)X_2]^T (X_2 - X_1) \geq \nabla f(X_1)^T (X_2 - X_1)$

Hence $\quad f(X_2) - f(X_1) \geq \nabla f(X_1)^T (X_2 - X_1)$

and by theorem 9.11, f (X) is convex.

It immediately follows that

Theorem 9.14: If f is a differentiable function on an open convex set S in R^n, then f is strictly convex if and only if for each $X_1, X_2, \in S$, $X_1 \neq X_2$

$$[\nabla f(X_2) - \nabla f(X_1)]^T (X_2 - X_1) > 0$$

Theorem 9.15: If f is a differentiable convex function on a convex set S in R^n, then $f(X_0)$, $X_0 \in S$ is a global minimum of f(X) if and only if $\nabla f(X_0)^T(X - X_0) \geq 0$ for all $X \in S$.

Proof: Suppose that

$$\nabla f(X_0)^T(X - X_0) \geq 0, \text{ for all } X \in S$$

Since f is convex, by theorem 9.11, we have

$$f(X) - f(X_0) \geq \nabla f(X_0)^T(X - X_0) \geq 0, \text{ for all } X \in S.$$

which implies that $f(X_0)$ is a global minimum.

To prove the converse, suppose that $f(X_0)$, is a global minimum of f(X) on S.

Then for all $X \in S$, $f(X_0) \leq f(X)$.

Since $\lambda X + (1-\lambda)X_0 \in S$, for any $X \in S$, $0 < \lambda < 1$

$$f(X_0) \leq f[\lambda X + (1-\lambda)X_0].$$

or $\quad f(X_0) \leq f[X_0 + \lambda(X - X_0)]$

or $\quad f[X_0 + \lambda(X - X_0)] - f(X_0) \geq 0$

Dividing the Taylor's expansion by λ and taking the limit as $\lambda \to 0$, we have

$$\nabla f(X_0)^T(X - X_0) \geq 0$$

Theorem 9.16: If a function ϕ is a differentiable function of single variable defined on an open interval $D \subset R^1$ then ϕ is convex on D if and only if ϕ', the derivative of ϕ is nondecreasing on D.

Proof: Let ϕ be a differentiable convex function on D.

Then by theorem 9.11 we have

$$\phi(\lambda_2) - \phi(\lambda_1) \geq \phi'(\lambda_1)(\lambda_2 - \lambda_1)$$

and $\quad \phi(\lambda_1) - \phi(\lambda_2) \geq \phi'(\lambda_2)(\lambda_1 - \lambda_2)$

for $\quad \lambda_1, \lambda_2 \in D$, $\lambda_1 < \lambda_2$

Hence

$$\varphi'(\lambda_1) \leq \frac{\varphi(\lambda_2) - \varphi(\lambda_1)}{\lambda_2 - \lambda_1} \leq \varphi'(\lambda_2), \qquad \text{for } \lambda_1 < \lambda_2$$

and therefore $\phi'(\lambda)$ is nondecreasing on D.

Conversely, suppose that ϕ' is nondecreasing on D. Let $\lambda_1, \lambda_2 \in D$, $\lambda_1 < \lambda_2$ and $\bar{\lambda} = \mu\lambda_1 + (1-\mu)\lambda_2$, $0 < \mu < 1$.

By the mean value theorem, we have

$$\phi(\lambda_2) - \phi(\bar{\lambda}) = \mu\,(\lambda_2 - \lambda_1)\,\phi'(\eta_1)\,, \quad \bar{\lambda} < \eta_1 < \lambda_2$$

$$\phi(\bar{\lambda}) - \phi(\lambda_1) = (1 - \mu)\,(\lambda_2 - \lambda_1)\,\phi'(\eta_2)\ \lambda_1 < \eta_2 < \bar{\lambda}$$

Since ϕ' is nondecreasing, $\phi'(\eta_1) \geq \phi'(\eta_2)$ for $\eta_1 > \eta_2$, and therefore from the above two expressions, we have

$$\begin{aligned}\phi(\bar{\lambda}) &= \phi[\mu\lambda_1 + (1-\mu)\lambda_2] \\ &\leq \mu\,\phi(\lambda_1) + (1-\mu)\,\phi(\lambda_2).\end{aligned}$$

Hence ϕ is convex.

Twice Differentiable Convex Functions

Theorem 9.17. If f is a twice differentiable function on an open convex set S in R^n, then f is convex on S if and only if the Hessian matrix H(X) is positive semidefinite for each $X \in S$.

Proof: Let f be convex on S and let $\bar{X} \in S$. We are then to show that for each $X \in R^n$, $X^T H(\bar{X})X \geq 0$. Since S is an open convex set, there exists a $\bar{\lambda} > 0$ such that for any $X \in R^n$

$$\bar{X} + \lambda X \in S\,, \text{ for } 0 < \lambda < \bar{\lambda}.$$

By theorem 9.11 we have

$$f(\bar{X} + \lambda X) \geq f(\bar{X}) + \lambda\, \nabla f(\bar{X})^T X\,, \text{ for } 0 < \lambda < \bar{\lambda}$$

and since f is twice differentiable

$$f(\bar{X} + \lambda X) = f(\bar{X}) + \lambda f(\bar{X})^T X + \tfrac{1}{2}\lambda^2 X^T H(\bar{X})X + \lambda^2 \|X\|^2 \beta(\bar{X}\,;\, \lambda X)$$

Hence $\frac{1}{2}\lambda^2 X^T H(\bar{X})X + \lambda^2\|X\|^2\beta(\bar{X};\lambda X) \geq 0.$

Dividing by λ^2 and taking the limit as $\lambda \to 0$, we have, since $\lim_{\lambda\to 0}\beta(\bar{X};\lambda X) = 0$

$$X^T H(\bar{X})X \geq 0, \quad \text{for all } X \in R^n$$

Hence H(X) is positive semidefinite for all $X \in S$.

Conversely, suppose that the Hessian matrix is positive semidefinite for each point in S. Let $X_1, X_2 \in S$ and consider

$$\hat{X} = X_1 + \lambda(X_2 - X_1), \quad 0 < \lambda < 1$$

obviously, $\hat{X} \in S$.

Now, by Taylor's theorem, we have

$$f(X_2) = f(X_1) + \nabla f(X_1)^T (X_2 - X_1) + \frac{1}{2}(X_2 - X_1)^T H(\hat{X})(X_2 - X_1)$$

Since $\hat{X} \in S$ and by assumption $H(\hat{X})$ is positive semidefinite, we get

$$f(X_2) - f(X_1) \geq \nabla f(X_1)^T (X_2 - X_1)$$

and by theorem 9.11 f is convex on S.

Theorem 9.18. Let f be a twice differentiable function on an open convex set S in R^n. Then f is strictly convex on S if the Hessian matrix H(X) is positive definite for each $X \in S$, that is, for each $X \in S$, $Y^T H(X)Y > 0$, for all $Y \in R^n$, $Y \neq 0$. The converse is not true.

Proof: If H(X) is positive definite for all $X \in S$, then from theorem 9.17, we find

$$f(X_2) - f(X_1) > \nabla f(X_1)^T (X_2 - X_1), \; X_1, X_2 \in S, \; X_1 \neq X_2$$

Hence by theorem 9.12, f is strictly convex

To see that the converse is not true, consider the function defined by $f(x) = x^4$, $x \in R^1$. f is strictly convex on R^1 but $H(x) = 12x^2$ is zero at $x = 0$ and hence is not positive definite.

9.3. Generalization of Convex Functions

In the beginning of the development of nonlinear programming, the theory and methods were mainly concerned with problems involving convex functions. It was however, gradually realized that not all the properties of convex functions are needed to establish many of the results of nonlinear programming—only some weaker properties are required. In this section, we present various types of functions, similar to convex functions, sharing some of their weaker properties.

Quasiconvex Function

A function f defined on a convex set S in R^n is said to be quasiconvex if for each $X_1, X_2 \in S$,

$$f(X_1) \le f(X_2) \Rightarrow f[\lambda X_1 + (1-\lambda)X_2] \le f(X_2)$$

for each $\lambda \in [0, 1]$

or equivalently, if for each $X_1, X_2 \in S$

$$f[\lambda X_1 + (1-\lambda)X_2] \le \text{Maximum } \{f(X_1), f(X_2)\}.$$

for each $\lambda \in [0, 1]$

The function f is said to be quasiconcave if – f is quasiconvex.

Theorem 9.19. Let f be a function defined on a convex set S in R^n, Then f is quasiconvex on S if and only if the set

$$S_\alpha = \{X \mid X \in S, f(X) \le \alpha\} \text{ is convex for each real number } \alpha.$$

Proof: Suppose that S_α is convex for each real number α. Let $X_1, X_2 \in S$, then $X = \lambda X_1 + (1-\lambda)X_2 \in S$ for $\lambda \in [0, 1]$

Further, let α = maximum $\{f(X_1), f(X_2)\}$, then $X_1, X_2 \in S_\alpha$ and since S_α is convex by assumption, $X \in S_\alpha$. Therefore,

$$f(X) \le \alpha = \max \{f(X_1), fX_2)\}.$$

Hence f(X) is quasiconvex.

Conversely, suppose that f is quasiconvex on S and let $X_1, X_2 \in S_\alpha$. Without loss of generality let us assume that $f(X_2) \le f(X_1)$, and since $X_1, X_2 \in S_\alpha$, $f(X_2) \le f(X_1) \le \alpha$.

Now since $X_1, X_2 \in S$ and S is convex,

$$f[\lambda X_1 + (1-\lambda)X_2] \in S \text{ for } \lambda \in [0, 1]$$

and since f is quasiconvex,

$$f[\lambda X_1 + (1-\lambda)X_2] \le f(X_1) \le \alpha$$

Hence $\lambda X_1 + (1-\lambda)X_2 \in S_\alpha$ for $\lambda \in [0, 1]$

and therefore S_α is convex.

Theorem 9.20: A function f is quasiconvex on a convex set $S \subset R^n$ if and only if for all $X_1, X_2 \in S$, the function, ϕ given by $\phi(\lambda) = f[\lambda X_1 + (1-\lambda)X_2]$ is quasiconvex in [0,1].

Proof: Suppose that f is quasiconvex on S and let $X_1, X_2 \in S$, $X_1 \ne X_2$ and

$$Y_1 = \lambda_1 X_1 + (1-\lambda_1)X_2$$
$$Y_2 = \lambda_2 X_1 + (1-\lambda_2)X_2$$

for $0 \le \lambda_1 < \lambda_2 \le 1$.

Further let $Y_0 \in (Y_1, Y_2)$, so that Y_0 can be expressed as.

$$Y_0 = \lambda_0 X_1 + (1-\lambda_0)X_2, \lambda_1 < \lambda_0 < \lambda_2$$

Thus, we have $Y_1, Y_2 \in [X_1, X_2]$ and $Y_0 \in (Y_1, Y_2)$ and hence $f(Y_1) \le f(Y_2)$ implies $f(Y_0) \le f(Y_2)$ which means $\phi(\lambda_1) \le \phi(\lambda_2) \Rightarrow \phi(\lambda_0) \le \phi(\lambda_2)$, for all $\lambda_1\ \lambda_2 \in [0, 1]$, $\lambda_0 \in (\lambda_1, \lambda_2)$

Hence ϕ is quasiconvex in [0,1]

Conversely, let ϕ be quasiconvex in [0,1]. Putting $\lambda_1 = 0$, $\lambda_2 = 1$, we get

$$\phi(0) \le \phi(1) \Rightarrow \phi(\lambda_0) \le \phi(1) \text{ for } \lambda_0 \in [0, 1]$$

which means

$$f(X_2) \le f(X_1) \Rightarrow f[\lambda_0 X_1 + (1-\lambda_0)X_2] \le f(X_1)$$

Thus f is quasiconvex.

Theorem 9.21. Let f be a differentiable function on an open convex set S in R^n. Then *f* is quasiconvex if and only if for all $X_1, X_2 \in S$.

$$f(X_1) \le f(X_2) \text{ implies } \nabla f(X_2)^T(X_1 - X_2) \le 0.$$

Proof: Let f be quasiconvex and let $X_1, X_2 \in S$ be such that $f(X_1) \le f(X_2)$.

Then $$f[\lambda X_1 + (1-\lambda)X_2] \le f(X_2), \text{ for } \lambda \in (0, 1). \tag{9.1}$$

By differentiability of f(X) at X_2, we have for $\lambda \in (0, 1)$.

$$f[\lambda X_1 + (1-\lambda)X_2] - f(X_2) = \lambda \nabla f(X_2)^T(X_1-X_2) + \lambda\|X_1-X_2\|\alpha[X_2; \lambda(X_1-X_2)]$$

where $\alpha[X_2; \lambda(X_1-X_2)] \to 0$ as $\lambda \to 0$

By (9.1), we have

$$\lambda \nabla f(X_2)^T (X_1 - X_2) + \lambda \|X_1 - X_2\| \alpha [X_2; \lambda(X_1 - X_2)] \le 0$$

Dividing by λ and letting $\lambda \to 0$, we get

$$\nabla f(X_2)^T (X_1 - X_2) \le 0.$$

To prove the converse, suppose that $X_1, X_2 \in S$ be such that $f(X_1) \le f(X_2)$ and $\nabla f(X_2)^T(X_1-X_2) \le 0$.

We are then to show that

$$f[\lambda X_1 + (1-\lambda)X_2] \le f(X_2), \text{ for each } \lambda \in (0, 1)$$

We establish the quasiconvexity of f by showing that

$$\Omega = \{X \mid X = \lambda X_1 + (1-\lambda)X_2, \lambda \in (0, 1), f(X) > f(X_2)\}$$

is empty.

We assume by contradiction that there exists an $\bar{X} \in \Omega$, so that

$$\bar{X} = \lambda X_1 + (1-\lambda)X_2, \text{ for some } \lambda \in (0, 1) \tag{9.2}$$

and $$f(\bar{X}) > f(X_2) \tag{9.3}$$

Since f is differentiable, it is continuous and there must exist a $\delta \in (0, 1)$, such that

$$f[\mu \bar{X} + (1-\mu)X_2] > f(X_2) \text{ for each } \mu \in [\delta, 1] \tag{9.4}$$

and $$f(\bar{X}) > f[\delta \bar{X} + (1-\delta)X_2] \tag{9.5}$$

Let $\hat{X} = \theta\bar{X} + (1 - \theta)X_2,$ for some $\theta \in (\delta, 1)$

By the mean value theorem, we then have,

$$f(\bar{X}) = f[\delta\bar{X} + (1-\delta)X_2] + (1-\delta)\nabla f(\hat{X})^T(\bar{X} - X_2) \tag{9.6}$$

Hence $$(1-\delta)\nabla f(\hat{X})^T(\bar{X} - X_2) = f(\bar{X}) - f[\delta\bar{X} + (1-\delta)X_2] > 0, \text{ by (9.5)} \tag{9.7}$$

and (9.4) implies that

$$f(\hat{X}) > f(X_2) \tag{9.8}$$

Dividing (9.7), by (1–δ), we have

$$\nabla f(\hat{X})^T (\bar{X} - X_2) > 0 \tag{9.9}$$

Now, since $\quad \bar{X} - X_2 = \lambda X_1 + (1-\lambda)X_2 - X_2$

$$= \lambda (X_1 - X_2) \text{ and } \lambda \in (0, 1),$$

we have

$$\nabla f(\hat{X})^T (X_1 - X_2) > 0 \tag{9.10}$$

and further $\quad f(\hat{X}) > (X_2) \geq f(X_1) \qquad (9.11)$

By the assumption of the theorem, we have

$$\nabla f(\hat{X})^T (X_1 - \hat{X}) \leq 0 \tag{9.12}$$

Now, $\quad \hat{X} = \theta \bar{X} + (1 - \theta)X_2$, for some $\theta \in (\delta, 1)$

$$= \theta[\lambda X_1 + (1 - \lambda)X_2] + (1 - \theta)X_2 \text{ for soem } \lambda \in (0, 1)$$

$$= \lambda^1 X_1 + (1 - \lambda^1)X_2, \qquad \text{where } \lambda^1 \in (0, 1)$$

and therefore we must have

$$\nabla f(\hat{X})^T (X_1 - \hat{X}) = (1 - \lambda^1) \nabla f(\hat{X})^T (X_1 - X_2)$$

Hence $\quad \nabla f(\hat{X})^T (X_1 - X_2) \leq 0. \qquad (9.13)$

which contradicts (9.10) implying that Ω is empty.

This completes the proof.

Strictly Quasiconvex Functions

A function f defined on a convex set S in R^n, is said to be strictly quasiconvex, if for each $X_1, X_2 \in S$,

$$f(X_1) < f(X_2) \Rightarrow f[\lambda X_1 + (1-\lambda)X_2] < f(X_2), \text{ for each } \lambda \in (0,1)$$

or equivalently,

$$f(X_1) \neq f(X_2) \Rightarrow f[\lambda X_1 + (1-\lambda)X_2] < \max\{f(X_1), f(X_2)\} \text{ for each } \lambda \in (0,1)$$

The function f is said to be strictly quasiconcave if –f is strictly quasiconvex.

Theorem 9.22: If f is a strictly quasiconvex function on a convex set S in R^n, then any local minimum point is also a global minimum point of f on S.

Proof: Let $\bar{X}$ be a local minimum point of a strictly quasiconvex function f on S in R^n.

Then there exists an $\in$-neighbourhood $N_\in(\bar{X})$ such that

$$f(\bar{X}) \leq f(X), \text{ for all } X \in N_\epsilon(\bar{X}) \tag{9.14}$$

Let us assume that there exists a point $\hat{X} \in S$ but not in $N_\epsilon(\bar{X})$, such that

$$f(\hat{X}) < f(\bar{X}). \tag{9.15}$$

Let $X_0 \in N_\epsilon(\bar{X})$, such that

$$X_0 = \lambda\bar{X} + (1-\lambda)\hat{X}, \quad 0 < \lambda < 1$$

By strict quasiconvexity of f, we then have

$$f(X_0) = f[\lambda\bar{X} + (1-\lambda)\hat{X}] < f(\bar{X})$$

i.e.
$$f(X_0) < f(\bar{X}) \tag{9.16}$$

which contradicts that

$$f(X_0) \geq f(\bar{X}) \qquad \text{by (9.14)}$$

Hence $\bar{X}$ is a global minimum point.

It should be noted that every strictly quasiconvex function need not be quasiconvex. For example [266] consider the function f defined on R^1, such that,

$$f(x) = \left.\begin{matrix} 1 & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{matrix}\right\} \tag{9.17}$$

f is then strictly quasiconvex on R' but is not quasiconvex, since for $x_1 = a$, $x_2 = -a$, $a > 0$

$$f(x_1) = f(x_2) = 0, \text{ but } f[\tfrac{1}{2}x_1 + \tfrac{1}{2}x_2] = f(0) = 1 > f(x_2).$$

If however, f is lower semicontinuous, then strict quasiconvexity implies quasiconvexity as is shown in the following theorem.

Theorem 9.23. [266] Let f be a lower semicontinuous function defined on a convex set S in R^n and strictly quasiconvex. Then f is quasiconvex on S.

Proof: Let $X_1, X_2 \in S$. If $f(X_1)<f(X_2)$, strict quasiconvexity of f implies quasiconvexity.

Now, suppose that $f(X_1)=f(X_2)$ and assume by contradiction that f is not quasiconvex. Hence there exist $X_1, X_2 \in S$, $X_0 \in (X_1, X_2)$ such that

$$f(X_1) = f(X_2) < f(X_0) \tag{9.18}$$

Since f is lower semicontinuous, there exists an $\bar{X} \in (X_1, X_0)$ such that

$$f(X_0) > f(\bar{X}) > f(X_1) = f(X_2). \tag{9.19}$$

It is easy to see that X_0 can be expressed as a convex combination of $\bar{X}$ and X_2, that is $X_0 \in (\bar{X}, X_2)$.

Now, since $f(X_2) < f(\bar{X})$, by strict quasiconvexity of f we have $f(X_0) < f(\bar{X})$ which contradicts (9.19)

Hence f is quasiconvex on S.

Explicitly quasiconvex

A function defined on a convex set S in R^n is explicitly quasiconvex if it is quasiconvex and if $f(X_1) < f(X_2)$ implies $f(X_0) < f(X_2)$., for all $X_1, X_2 \in S$, $X_0 \in (X_1, X_2)$ or equivalently if $f(X_1) \neq f(X_2)$ implies $f(X_0) < \max[f(X_1), f(X_2)]$. In other words, f is explicitly quasiconvex if it is quasiconvex and also strictly quasiconvex. However, a strictly quasiconvex function need not be explicity quasiconvex.[See (9.17)]

Further, an explicitly quasiconvex function on S is quasiconvex by definition, but every quasiconvex on S is not explicitly quasiconvex. This can be seen from the following example.

Let the function f be defined on R^1_+ as

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{if } \quad x > 1 \end{cases}$$

It can be easily seen that this function is quasiconvex on R^1_+

Now for $x_1 = 2$, $x_2 = 0$, $f(2) = 0 < f(0) = 1$,

but for $\lambda = 1/4$, $x_0 = \lambda x_1 + (1-\lambda)x_2 = 1/2$,

$$f(x_0) = f(1/2) = 1 \not< f(0) = 1$$

Thus the function f is quasiconvex in R^1_+ but not strictly quasiconvex and hence not explicitly quasiconvex.

Theorem 9.24. If f is a convex function defined on a convex set S in R^n, it is explicitly quasiconvex, that is, quasiconvex and also strictly quasiconvex but not conversely.

Proof: Let f be convex on S and $f(X_1) \le f(X_2)$.

Then for all $\lambda \in (0,1)$,

$$\begin{aligned} f[\lambda X_1 + (1-\lambda)X_2] &\le \lambda f(X_1) + (1-\lambda) f(X_2) \\ &= f(X_2) + \lambda [f(X_1) - f(X_2)] \\ &\le f(X_2). \end{aligned}$$

Hence f is quasiconvex on S.

If $f(X_1) < f(X_2)$, the above inequality is strict and thus f(X) is strictly quasiconvex.

To prove that the converse is not true, consider the following example.

Let $f(x) = -x^2$, $x \in R^1_+$. Then f is lower semicontinuous in R^1_+ (indeed continuous) and for $x_1, x_2 \ge 0$, $-x^2_1 < -x_2^2$ implies $x_1 > x_2$.

Thus for $x_0 = (1-\lambda)x_1 + \lambda x_2$, $0 < \lambda < 1$

$$f(x_0) - f(x_2) = -[(1-\lambda)x_1 + \lambda x_2]^2 + x_2^2$$

$$= [(1-\lambda)\, x_1 + (1+\lambda)x_2]\, [(1-\lambda)\, (x_1 - x_2)]$$
$$< 0.$$

Hence f is strictly quasiconvex and by theorem 9.23 it is quasiconvex on R^1_+

But for $x_1 = 2$, $x_2 = 0$ and $\lambda = ½$, we have

$$f[½.2 + ½\,.0] = f(1) = -1 > ½\, f(2) + ½\, f(0) = ½\,(-4) + ½\,(0) = -2$$

and therefore, f is not convex on R^1_+.

Theorem 9.25. If f is convex and nonnegative and g is concave and positive functions on a convex set S in R^n, then the function F defined by

$$F(X) = \frac{f(X)}{g(X)}$$

is explicitly quasiconvex on S.

Proof: Let $X_1, X_2 \in S$ and $X_0 = \lambda X_1 + (1-\lambda)X_2$, $\lambda \in (0,1)$.

$$0 \le f(X_0) \le \lambda\, f(X_1) + (1-\lambda)\, f(X_2).$$

and $$g(X_0) \ge \lambda\, g(X_1) + (1-\lambda)\, g(X_2) > 0$$

Then

$$F(X_0) = \frac{f(X_0)}{g(X_0)} \le \frac{\lambda f(X_1) + (1-\lambda) f(X_2)}{\lambda g(X_1) + (1-\lambda) g(X_2)}$$

$$= \frac{\lambda g(X_1) F(X_1) + (1-\lambda) g(X_2) F(X_2)}{\lambda g(X_1) + (1-\lambda) g(X_2)}$$

$$= \frac{\lambda g(X_1)}{\lambda g(X_1) + (1-\lambda) g(X_2)} F(X_1) + \frac{(1-\lambda) g(X_2)}{\lambda g(X_1) + (1-\lambda) g(X_2)} F(X_2)$$

$$= \theta\, F(X_1) + (1-\theta)\, F(X_2)\,,\ \theta \in (0,1).$$

$$\le \max\{F(X_1), F(X_2)\}$$

Hence F(X) is quasiconvex on S.

If $F(X_1) \ne F(X_2)$, the above inequality is strict implying that F is strictly quasiconvex.

F is therefore explicitly quasiconvex on S.

Pseudoconvex Function

Let a function f defined on some open set S in R^n be differentiable on S. The function f is then said to be pseudoconvex if for each $X_1, X_2 \in S$.

$$\nabla f(X_1)^T (X_2 - X_1) \ge 0 \text{ implies } f(X_2) \ge f(X_1).$$

or equivalently, if for each $X_1, X_2 \in S$,

$$f(X_2) < f(X_1) \text{ implies } \nabla f(X_1)^T (X_2 - X_1) < 0$$

The function f is said to be strictly pseudoconvex if for each $X_1, X_2 \in S$, $X_1 \ne X_2$,

$$\nabla f(X_1)^T (X_2 - X_1) \ge 0 \text{ implies } f(X_2) > f(X_1).$$

or equivalently, if for each $X_1, X_2 \in S, X_1 \neq X_2$,

$$f(X_2) \leq f(X_1) \text{ implies } \nabla f(X_1)^T (X_2 - X_1) < 0.$$

The function f is said to be pseudoconcave (strictly pseudoconcave) if $-f$ is pseudoconvex (strictly pseudoconvex).

Theorem 9.26. Let a function f be defined on an open convex set S in R^n. If f is pseudoconvex on S, then f is both strictly quasiconvex and quasiconvex on S but not conversely.

Proof: Suppose that f is pseudoconvex but not strictly quasiconvex. It then follows that there exists $X_1, X_2 \in S$ such that

$$f(X_1) < f(X_2) \tag{9.20}$$

and $$f(\hat{X}) \geq f(X_2) > f(X_1)$$

where $\hat{X} = \lambda X_1 + (1 - \lambda)X_2$, for some $\lambda \in (0, 1)$

By pseudoconvexity of f we then have

$$\nabla f(\hat{X})^T (X_1 - \hat{X}) < 0 \tag{9.21}$$

Since $$X_1 - \hat{X} = -\frac{1-\lambda}{\lambda}(X_2 - \hat{X}), \text{ we have}$$

$$\nabla f(\hat{X})^T (X_2 - \hat{X}) > 0 \tag{9.22}$$

and then by pseudoconvexity of f

$$f(X_2) \geq f(\hat{X})$$

and hence by (9.20),

$$f(X_2) = f(\hat{X}) \tag{9.23}$$

Now, since $\nabla f(\hat{X})^T (X_2 - \hat{X}) > 0$, there exists a point $\bar{X} = \mu\hat{X} + (1 - \mu) X_2$, $\mu \in (0, 1)$ such that

$$\nabla f(\hat{X})^T (\bar{X} - \hat{X}) > 0 \tag{9.24}$$

Hence, $$f(\bar{X}) > f(\hat{X}) = f(X_2). \tag{9.25}$$

Now, since $f(X_2) < f(\bar{X})$, by pseudoconvexity of f, we have

$$\nabla f(\bar{X})^T (X_2 - \bar{X}) < 0 \tag{9.26}$$

and since $$X_2 - \bar{X} = -\frac{\mu}{1-\mu}(\hat{X} - \bar{X}),$$

$$\nabla f(\bar{X})^T (\hat{X} - \bar{X}) > 0 \tag{9.27}$$

Similarly, $f(\hat{X}) < f(\bar{X})$ implies

$$\nabla f(\bar{X})^T(\hat{X} - \bar{X}) < 0 \tag{9.28}$$

which contradicts (9.27).

Hence f is strictly quasiconvex.

By theorem 9.23, then f is also quasiconvex.

To show that the converse is not true consider the example $f(x) = x^3$, $x \in R^1$. It can be easily seen that f is strictly quasiconvex but is not pseudoconvex on R^1.

Theorem 9.27. If a function f defined on an open convex set S in R^n is pseudoconvex, then a local minimum of f on S is also a global minimum.

Proof: By theorem 9.26, the function f is strictly quasiconvex function on S and therefore by theorem 9.22, a local minimum of f on S is also a global minimum.

Theorem 9.28. If a function f defined on an open convex set S in R^n is convex and differentiable, then f is pseudoconvex on S.

Proof: Let f be a differentiable convex function on S, then

$$f(X_2) - f(X_1) \geq \nabla f(X_1)^T (X_2 - X_1)$$

Then , $\nabla f(X_1)^T (X_2 - X_1) \geq 0$, implies that

$$f(X_2) \geq f(X_1)$$

and hence f is pseudoconvex on S.

Theorem 9.29. Let S be a convex set in R^n and let a function F on S be defined by

$$F(X) = \frac{C^T X + \alpha}{D^T X + \beta}$$

where $C, D \in R^n$ and α, β are scalars.

If $D^T X + \beta \neq 0$ for all $X \in S$, then F is both pseudoconvex and pseudoconcave on S.

Proof: Since $D^T X + \beta \neq 0$, for all $X \in S$, it is either positive for all $X \in S$ or negative for all $X \in S$. For if there exist $X_1, X_2 \in S$ such that $D^T X_1 + \beta > 0$ and $D^T X_2 + \beta < 0$, then for X, a convex combination of X_1 and X_2, $D^T X + \beta$ would be zero, contradicting the assumption.

To show that F is pseudoconvex on S, we need to show that for $X_1, X_2 \in S$, $\nabla f(X_1)^T (X_2 - X_1) \geq 0$ implies $F(X_2) \geq F(X_1)$.

$$\text{Now, } 0 \leq (X_2 - X_1)^T \Delta F(X_1) = (X_2 - X_1)^T \left[\frac{(D^T X_1 + \beta)C - (C^T X_1 + \alpha)D}{(D^T X_1 + \beta)^2} \right]$$

Since $(D^T X_1 + \beta)^2 > 0$, we have

$$(X_2 - X_1)^T[(D^T X_1 + \beta)\, C - (C^T X_1 + \alpha)D] \geq 0$$

or

$$(C^T X_2 + \alpha)\,(D^T X_1 + \beta) - (D^T X_2 + \beta)\,(C^T X_1 + \alpha) \geq 0$$

Since $D^T X + \beta$ and $D^T X_2 + \beta$ are both either positive or negative, dividing by $(D^TX_1 + \beta)(D^TX_2 + \beta) > 0$, we get

$$\frac{C^TX_2 + \alpha}{B^TX_2 + \beta} \geq \frac{C^TX_1 + \alpha}{D^TX_1 + \beta}$$

Hence F is pseudoconvex

Similarly, it can be shown that $\nabla f(X_1)^T(X_2 - X_1) \leq 0$ implies $F(X_2) \leq F(X_1)$ and therefore F is pseudoconcave.

Corollary 9.2: The function F is strictly quasiconvex, quasiconvex and also strictly quasiconcave, quasiconcave.

Proof: Follows from theorem 9.26.

The ordering among the generalized convex functions, we have discussed may now be given as follows.

(a) If a function is defined on an open convex set and is differentiable, we have the relations
strictly convex $\Rightarrow$ convex $\Rightarrow$ pseudoconvex
$\Rightarrow$ explicitly quasiconvex
$\Rightarrow$ quasiconvex.

(b) If a function is defined on an open convex set but is not differentiable, we have
strictly convex $\Rightarrow$ convex $\Rightarrow$ explicitly quasiconvex
$\Rightarrow$ quasiconvex.

9.4. Exercises

1. Show that a function f on R^n is a linear function: if and only if f is both convex and concave.
2. Classify the following functions as convex and concave.
   (a) x^2 , $x \in R^1$
   (b) $|X|$, $x \in R^1$
   (c) e^x , $x \in R^1$
   (d) e^{-x^2} , $x \in R^1$
   (e) $\log x$, $x > 0$
3. If f_i, $i \in I$ are convex functions on a convex set S in R^n and b_i are scalars, then show that the set
   $R = \{X \mid f_i(X) \leq b_i, i \in I\}$ is convex and closed.
4. Show that a function f defined on a convex set S in R^n is concave if and only if the hypograph of f defined by the set
   $$S_{hyp} = \{(X, \xi) \mid X \in S, \xi \in R^1, f(X) \geq \xi\}$$
   is convex.
5. Consider the function $f(X) = C^TX + \frac{1}{2} X^TBX$, where C is an n-vector and B is an $n \times n$ symmetric matrix. Show that f is concave on R^n if and only if B is negative semidefinite.
6. Let f be a convex function on a convex set S in R^n and g be a

nondecreasing convex function on R^1. Show that the composite function h defined by $h(X) = g[f(X)]$ is convex on S.

7. Let $f_1, f_2, .. f_m$ be convex functions on a convex set S in R^n. Show that the function f defined by $f(X) = \max\{f_1(X), f_2(X), .. f_m(X)\}$ is convex on S.
8. Let g be a concave function on $S = \{X: X \in R^n, g(X) > 0\}$. Show that the function f defined by $f(X) = 1/g(X)$ is convex on S.
9. Show that the function f defined by

$$f(X) = \sum_{i-1}^{m} c_i \exp\left(\sum_{j=1}^{n} a_{ij} x_j\right)$$

where a_{ij} and $c_i > 0$ are scalars, is convex on R^n.

10. Let $f(X, \alpha)$ be convex on a convex set S for each α, $a \leq \alpha \leq b$. If the function $\phi(\alpha) \geq 0$, then prove that the function

$$F(X) = \int_a^b f(X, \alpha).\varphi(\alpha) d\alpha$$

is convex on S.

11. Show that no polynomial of odd degree (≥ 3) is a convex function on R^1.
12. Show that the function f defined by

$$f(X) = \ln\left(\sum_{i=1}^{n} e^{x_i}\right)$$

is a convex function on R^n

13. Let the function ϕ be defined on $S = \{X \in R^n \mid X > 0\}$. Then show that

$$\phi(X) = (x_1)^{\alpha_1}(x_2)^{\alpha_2}...(x_n)^{\alpha_n}$$

where $\alpha_i \geq 0$, $(i = 1,2,..n)$ is explicitly quasiconcave.

14. If functions f and g defined on a convex set S is R^n are both concave and nonnegative, then prove that $\phi(X) = f(X).g(X)$ is explicitly quasiconcave on S.
15. Let the functions f and g be defined on a convex set S in R^n. Show that the function ϕ defined by $\phi(X) = f(X)/g(X)$ is explicitly quasiconvex if any one of the following conditions holds true.
    (a) f is concave and nonnegative, g is concave and negative on S.
    (b) f is convex and nonpositive, g is convex and positive on S.
    (c) f is concave and nonpositive, g is convex and negative on S.
16. If in Q15, S is open and f and g are differentiable on S, then show that ϕ is pseudoconvex.

PART – 2

LINEAR PROGRAMMING

This Page Intentionally Left Blank

Chapter 10

Linear Programming Problems

As we have seen in Chapter 1, a wide class of decision problems can be formulated as linear programming problems where we are to maximize or minimize a linear function subject to certain inequality constraints. In this chapter we give equivalent formulations of the general linear programming problem and discuss some of its fundamental properties which make it possible to develop the most widely used method—the simplex method, for solving the problem (discussed in Chapter 11).

10.1. The General Problem

The general linear programming problem is to find values of a set of variables $x_1, x_2, .. x_n$ which optimizes (maximizes or minimizes) a linear function

$$z = \sum_{j=1}^{n} c_j x_j \tag{10.1}$$

Subject to

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i = 1, 2, ...m_1 \tag{10.2}$$

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = m_1 + 1, ..m_2 \tag{10.3}$$

$$\sum_{j=1}^{n} a_{ij} x_j \geq b_i, \quad i = m_2 + 1, ..m \tag{10.4}$$

$$x_j \geq 0, \quad j = 1, 2, ..n_1 \tag{10.5}$$

$$x_j \text{ arbitrary}, j = n_1, + 1, ..n \tag{10.6}$$

where all the c_j, a_{ij} and b_i are assumed to be known constants.

10.2. Equivalent Formulations

It has been shown in the examples of Chapter 1 that the linear programming problem can be presented in variety of forms. It may be a problem of maximization or minimization under the conditions with $\leq$, $=$, and/or $\geq$ type of inequalities.

Different aspects of development in linear programming are based on different forms of the problem. It can however be easily shown that these different forms are equivalent to each other and the results attained with one form of the problem, therefore are valid for all types of linear programming problem. In all the cases, the condition of nonnegativity of all the variables are imposed as the nonnegativity conditions $x_j \geq 0$, play a special role in linear programming.

Three equivalent formulations of the problem are considered which are also equivalent to the general form: the canonical form, which is primarily used in the development of the theory of duality; the standard form, which is used in the development of the methods of computation and the mixed form which contains both the conditions of equalities and inequalities, is used to represent some practical situations.

Canonical Form

$$\text{Maximize} \quad z = \sum_{j=1}^{n} c_j x_j$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i = 1, 2, ...m$$

$$x_j \geq 0, \quad j = 1, 2, ..n.$$

In matrix notation, it can be expressed as

$$\begin{aligned} &\text{Maximize} && z = c^T X \\ &\text{Subject to} && AX \leq b \\ & && X \geq 0 \end{aligned} \tag{10.7}$$

where $c^T = (c_1, c_2, ..c_n)$; $b^T = (b_1, b_2, ..b_m)$;
$X^T = (x_1, x_2 .. x_n)$ are row vectors and
$A = (a_{ij})$ is an $m \times n$ matrix

Standard Form

$$\text{Minimize} \quad z = \sum_{j=1}^{n} c_j x_j$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = 1, 2, ...m$$

$$x_j \geq 0, \quad j = 1, 2, ..n.$$

and in matrix notation it can be stated as,

$$\begin{aligned} &\text{Minimize } z = c^T X \\ &\text{Subject to} \quad AX = b \\ &\qquad\qquad\quad X \geq 0 \end{aligned} \tag{10.8}$$

Mixed Form

$$\text{Maximize} \quad z = \sum_{j=1}^{n} c_j x_j$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad i = 1, 2, \ldots m_1$$

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = m_1 + 1, \ldots m.$$

$$x_j \ge 0, \quad j = 1, 2, \ldots n.$$

In matrix notation, we have

$$\begin{aligned} \text{Maximize} \quad & z = c^T X \\ \text{Subject to} \quad & A_1 X \le b_1 \\ & A_2 X = b_2 \\ & X \ge 0 \end{aligned} \tag{10.9}$$

where c^T, A_1, A_2, b_1 and b_2 are appropriate matrices or vectors.

The equivalence of the above different forms may easily be shown by the following elementary operations.

Operation 1: Since minimum f(X) = – maximum [–f(X)], where f(X) is a linear function, any problem of maximization may be expressed as a problem of minimization and vice-versa.

Operation 2: An inequality in one direction may be changed to an inequality in the opposite direction by multiplying both sides of the inequality by –1.

Thus $a_1 x_1 + a_2 x_2 \le b_1$,

is equivalent to $-a_1 x_1 - a_2 x_2 \ge -b_1$.

Similarly, $d_1 x_1 + d_2 x_2 \ge b_2$,

is equivalent to $-d_1 x_1 - d_2 x_2 \le -b_2$.

Operation 3: A variable of arbitrary sign can always be expressed as the difference between two nonnegative variables. Thus if a variable x is unconstrained in sign, it may be replaced by $(x^+ - x^-)$, where x^+ and x^- are both nonnegative variables, i.e. $x^+ \ge 0$, $x^- \ge 0$

Operation 4: An equation

$$a_1 x_1 + a_2 x_2 = b$$

is equivalent to the two inequalities

$$a_1 x_1 + a_2 x_2 \le b$$
$$-a_1 x_1 - a_2 x_2 \le -b.$$

Operation 5: An inequality of less than or equal to type, may be changed to an equation by adding an additional nonnegative variable called the *slack variable*.

For example, $a_1x_1 + a_2x_2 \le p_1$,

may be replaced by $a_1x_1 + a_2x_2 + x_3 = p_1$,

where $x_3 \ge 0$ is the additional variable known as slack variable.

Similarly, an inequality of greater than or equal to type may be changed to an equation by subtracting, an additional nonnegative variable called the *surplus variable.*

Thus, $d_1x_1 + d_2x_2 \ge q$

may be replaced by, $d_1x_1 + d_2x_2 - x_3 = q$

where $x_3 \ge 0$ is the additional variable called the surplus variable.

It should be noted that the slack and surplus variables are given zero coefficients in the linear function to be optimized.

Using the preceding operations, the following transformations from one form to another may easily be obtained.

General form	to	Mixed form (by operations 1, 2 and 3)
Standard form	to	Canonical form (by operations 1 and 4)
Mixed form	to	Canonical form (by operations 4)
Mixed and Canonical form	to	Standard form (by operations 1 and 5)

10.3. Definitions and Terminologies

The linear function to be optimized (minimized or maximized) is called the *objective function.*

Mathematically, the conditions of nonnegativity do not differ from other conditions but since nonnegativity conditions play a special role in the development of the methods of computations for linear programming problems, (c.f. Chapter 11), $x_j \ge 0$ are called *nonnegative restrictions* or simply *restrictions* and the other conditions are referred to as *constraints.*

A set of values of the variables which satisfies all the constraints and the nonnegative restrictions is called a *feasible solution.*

A feasible solution which optimizes the objective function if it exists, is called an *optimal solution,* that is, an optimal solution is a feasible solution in which all variables are finite, which optimizes the objective function.

A set of feasible solutions is said to be a *feasible region,* a *feasible set* or a *constraint set.*

The vector 'b' on the right hand side of the constraints is called the *requirement vector* and the columns of A are called *activity vectors.*

A linear programming problem is often referred to as a *linear program.*

10.4. Basic Solutions of Linear Programs

By a basic solution of a linear programming problem, we mean a basic solution of the system of linear constraints of the problem in standard form. (c.f. section 7.3).

Consider the problem,

$$\begin{aligned} &\text{Minimize} && z = c^T X \\ &\text{Subject to} && AX = b. \\ &&& X \geq 0 \end{aligned} \tag{10.10}$$

where $c^T = (c_1, c_2, .. c_n)$, $b^T = (b_1, b_2, .. b_m)$ and A is an $m \times n$ matrix.

It is assumed that $n > m$ and $r(A) = m$, so that the system of equations are nonredundant and the problem has more than one, in fact an infinite number of solutions.

Suppose that any m linearly independent columns from A are selected to form a matrix B. B is then a nonsingular submatrix of order m of A and is called a basis of the linear system or a basis matrix of the problem.

The matrix A can then be written as

$$A = [B, N]$$

where B is an $m \times m$ matrix, and N is an $m \times (n-m)$ matrix.

Further, suppose that X_B is the vector of variables associated with the columns of B and X_N be the vector of the remaining variables. The system of equations can then be expressed as,

$$BX_B + NX_N = b, \tag{10.11}$$

and all the solutions of the linear system can be generated by assigning arbitrary values to X_N. If we set $X_N = 0$, we have $BX_B = b$ and since B is nonsingular, a unique solution $X_B = B^{-1}b$ can be obtained. $[X_B, 0]$, is then called a basic solution of the system $AX = b$ or the linear program. For convenience, $X_B = B^{-1}b$, is often referred to as the basic solution, understanding that the basic solution actually is $[X_B, 0]$, with (n–m) remaining variables having zero values.

The m components of X_B are called basic variables and the remaining (n–m) variables (the components of X_N) are known as nonbasic variables.

The maximum number of basic solutions is equal to $\binom{n}{m} = \frac{n!}{m!(n-m)!}$, which is the number of nonsingular matrices possible to extract from A.

A basic solution is called a basic feasible solution, if it satisfies the nonnegative restrictions also, that is, if all the basic variables are nonnegative.

A basic feasible solution is called a nondegenerate basic feasible solution if all the basic variables are strictly positive. If one or more of the basic variables are zero, it is known as degenerate basic feasible solution.

10.5. Fundamental Properties of Linear Programs

In this section we shall discuss some fundamental properties of linear programming problems which lead to the development of the simplex method devised by Dantzig [96] for solving linear programs.

Theorem 10.1. The set of feasible solutions to a linear programming problem is a closed convex set bounded from below.

Proof: Every feasible solution which must satisfy the constraints of the problem (10.9) belongs to the intersection of the closed half-spaces with the hyperplanes and with the nonnegative orthant $X \geq 0$. All these are closed convex sets and hence their intersection is. Furthermore, the set is bounded from below since $X \geq 0$.

Let S denote the set of feasible solutions to a linear programming problem. If S is nonempty, S may either be a convex region unbounded in some direction or a convex polyhedron. If S is unbounded, the problem has a feasible solution but the value of the objective function might be unbounded and if S is a convex polyhedron, then the problem has a feasible solution with an optimal value of the objective function.

Let us assume that S is a convex polyhedron.

Note that the objective function of the linear program which is to be optimized is a hyperplane

$$c^T X = z$$

where z is a parameter.

This hyperplane is moved parallel to itself over the convex set of the feasible solutions S until $z = z_0$, the minimum value of z on S. (for minimization problem)

Thus X_0 is an optimal solution of the minimization problem if and only if

$$c^T X_0 = z_0,$$
$$c^T X \geq z, \text{ for every } X \in S.$$

which implies that X_0 is a boundary point of S. By definition, then $c^T X = z_0$ is a supporting hyperplane which contains at least one point of S on the boundary. Thus, there may be more than one optimal solution of the linear program and at least one extreme point of S is an optimal solution. (See Chapter 8)

This can also be seen from theorem 10.2, which follows.

Theorem 10.2: Let the set S of feasible solutions to the linear programming problem be nonempty, closed and bounded. Then the problem has an optimal solution which is attained at an extreme point of S.

Proof: Since S is a nonempty, closed, bounded convex set and the objective function is linear, an optimal solution of the problem exists. S is a convex polyhedron and from corollary 10.2, it follows that S has a finite number of extreme points. Let these extreme points be denoted by $X_1^*, X_2^*, .. X_p^*$. Let X_0 be a feasible point minimizing (for the minimization problem) the objective function $c^T X$ on S, so that

$$c^T X_0 \leq c^T X \text{ , for all } X \in S \tag{10.12}$$

If X_0 is an extreme point, the theorem is true.

Suppose X_0 is not an extreme point of S and hence X_0 can be expressed as a convex combination of the extreme points of S (see Theorem 8.12). Thus, we have,

$$X_0 = \sum_{i=1}^{P} \lambda_i X_i^*, \quad \text{for } \lambda_i \geq 0, \ \sum_{i=1}^{P} \lambda_i = 1$$

Then

$$c^T X_0 = c^T \left(\sum_{i=1}^{P} \lambda_i X_i^* \right) = \sum_{i=1}^{P} \lambda_i c_i^T X_i^*$$

$$\geq c^T X_r^* \sum_{i=1}^{P} \lambda_i$$

$$= c^T X_r^* \tag{10.13}$$

where $= c^T X_r^* = \operatorname*{Min}_i c^T X_i^*$,

$$\text{By (10.12)}, \quad c^T X_0 = c^T X_r^*$$

and therefore, there is an extreme point X_r^* of S, at which the objective function attains its minimum.

Theorem 10.3: The set of optimal solutions to a linear programming problem is convex.

Proof: Let K be the set of optimal solutions and X_1^0, $X_2^0 \in K$. X_1^0, X_2^0 are then optimal solutions to the problem and $c^T X_1^0 = c^T X_0 = z_0 = \min z$.

Since X_1^0, X_2^0 are feasible solutions, X_1^0, X_2^0 belong to S, the set of feasible solutions of the problem and since S is convex,

$$\lambda X_1^0 + (1-\lambda) X_2^0 \in S, \quad \text{for all } \lambda, \ 0 \leq \lambda \leq 1.$$

Also,

$$c^T \left[\lambda X_1^0 + (1-\lambda) X_2^0 \right] = \lambda c^T X_1^0 + (1-\lambda) c^T X_2^0 = z_0$$

Hence,

$$\lambda X_1^0 + (1-\lambda) X_2^0 \in K, \quad \text{for all } \lambda, 0 \leq \lambda \leq 1$$

and therefore K is convex.

Corollary 10.1: If there exist more then one optimal solutions, then an infinite number of optimal solutions can be obtained.

Proof: Every convex combination of the known optimal solutions is optimal.

Now, a convex polyhedron has a finite number of extreme points (corollary 10.2) and hence if we had an analytical method to find the extreme points of the feasible set of the problem, we would only need to examine a finite number of feasible solutions to find an optimal solution. That, it is indeed possible to have such a method can be seen from the following theorems.

Theorem 10.4: If a standard linear program has a feasible solution it also has a basic feasible solution.

Proof: Let the linear constraints of the standard linear program be given by,

$$AX = b,$$

where A is an $(m \times n)$ matrix and rank $(A) = m$

Let $X = (x_1, x_2, ..x_n)^T$ be a feasible solution and suppose that the variables are so numbered that the positive components of X are the first k components and the last (n–k) components have the value zero, so that

$$a_1x_1 + a_2x_2 + .. + a_kx_k = b \tag{10.15}$$

where a_j is the j th column of A and $x_j > 0$, $j = 1,2,...k$.

Two cases may now arise

Case 1. The vectors $a_1, a_2, ..a_k$ associated with the positive variables are linearly independent. This implies $k \le m$.

If $k = m$, the assumed feasible solution is a nondegenerate basic feasible solution.

If $k < m$, a set of additional vectors say, $a_{k+1}, a_{k+2}, ..a_m$ can always be found (since $r(A) = m$) so that $a_1, a_2, ..a_k, a_{k+1}, ..a_m$ form a linearly independent set. Then,

$$a_1x_1 + .. + a_mx_m = b,$$

where $\quad x_j > 0, j = 1,2..k \quad (10.16)$

and $\quad x_j = 0, j = k+1, ..m.$

Thus we have a degenerate basic feasible solution with (m–k) of the basic variables having values zero.

Case 2. Let $a_1, a_2, ..a_k$ be linearly dependent. This means that there exist constants $\alpha_1, \alpha_2..\alpha_k$, not all zero, such that

$$\alpha_1a_1 + \alpha_2a_2 + ... + \alpha_ka_k = 0. \tag{10.17}$$

Multiplying this equation by –1, if necessary, we can always assume that one of the α_j is positive. Now, multiplying (10.17) by a scalar θ and subtracting it from the equation (10.15), we get,

$$\sum_{j=1}^{k}(x_j - \theta\alpha_j)\, a_j = b \tag{10.18}$$

and if $x_j - \theta\, \alpha_j$, $j = 1, 2, ..k$ are all nonnegative, we have a new feasible solution. Since $x_1, x_2, .. x_k$ are all positive, we can always find a $\theta > 0$, for which the coefficients of the vectors a_j in (10.18) remain nonnegative.

$$\text{Let } \theta = \frac{x_r}{\alpha_r} = \underset{j}{\text{Min}}\left\{\frac{x_j}{\alpha_j}, \alpha_j > 0\right\}. \tag{10.19}$$

and then
$$x'_j = x_j - \frac{\alpha_j}{\alpha_r}x_r, \quad j = 1,2,..k \tag{10.20}$$

are all nonnegative

It is thus clear that $\quad (x_1', x_2', ... 0; 0...0)^T \quad (10.21)$

constitute a new feasible solution with at most (k–1) positive variables. If the vectors associated with these positive variables are linearly dependent, we repeat the forgoing procedure and continue the process until the vectors associated with the positive variables are linearly independent. Then case 1 applies and a basic feasible solution can be obtained.

Theorem 10.5. Let S be the convex polyhedron generated by the set of feasible solutions to the standard linear programming problem, i.e. $S = \{x \mid Ax = b, x \geq 0\}$, where A is an m x n matrix and $r(A) = m$. Then X is an extreme point of S if and only if X is a basic feasible solution of the problem.

Proof: Suppose that X is an extreme point of S. Let the nonzero components of X be the first k components so that

$$a_1x_1 + a_2x_2 + .. + a_kx_k = b \qquad (10.22)$$

Assume that the vectors $a_1,a_2,..a_k$ are linearly dependent. Then there exist constants $\alpha_1,\alpha_2..\alpha_k$ not all zero, such that

$$\alpha_1a_1 + \alpha_2a_2 + \quad + \alpha_ka_k = 0. \qquad (10.23)$$

Now, for some $\theta > 0$, multiplying (10.23) by θ and adding and subtracting the result from (10.22) we have,

$$\sum_{j=1}^{k}(x_j + \theta\alpha_j)\, a_j = b \qquad (10.24)$$

and

$$\sum_{j=1}^{k}(x_j - \theta\alpha_j)\, a_j = b \qquad (10.25)$$

Let $X_1 = (x_1 + \theta\,\alpha_1, x_2 + \theta\,\alpha_2, .. x_k + \theta\,\alpha_k, 0 \ldots 0)^T$ (10.26)

and $X_2 = (x_1 - \theta\,\alpha_1, x_2 - \theta\,\alpha_2, .. x_k - \theta\,\alpha_k, 0 \ldots 0)^T$ (10.27)

Since $x_j > 0$, $j = 1,2..k$, it is always possible to find a $\theta > 0$, such that the first k components of both X_1 and X_2 are positive. For example, if we take

$$\theta \text{ as } 0 < \theta < \operatorname*{Min}_{j} \frac{x_j}{|\alpha_j|}, \quad \alpha_j \neq 0, \quad j = 1,2,..k \qquad (10.28)$$

then, $x_j + \theta\,\alpha_j$ and $x_j - \theta\,\alpha_j$ are both positive for $j = 1,2..k$.

Hence X_1 and X_2 are feasible solutions of the problem ($X_1 \neq X_2$).

But $X = \frac{1}{2}X_1 + \frac{1}{2}X_2$, i.e. X is a convex combination of two distinct feasible solutions different from X and therefore contradicts that X is an extreme point of S. Hence $a_1, a_2,..a_k$ are linearly independent.

Since $r(A) = m$, if $k < m$, we can always include additional $(m-k)$ vectors say, $a_{k+1},..a_m$ so that the set of vectors $a_1,a_2..a_k, a_{k+1},..a_m$ are linearly independent. $B = (a_1, a_2, ..a_k, a_{k+1}..a_m)$ then forms a basis matrix and $X = (x_1, x_2..x_k,0..0)^T$ is a basic feasible solution.

Now, suppose that X^B is a basic feasible solution to the problem, and let

$$X^B = (x_1^B, x_2^B, .. x_m^B, 0..0)^T \tag{10.29}$$

(after relabelling the variables if necessary).

The values of the components of X^B are given by the solution of

$$\sum_{j=1}^{m} a_{ij} x_j = b_i, \quad i = 1, 2, ..m \tag{10.30}$$

Since a_1, a_2..a_m are linearly independent vectors forming the basis, this solutions must be unique.

Suppose that X^B is not an extreme point of S. Then there exist two distinct feasible solutions.

$$X^{F1} = (x_1^{F1}, x_2^{F1} .. x_n^{F1})^T \tag{10.31}$$

and
$$X^{F2} = (x_1^{F2}, x_2^{F2} .. x_n^{F2})^T$$

$$X^{F1} \neq X^{F2}$$

so that we have,

$$x_j^B = \lambda x_j^{F1} + (1-\lambda) x_j^{F2}, \quad j = 1, 2, ..m \tag{10.32}$$

$$0 = \lambda x_j^{F1} + (1-\lambda) x_j^{F2}, \quad j = m+1, ..n. \tag{10.33}$$

for some λ, $0 < \lambda < 1$.

Since $x_j^{F1} \geq 0$, $x_j^{F2} \geq 0$ for all j and $0 < \lambda < 1$, (10.33) implies that

$$x_j^{F1} = x_j^{F2} = 0, \qquad \text{for} \quad j = m+1, ...n.$$

and the values of the remaining components of X^{F1} and X^{F2} are also determined by the solution of (10.30). As this solution is unique, we must have

$$x_j^B = x_j^{F1} = x_j^{F2}, \qquad \text{for all } j = 1, 2, ..n.$$

which implies that X^B cannot be expressed as a convex combination of two distinct feasible solutions. Hence X^B is an extreme point of S.

Corollary 10.2. The number of extreme points of the convex polydron $S = \{X \mid AX = b, X \geq 0\}$ where A is an $m \times n$ matrix and $r(A) = m$ is finite.

Proof: The number of extreme points of S is at most $\binom{n}{m}$, which is the maximum number of ways to choose m independent columns from A to from the basis matrix B.

10.6. Exercises

1. A manufacturer produces three products A, B and C which are to be

processed by three machines M_1, M_2 and M_3. The machine time in hours per unit produced, the total machine times available on machines and profit per unit from the products are as follows:

Products	*Machine time in hours*			*Profit per unit*
	M_1	M_2	M_3	
A	3	0	1	8
B	8	4	2	20
C	2	3	0	6
Machine hours. Available	250	150	50	

Formulate the problem as a linear programming problem to determine the amount of each product to be produced to maximize the total profit.

2. A manufacturer of biscuits is considering four types of gift packs containing three types of biscuits, Orange Cream (OC), Chocolate Cream (CC) and Wafers (W). Market research study conducted recently to assess the preferences of the consumers shows the following types of assortments to be in good demand:

Assortments	*Contents*	*Selling price per unit in dollar*
	Not less than 40% of OC	
A	Not more than 20% of CC	10
	Any quantity of W	
	Not less than 20% of OC	
B	Not more than 40% of CC	12.5
	Any quantity of W	
	Not less than 50% of OC	
C	Not more than 10% of CC	11
	Any quantity of W	
D	No restrictions	6

For the biscuits, the manufacturing capacity and costs are given below:

Biscuit variety	*Plant Capacity units per day*	*Manufacturing cost in dollar per unit*
OC	200	4
CC	200	4.5
W	150	3.5

Formulate a linear programming model to find the production schedule which maximizes the profit assuming that there are no market restrictions.

3. An oil refinery has two types of crude oil A and B and wants to optimally mix them to produce fuel oil and gasoline from the data given below:

Mix	*Input in units*		*Output in units*	
	Crude A	*Crude B*	*Fuel oil*	*Gasoline*
M_1	2	3	3	4
M_2	3	2	6	2
M_3	3	3	3	2

The amount of crude oil available are 15 units of A and 20 units of B. The fuel oil sells for \$3 per unit and gasoline sells for \$18 per unit. Formulate the problem to decide on the production scheme to maximize the profit.

4. A paper company produces rolls of paper of standard width of 200 cm and the following orders have been received:

Width in cm:	65	60	55
No. of rolls:	250	150	100

To satisfy these orders the rolls are to be cut in such a way that the wastage of paper is minimum.

Obtain a linear programming formulation of the problem to determine an optimal cutting pattern.

5. A ship has three cargo holds: forward, centre and aft. The capacity limits are as follows:

Cargo holds	*Weight capacity in tonnes*	*Volume capacity in cubic meters.*
Forward	2,000	100,000
Centre	3,000	135,000
Aft	1,500	30,000

The following cargoes are offered; the ship owners may accept all or any part of each commodity:

Commodity	*Amount (tonnes)*	*Volume per tonne*	*Profit per tonne (\$)*
A	6,000	60	60
B	4,000	50	80
C	2,000	25	50

In order to preserve the trim of the ship, the weight in each hold must be proportional to the capacity in tonnes. The problem facing the owners of the ship is to decide how much of each commodity to accept in order to maximize the total profit.

Formulate a linear programming model for this problem.

6. A television set manufacturing company produces two types of television sets—color and black-and-white. A color T.V. requires 16 man-hours and a black-and-white requires 12 man-hours and the total number of man-hours available is 4,000 per month. A market research indicates that at most 800 units of color T.V. and 3200 units of black-and-white T.V. can be sold per month. The unit profits of the color and black-and-white T.V.'s are \$48 and \$24 respectively. The company wants to determine the number of units of the two types of T.V.'s that are to be produced so that the total profit is maximum. Formulate the problem as a linear program.

7. A company manufactures products A, B, C and D which are processed by planer, milling, drilling and assembly departments. The requirements per unit of products in hours and the capacities of the departments and profits per unit of products are as follows:

Product	*Requirements per unit (hours)*				*Profit per*
	Planner	*Milling*	*Drilling*	*Assembly*	*unit in dollars*
A	0.5	2.0	0.5	3.0	8
B	1.0	1.0	0.5	1.0	9
C	1.0	1.0	1.0	2.0	7
D	0.5	1.0	1.0	30	6
Capacity (in hours)	1,800	2,800	3,000	6,000	

Minimum sales requirements of the products A, B, C and D are 100 units, 600 units, 500 units and 400 units, respectively.

Formulate the problem of determining the number of units of the products to be produced to maximize profit as a linear programming problem.

8. Express the following linear programs in standard form.

(a) Minimize $3x_1 + 2x_2 - 4x_3$

Subject to
$$5x_1 - x_2 + 3x_3 \geq 8$$
$$-4x_1 + 2x_2 + 5x_3 \geq 4$$
$$2x_1 + 5x_2 - 6x_3 \geq 5.$$
$$x_1, x_2, x_3 \geq 0.$$

(b) Maximize $3x_1 - 6x_2 - 2x_3$

Subject to
$$3x_1 - 2x_2 + x_3 \leq 5$$
$$4x_1 + 6x_2 - 3x_3 \geq -2$$

$$x_1 + 2x_2 + 2x_3 \geq 11$$
$$x_1, x_2, x_3 \geq 0$$

(c) Maximize $5x_1 + 2x_2 + 3x_3$

Subject to
$$x_1 - x_2 - x_3 \leq 5$$
$$9x_1 + 6x_2 - 7x_3 \geq -4$$
$$4x_1 + x_2 + x_3 = 10$$
x_1 unrestricted; $x_2, x_3 \geq 0$.

(d) Maximize $3x_1 + x_2$

Subject to
$$x_1 - x_2 \leq 4$$
$$x_1 + x_2 \leq 2$$
$x_1 \geq 0$, x_2 unrestricted

9. Convert the following linear program into canonical form.

Maximize $3x_1 + 3x_2 + 2x_3$

Subject to
$$2x_1 + 3x_2 + 4x_3 = 8$$
$$x_1 + x_2 + 3x_3 = 6$$
$$x_1 + 2x_2 + x_3 = 2$$
$$x_1, x_2, x_3 \geq 0.$$

10. Solve the following problems graphically.

(a) Maximize $2x_1 + 3x_2$

Subject to
$$x_1 + x_2 \leq 3$$
$$x_1 - 2x_2 \leq 1$$
$$x_2 \leq 2$$
$$x_1, x_2 \geq 0.$$

(b) Minimize $x_1 + 2x_2$

Subject to
$$x_1 + x_2 \geq 1$$
$$2x_1 + x_2 \leq 9$$
$$x_1 \leq 4$$
$$x_1, x_2 \geq 0.$$

(c) Maximize $x_1 + 3x_2$

Subject to
$$x_1 + x_2 \geq 3$$
$$-x_1 + x_2 \leq 2.$$
$$x_1 - 2x_2 \leq 2$$
$$x_1, x_2 \geq 0$$

11. Obtain basic feasible solutions of the following linear programs and determine their optimal solutions. Verify them graphically

(a) Maximize $2x_1 + 3x_2$

Subject to $2x_1 + x_2 \leq 2$

$$x_1 + 3x_2 = 5$$
$$x_1, x_2 \geq 0$$

(b) Maximize $5x_1 + 2x_2$

Subject to
$$2x_1 + x_2 \leq 4$$
$$2x_1 + 3x_2 \geq 6$$
$$x_1, x_2 \geq 0$$

(c) Maximize $3x_1 + 4x_2$

Subject to
$$x_1 + x_2 \leq 4$$
$$2x_1 + 5x_2 \leq 10$$
$$x_1, x_2 \geq 0.$$

12. Find the extreme points of the set of solutions of the following system:

$$8x_1 + 3x_2 - 4x_3 + x_4 = 10$$
$$4x_2 - 2x_3 + x_5 = 12$$
$$2x_1 - x_2 + 3x_3 + x_6 = 7$$
$$x_j \geq 0,$$
$$j = 1,2,...6.$$

13. Show graphically that the maximum or minimum values of the objective function of the following problem are same.

Maximize (or Minimize) $3x_1 + 5x_2$

Subject to
$$3x_1 + 2x_2 \geq 3$$
$$x_1 + x_2 \leq 6$$
$$x_1 \geq 3$$
$$x_2 \geq 3$$
$$x_1, x_2 \geq 0$$

14. Construct an example of linear program
   (a) which has no feasible solution.
   (b) which is feasible but has no optimal solution.

CHAPTER 11

Simplex Method: Theory and Computation

11.1. Introduction

From the theorems discussed in Chapter 10, we know that

(a) an optimal solution of a linear program, if it exists, is attained at one of the extreme points of the convex set of the feasible solutions,
(b) every extreme point of the set corresponds to a basic feasible solution, and
(c) if there exists a feasible solution to the problem, there also exists a basic feasible solution.

It is therefore, clear that we need only to investigate extreme point solutions and hence only basic feasible solutions to obtain an optimal solution of the problem. Now, the maximum number of bases for a system of m equations in n unknowns is equal to $\binom{n}{m}$, which increases rapidly with the increase of m and n. For large m and n it would therefore be an impossible task to evaluate all the extreme points and select one that optimizes the objective function. What we need therefore, is a computational scheme which selects in a systematic way, a subset of extreme points where the value of the objective function continues improving and ultimately converges to an optimal solution if it exists. The simplex method devised by G.B. Dantzig [96] is such a procedure. The procedure finds an extreme point and checks whether it is optimal or not. If it is not, the procedure finds a neighbouring extreme point whose corresponding value of the objective function is at least as good as the one just obtained. The process is continued until an optimal solution is obtained or it indicates that the problem is unbounded. In each step, the procedure puts many extreme points out of consideration and it makes it possible to find an optimal solution in a finite number of steps. The simplex method also makes it possible to find whether the problem has no feasible solution.

11.2. Theory of the Simplex Method

To develop the simplex method, we proceed with the standard form of the linear programming problem

$$\begin{aligned} &\text{Minimize } z = c^T X. \\ &\text{Subject to } AX = b. \\ &\qquad X \geq 0. \end{aligned} \tag{11.1}$$

where A is an m x n matrix, m < n and r(A) = m.

We assume that the problem is feasible and that we are given a basic feasible solution

$$X_0 = (x_{10}, x_{20}..x_{m0}, 0, 0,...0)^T$$

We then have

$$a_1x_{10} + a_2x_{20} +..+ a_mx_{m0} = b. \tag{11.2}$$

$$c_1x_{10} + c_2x_{20} +..+ c_mx_{m0} = z_0 \tag{11.3}$$

where a_j are the linearly independent column vectors of A, associated with the basic variables c_j is the component of c corresponding to the basic variable x_j and z_0 is the corresponding value of the objective function.

Since a_1, a_2.. a_m are linearly independent vectors, any vector from the set a_1, a_2,.. a_m,..a_n can be expressed in terms of a_1, a_2,..a_m.

Let $$\alpha_{ij}a_1 + \alpha_{2j}a_2 +..+ \alpha_{mj}a_m = a_j, j = 1, 2,..n \tag{11.4}$$

and define $$\alpha_{ij}c_1 + \alpha_{2j}c_2 +..+ \alpha_{mj}c_m = z_j, j = 1, 2,..n \tag{11.5}$$

Theorem 11.1. Given a basic feasible solution to a standard linear programming problem, if for some j associated with the nonbasic variables, $z_j - c_j > 0$ and all $\alpha_{ij} \leq 0$, i = 1, 2,..m, then a class of new feasible solutions can be constructed in which (m + 1) variables are strictly positive and the corresponding value of the objective function can be made arbitrarily small.

Proof: Multiplying (11.4) by some number θ and subtracting from (11.2) and similarly multiplying (11.5) by the same θ and subtracting from (11.3), we get,

$$(x_{10} - \theta\, \alpha_{ij})\, a_1 + (x_{20} - \theta\, \alpha_{2j})\, a_2 + ...+ (x_{m0} - \theta\, \alpha_{mj})\, a_m + \theta\, a_j = b, j = 1, 2,...n \tag{11.6}$$

$$(x_{10} - \theta\, \alpha_{ij})\, c_1 + (x_{20} - \theta\, \alpha_{2j})\, c_2 + ...+ (x_{m0} - \theta\, \alpha_{mj})\, c_m + \theta\, c_j = z_0 - \theta\,(z_j - c_j)\,,\ 1, 2,...n \tag{11.7}$$

where $\theta\, c_j$ has been added to both sides of (11.7).

Now, if at any stage, we have for some j associated with the nonbasic variables, $z_j - c_j > 0$ and all $\alpha_{ij} < 0$, then for any $\theta > 0$, the coefficients of a_j's in (11.6) will constitute a new feasible solution. It is clear that there is no upper bound to θ and that a class of new feasible solutions has been constructed in which (m + 1) variables are strictly positive and $z \to -\infty$ as $\theta \to +\infty$.

Theorem 11.2. Let every basic feasible solution to the linear programming problem (11.1) be nondegenerate. Then, if for some j associated with the nonbasic variables, $Z_j - c_j > 0$ and if $\alpha_{ij} > 0$ for at least one i, i = 1,2,..m then a new basic feasible solution with exactly m positive variables can be constructed whose value of the objective function is less then the value for the proceeding solution.

Proof: Since by assumption, the basic variables $x_{10}, x_{20},..x_{mo}$ are all positive, it is always possible to find a $\theta > 0$, for which the coefficients of the vectors in (11.6) remain positive. Thus, a new feasible solution is obtained and since $Z_j - c_j > 0$, we have from (11.7),

$$z = z_0 - \theta (z_j - c_j) < z_0 \tag{11.8}$$

Therefore, θ should be made as large as possible to have a maximum reduction in the value of z. However, if at least one $\alpha_{ij} > 0$, it is not possible to assign any positive value to θ, without violating the nonnegativity restrictions. The largest value of θ for which the coefficients of the vectors in (11.6) remain nonnegative is

$$\theta_0 = \operatorname*{Min}_i \frac{x_{i0}}{\alpha_{ij}} > 0 \qquad \text{for } \alpha_{ij} > 0. \tag{11.9}$$

Since the problem is nondegenerate, the minimum in (11.9) will be obtained for unique i, say for $i = r$ $(1 \leq r \leq m)$. If θ_0 is substituted for θ in (11.6) and (11.7), the coefficient corresponding to $i = r$ will vanish, which means that the vector a_r is replaced by a_j and we have a new basic feasible solution with exactly m positive variables,

$$x_i = x_{i0} - \theta_0 \alpha_{ij}, \quad i = 1, 2,..m, \quad i \neq r.$$

$$x_j = \theta_0$$

and other variables = 0

with $z = z_0 - \theta_0 (z_j - c_j) < z_0$

Note that if the problem is not nondenerate, $z \leq z_0$

Theorem 11.3. A necessary and sufficient condition for a basic feasible solution to the linear programming problem (11.1) to be minimal is that $z_j - c_j \leq 0$ for every j associated with a nonbasic variable.

Proof: It follows immediately from theorems 11.1 and 11.2 that the condition is necessary if we assume that every basic feasible solution to the problem (11.1) is nondegenerate.

Let us now show that the condition is also sufficient.

Let $Y = (y_1, y_2,..y_n)^T$ be any other feasible solution of the problem so that

$$y_1 a_1 + y_2 a_2 + ... + y_n a_n = b \tag{11.10}$$

$$y_j \geq 0, \quad j = 1,2..n$$

$$\text{and let } y_1 c_1 + y_2 c_2 + ... + y_n c_n = z^* \tag{11.11}$$

where z* is the corresponding value of the objective function.

By hypothesis $z_j - c_j \leq 0$, for every j associated with a nonbasic variable and since it is always true that $z_j - c_j = 0$ for every j, associated with a basic variable, we have,

$$z_j - c_j \leq 0 \quad \text{for all } j, j = 1,2..n. \tag{11.12}$$

Replacing c_j by z_j in (11.11), we then get

$$y_1 z_1 + y_2 z_2 + ... + y_n z_n \leq z^* \tag{11.13}$$

Further, on substituting (11.4) into (11.10) we obtain

$$y_1\left(\sum_{i=1}^{m}\alpha_{i1}a_i\right)+y_2\left(\sum_{i=1}^{m}\alpha_{i2}a_i\right)+\ldots+y_n\left(\sum_{i=1}^{m}\alpha_{in}a_i\right)=b$$

or $$\left(\sum_{j=1}^{n}y_j\alpha_{1j}\right)a_1+\left(\sum_{j=1}^{n}y_j\alpha_{2j}\right)a_2+\ldots+y_n\left(\sum_{j=1}^{n}y_j\alpha_{mj}\right)a_m=b \qquad (11.14)$$

Similarly, on substituting (11.5) into (11.13), we obtain

$$\left(\sum_{j=1}^{n}y_j\alpha_{1j}\right)c_1+\left(\sum_{j=1}^{n}y_j\alpha_{2j}\right)c_2+\ldots+\left(\sum_{j=1}^{n}y_j\alpha_{mj}\right)c_m\le Z^* \qquad (11.15)$$

Now, $a_1,a_2,..a_m$ are basis vectors and since the expression of any vector in terms of basis vectors in unique, the coefficients of the corresponding vectors in (11.2) and (11.14) must be equal, that is

$$x_{i0}=\sum_{j=1}^{n}y_j\alpha_{ij},\qquad i=1,2..m \qquad (11.16)$$

Hence from (11.15), we have

$$x_{10}c_1 + x_{20}c_2 +\ldots+ x_{m0}c_m \le z^*$$

or $$z_0 \le z^*.$$

Note that the nondegeneracy assumption is not required for the proof of sufficiency.

Corollary 11.1: Given an optimal basic feasible solution to the linear programming problem (11.1), a necessary and sufficient condition that there is another basic feasible solution to be optimal is that for some j associated with the nonbasic variables, $z_j - c_j = 0$ for which $\alpha_{ij} > 0$ for at least on i.

Proof: The proof follows from theorem 11.3 and the relation

$$z^* = z_0 - \theta_0\,(z_j - c_j).$$

Corollary 11.2: A necessary and sufficient condition for an optimal solution to the linear programming problem (11.1) to be unique is that $z_j - c_j < 0$, for every j, associated with the nonbasic variables.

Proof: The proof follows immediately from corollary 11.1 and from the fact that for a given basis, the solution is unique.

Theorem 11.4. Under the assumption of nondegeneracy at each iteration, the simplex procedure terminates in a finite number of steps.

Proof: The simplex procedure moves from one basis to another and under the assumption of nondegeneracy, the value of the objective function is decreased at each step. In the absence of degeneracy therefore, no basis can ever be repeated and since there is only a finite number of bases, the process terminates in a finite number of steps either with an optimal basic solution or an indication that there is an unbounded solution of the problem.

In the presence of degeneracy, one cannot be sure that the simplex procedure will necessarily terminate in a finite number of steps. If a basic feasible solution is degenerate, that is one or more of the basic variables are zero, then θ_0 in (11.9) is zero and the new solution obtained is again degenerate with no improvement in the value of the objective function. This may occur in several successive iterations and it is possible to return to a basis already obtained and then the procedure may cycle indefinitely. It is therefore desirable to develop a procedure to avoid degeneracy in the problem.

The techniques to resolve degeneracy in linear programming has been discussed in Chapter 13.

11.3. Method of Computation: The Simplex Algorithm

The theorems discussed in the previous section indicate how to proceed step by step so that the procedure converges to an optimal solution to the linear programming problem or determines that the problem has no optimal solution.

We summarize them as follows:

The Simplex Algorithm

1. Convert all the inequality constraints into equations by introducing slack or surplus variables in the constraints.
2. Check if all b_i $(i = 1,2..m)$ are nonnegative. If any one of b_i is negative, multiply the corresponding equation by -1.
3. Express the problem in tabular form known as Simplex Tableau. (See section 11.4)
4. Obtain an initial basic feasible solution X_B to the problem where B is the initial basis matrix (in the case where there is no obvious feasible solution, see Chapter 12)
5. Calculate the matrix (α_{ij}) and the quantities $z_j - c_j$ according to (11.4) and (11.5). Test the $z_j - c_j$,
   (a) If for every j, $z_j - c_j \leq 0$ (≥ 0), the present solution is minimal (maximal).
   (b) If for some j, $z_j - c_j > 0$ (<0), proceed as follows.
      (i) If all $\alpha_{ij} \leq 0$, the problem is unbounded.
      (ii) If $\alpha_{ij} > 0$, for at least one i, associated with the basic variables, the solution can be improved and proceed to step 6.
6. If $z_j - c_j > 0$ (< 0) holds for more than one j, we introduce the vector a_s in the basis for which

$$|z_s - c_s| = \max_j |z_j - c_j|. \quad \text{entry criterion}$$

and depart the vector a_r from the basis for which

$$\frac{x_r}{\alpha_{rs}} = \underset{i}{\text{Min}}\left[\frac{x_i}{\alpha_{is}},\quad \alpha_{is} > 0\right] \qquad \text{exit criterion}$$

to obtain a new set of basis vectors $a_1, a_2, \ldots a_{r-1}\, a_s, a_{r+1}, \ldots a_m$.
α_{rs} is called the pivot element

7. Calculate the new values of the basic variables, X_B, and the new values α'_{ij} and $z_j' - c_j$ and enter them into the next tableau [See section 11.5].
8. Repeat the process from step 5 until either an optimal solution is obtained or there is an indication that the problem is unbounded.

11.4. The Simplex Tableau

To apply simplex algorithm in a systematic and efficient manner, the linear programming problem is expressed in a tabular form known as Simplex Tableau. The simplex tableau is a very useful tabular form displaying all the quantities of interest in a linear programming problem so that the computation can be carried out in a systematic way.

Let $\beta_1, \beta_2 .. \beta_m$ be a set of linearly independent vectors and let $a_1, a_2, \ldots a_n$ be a set of vectors each of which is a linear combination of β_i. The *tableau* of the vectors a_j, with respect to the basis β_i is the matrix T_0, expressing each of the a_j as a linear combination of the β_i. It is expressed as

$$\begin{array}{c|cccccc|l}
 & a_1 & a_2 & .. & a_j & .. & a_m & \\
\hline
\beta_1 & \alpha_{11} & \alpha_{12} & & \alpha_{1j} & & \alpha_{1n} & \\
\beta_2 & \alpha_{21} & \alpha_{22} & & \alpha_{2j} & & \alpha_{2n} & = T_0. \\
\beta_i & \alpha_{i1} & \alpha_{i2} & & \alpha_{ij} & & \alpha_{in} & \\
\beta_m & \alpha_{m1} & \alpha_{m2} & & \alpha_{mj} & & \alpha_{mn} & \\
\hline
\end{array}$$

The elements α_{ij} are the coefficients of β_i, in the expression for a_j, that is,

$$a_j = \sum_{i=1}^{m} \alpha_{ij}\beta_i$$

In simplex tableau, a few more rows and columns are added to give all the relevant information about the problem. The first column of the tableau gives the costs c_B, corresponding to the vectors in the basis. The second column shows the vectors that are in the basis. Thus a_{B_i} is the vector from A which is in the ith column of the basis matrix B. The succeeding n–columns are the coefficients of a_{B_i} in the expression of a_j in terms of a_{B_i}. The next column under the heading 'b' or X_B gives the current values of the basic variables and the last column of the tableau gives the values of the ratio b_i/α_{ij} where a_j is the vector to be introduced in the basis.

The (m + 1)th row of the tableau gives the cost elements of the problem and the last row provides the values of $z_j - c_j$.

The simplex tableau then looks like

c_B	a_B	a_1	a_2	...	a_j	...	a_n	$X_B = b$	b/α_{ij}
c_{B_1}	a_{B_1}	α_{11}	α_{12}		α_{1j}		α_{1n}	b_1	b_1/α_{1j}
c_{B_2}	a_{B_2}	α_{21}	α_{22}		α_{2j}		α_{2n}	b_2	b_2/α_{2j}
c_{B_i}	a_{B_i}	α_{i1}	α_{i2}		α_{ij}		α_{in}	b_i	b_i/α_{ij}
c_{B_m}	a_{B_m}	α_{m1}	α_{m2}		α_{mj}		α_{mn}	b_m	b_m/α_{mj}
	c	c_1	c_2		c_j		c_n		
	$z_j - c_j$	z_1-c_1	z_2-c_2		z_j-c_j		z_n-c_n		(11.17)

Table 11.1 Simplex Tableau.

11.5. Replacement Operation

Suppose that a basic feasible solution to the linear programming problem (11.1) is given by $X_B = B^{-1}b$, where B is a basis matrix whose columns are the first m columns of A. For any column vector a_j from A, we then have.

$$a_j = \alpha_{1j}\, a_1 + \alpha_{2j}\, a_2 + ... + \alpha_{rj}\, a_r + ... + \alpha_{mj}\, a_m, j = 1, 2,...n \tag{11.18}$$

where α_{ij} are scalars.

Suppose that the present basic feasible solution is not optimal and we want to examine the possibility of finding another basic feasible solution with an improved value of z. In Chapter 3, we have discussed how one of the vectors in a basis can be replaced by another vector to form a new basis. In the present case let a_s be a vector associated with the nonbasic variable x_s so that

$$a_s = \alpha_{1s}\, a_1 + \alpha_{2s}\, a_2 + ... + \alpha_{rs}\, a_r + ... + \alpha_{ms}\, a_m \tag{11.19}$$

and let $\alpha_{rs} \neq 0$. We then get a new basis consisting of vectors a_1, a_2.. a_{r-1}, $a_s a_{r+1}..a_m$, by replacing the vector a_r by a_s and entries in the new tableau, i.e. the new values of α_{ij}, X_B, and $z_j - c_j$ can be easily calculated.

From (11.19), we get

$$a_r = \frac{a_s}{\alpha_{rs}} - \sum_{\substack{i=1 \\ i \neq r}}^{m} \frac{\alpha_{is}}{\alpha_{rs}} a_i \tag{11.20}$$

Now, any vector a_j can be expressed as

$$a_j = \sum_{\substack{i=1 \\ i \neq r}}^{m} \alpha_{ij} a_i + a_{rj} a_i$$

Substituting the value of a_r from (11.20), we have,

$$a_j = \sum_{\substack{i=1 \\ i \neq r}}^{m} \alpha_{ij} a_i + \left[\frac{a_s}{\alpha_{rs}} \sum_{\substack{i=1 \\ i \neq r}}^{m} \frac{\alpha_{is}}{\alpha_{rs}} a_i \right] \alpha_{rj}$$

$$= \sum_{\substack{i=1 \\ i \neq r}}^{m} \left(\alpha_{ij} - \frac{\alpha_{is}\alpha_{rj}}{\alpha_{rs}} \right) a_i + \frac{\alpha_{rj}}{\alpha_{rs}} a_s \tag{11.21}$$

Thus the new values α'_{ij} are given by,

$$\alpha'_{ij} = \alpha_{ij} - \frac{\alpha_{is}\alpha_{rj}}{\alpha_{rs}}, \quad i = 1, 2, ..m; \ i \neq r. \tag{11.22}$$

$$\alpha'_{rj} = \frac{\alpha_{rj}}{\alpha_{rs}} \tag{11.23}$$

From the original basic feasible solution X_B, we have

$$\sum_{i=1}^{m} x_{B_i} a_i = b.$$

or
$$\sum_{\substack{i=1 \\ i \neq r}}^{m} x_{B_i} a_i + x_{B_r} a_r = b \tag{11.24}$$

Now, eliminating a_r in (11.24) by (11.20), we have

$$\sum_{\substack{i=1 \\ i \neq r}}^{m} \left(x_{B_i} - \frac{\alpha_{is}}{\alpha_{rs}} x_{B_r} \right) a_i + \frac{x_{B_r}}{\alpha_{rs}} a_s = b \tag{11.25}$$

and if the new basic solution is to be feasible, we must have

$$x'_i = x_{B_i} - \frac{\alpha_{is}}{\alpha_{rs}} x_{B_r} \geq 0, \quad i = 1, 2..m; \quad i \neq r \tag{11.26}$$

$$x'_s = \frac{x_{B_r}}{\alpha_{rs}} \geq 0 \tag{11.27}$$

It is clear that if we are interested in having the new basic solution nonnegative, we cannot replace the vector a_r by a_s when α_{rs} is not positive. In fact, we immediately note that to satisfy (11.27), we must have $\alpha_{rs} > 0$, since $x_{Br} > 0$. Now, if $\alpha_{rs} > 0$ and for some i, $\alpha_{is} \leq 0$ ($i \neq r$) then (11.26) is automatically satisfied for that i. We therefore need to be concerned with the case when $\alpha_{is} > 0$. We then select from a_i with $\alpha_{is} > 0$, the vector a_r to be removed from the old basis so that (11.26) is satisfied.

If $\alpha_{is} > 0$, (11.26) can be written as

$$\frac{x_{Bi}}{\alpha_{is}} - \frac{x_{Br}}{\alpha_{rs}} \geq 0 \tag{11.28}$$

We therefore, determine

$$\operatorname*{Min}_{i} \left\{ \frac{x_{B_i}}{\alpha_{is}} \middle| \alpha_{is} > 0 \right\} \tag{11.29}$$

and if this minimum (unique under nondegenerate assumption) is obtained when $i = r$, then a_r is replaced by a_s. This implies that (11.26) are satisfied, i.e. the new basic solution is feasible.

We are now interested in getting the transformation formula of $z_j - c_j$ and of z and show that there is an improvement in the value of the objective function at the new basic feasible solution. Let x_j be a nonbasic variable, then $z'_j - c_j$ is given by

$$z'_j - c_j = \sum_{\substack{i=1\\ i\neq r}}^{m} \alpha'_{ij} c_i + \alpha'_{rj} c_s - c_j.$$

$$= \sum_{\substack{i=1\\ i\neq r}}^{m} \left(\alpha_{ij} - \frac{\alpha_{is}\alpha_{rj}}{\alpha_{rs}} \right) c_i + \frac{\alpha_{rj}}{\alpha_{rs}} c_s - c_j.$$

or $$z'_j - c_j = \sum_{i=1}^{m} \alpha_{ij} c_i - \alpha_{rj} c_r - \sum_{i=1}^{m} \frac{\alpha_{is}\alpha_{rj}}{\alpha_{rs}} c_i + \frac{\alpha_{rs}\alpha_{rj}}{\alpha_{rs}} c_r + \frac{\alpha_{rj}}{\alpha_{rs}} c_s - c_j$$

$$= \left[\sum_{i=1}^{m} \alpha_{ij} c_i - c_j \right] - \frac{\alpha_{rj}}{\alpha_{rs}} \left[\sum_{i=1}^{m} \alpha_{is} c_i - c_s \right]$$

$$= \left(z_j - c_j\right) - \frac{\alpha_{rj}}{\alpha_{rs}} \left(z_s - c_s\right) \tag{11.30}$$

Now the new value of the objective function is

$$z' = \sum_{i\neq r}^{m} c_i x'_i + c_s x'_s$$

Substituting the values of x_i', x'_s from (11.26) and (11.27) in the above expression we get

$$z' = z - \frac{x_{Br}}{\alpha_{rs}} (z_s - c_s)$$

and hence $z' \leq z$ if $z_s - c_s > 0$.

We therefore enter that vector a_s into the basis for which $z_s - c_s > 0$ and at least one ais (say α_{rs}) > 0, so that the value of z in the new solution is improved.

11.5.1 Replacement Rule

Suppose at some stage of computation for a solution of a linear programming problem by the simplex method, we find that to improve the solution we are to replace the vector a_r in the basis by the vector a_s. We are to calculate the new values of α_{ij}, X_b and $Z_j - c_j$ and enter them in the next tableau.

Let the column associated with the vector a_s be called the *distinguished column* and the row associated with the vector a_r be known as the *distinguished row*. The element α_{rs} at their intersection (distinguished position) is called the *distinguished element* or *the pivot element* and is indicated by an asterisk.

From the formulae derived in the previous section, it is clear that by following the rules stated below, the values of x_B', α'_{ij} and $z_j' - c_j$ can be easily obtained.

Rule 1: Divide all elements of the distinguished row by the distinguished element, thus obtaining an unity in the distinguished position.

Rule 2: Subtract from the elements of the other rows, the ratio of the product of the corresponding elements in the distinguished row and the distinguished column to the distinguished element.

Note that these operations will give zeros in th distinguished column except in the distinguished position, where we will get unity.

11.6. Example

Consider the problem

Minimize $z = x_1 - 3x_2 + 2x_3.$

Subject to

$$3x_1 - x_2 + 2x_3 \le 7$$
$$-4x_1 + 3x_2 + 8x_3 \le 10$$
$$-x_1 + 2x_2 \le 6$$
$$x_1, x_2, x_3 \ge 0.$$

After introducing the slack variables, the problem reduces to the standard form

Minimize $z = x_1 - 3x_2 + 2x_3$

Subject to

$$3x_1 - x_2 + 2x_3 + x_4 = 7$$
$$-4x_1 + 3x_2 + 8x_3 + x_5 = 10$$
$$-x_1 + 2x_2 + x_6 = 6$$
$$x_j \ge 0, j = 1,2,..6.$$

Tableau 1

c_B	a_B	a_1	a_2	a_3	a_4	a_5	a_6	b	b_i/a_{i2}
0	a_4	3	–1	2	1	0	0	7	–
0	a_5	–4	3	8	0	1	0	10	10/3
0	a_6	–1	2*	0	0	0	1	6	3 →
	c	1	–3	2	0	0	0		
	zj – cj	–1	3	–2	0	0	0		
			↑						

Tableau 2

c_B	a_B	a_1	a_2	a_3	a_4	a_5	a_6	b	b_i/a_{i1}
0	a_4	5/2*	0	2	1	0	1/2	10	4 →
0	a_5	–5/2	0	8	0	1	–3/2	1	–
–3	a_2	–1/2	1	0	0	0	1/2	3	–
	c	1	–3	2	0	0	0		
	zj – cj	1/2	0	–2	0	0	–3/2		
		↑							

Tableau 3

c_B	a_B	a_1	a_2	a_3	a_4	a_5	a_6	b
1	a_1	1	0	4/5	2/5	0	1/5	4
0	a_5	0	0	10	1	1	–1	11
–3	a_2	0	1	2/5	1/5	0	3/5	5
	c	1	–3	2	0	0	0	
	zj – cj	0	0	–12/5	–1/5	0	–8/5	

Since in tableau 3, $z_j - c_j \leq 0$ for all j, the solution obtained is optimal.

Thus the optimal solution is

$$x_1 = 4, x_2 = 5, x_3 = 0 \text{ and Min } z = -11.$$

11.7. Exercises

1. Consider the linear programming problem max c^TX, $AX = b$, $X \geq 0$. If X_0 is an optimal solution, will X_0 be an optimal solution if
   (i) c^TX is changed to $c^TX + k$, where k is a constant.
   (ii) c^TX is changed to kc^TX, where $k > 0$.
   (iii) c is changed to c + d, where $d \neq 0$ is an n component vector.
2. Under what condition a standard linear programming problem has
   (i) a unique optimal solution.
   (ii) more than one optimal solutions.
   (iii) an unbounded solution.
3. Solve the following linear programming problems by the simplex method.
   (a) Maximize $2x_1 + 4x_2 + 3x_3$
   Subject to $2x_1 + x_2 + 2x_3 \leq 4$

$$3x_1 + 4x_2 + 2x_3 \leq 6$$
$$x_1 + 3x_2 + 2x_3 \leq 8$$
$$x_1, x, x_3 \geq 0$$

(b) Maximize $z = 7x_1 + 5x_2$

Subject to
$$7x_1 + 10x_2 \leq 35$$
$$8x_1 + 3x_2 \leq 24$$
$$x_1 + x_2 \leq 4$$
$$x_1, x_2 \geq 0$$

(c) Maximize $2x_1 + 3x_2$

Subject to
$$x_1 + x_2 \leq 6$$
$$x_1 + 2x_2 \leq 10$$
$$x_1 \leq 4$$
$$x_1, x_2 \geq 0$$

(d) Maximize $3x_1 + 4x_2$

Subject to
$$2x_1 + 3x_2 \leq 14$$
$$2x_1 + x_2 \leq 8$$
$$x_1 + x_2 \leq 5$$
$$x_1, x_2 \geq 0$$

(e) Maximize $x_1 - 3x_2 + 2x_3$

Subject to
$$x_1 - x_2 + x_3 \leq 3$$
$$x_1 + 2x_2 + 4x_3 \leq 4$$
$$x_1 + 6x_2 + 3x_3 \leq 6$$
$$x_1, x_2 \geq 0.$$

4. Solving by the simplex method, show that

(a) the problem

Maximize $x_1 + 2x_2$

Subject to
$$x_1 - 2x_2 \leq 3$$
$$x_1 + 2x_2 \leq 12$$
$$-x_1 + 2x_2 \leq 8$$
$$x_1, x_2 \geq 0$$

has multiple optimal solution

(b) the problem

Maximize $2x_1 + 3x_2$

Subject to
$$x_1 + x_2 \leq 3$$
$$x_1 - 2x_2 \leq 1$$
$$x_2 \leq 2$$
$$x_1, x_2 \leq 0$$

has a unique solution

(c) the problem

Maximize $2x_1 + 3x_2$

Subject to $$\begin{aligned} &x_1 + x_2 \geq 4 \\ &x_1 - x_2 \leq 2 \\ &x_1, x_2 \geq 0 \end{aligned}$$

has no optimal solution.

5. Given the feasible solution $x_1 = 1$, $x_2 = 1$, $x_3 = 2$, $x_4 = 3$ to the following set of equations find a basic feasible solution of

$$\begin{aligned} x_1 - 4x_2 + 6x_3 + 2x_4 &= 15 \\ 5x_1 + 3x_3 - 4x_3 + x_4 &= 3 \\ 2x_1 + 5x_2 + x_3 - 3x_4 &= 0 \\ x_1, x_2, x_3, x_4 &\geq 0 \end{aligned}$$

CHAPTER 12

Simplex Method: Initial Basic Feasible Solution

12.1. Introduction: Artificial Variable Techniques

Application of the simplex algorithm requires knowledge of an initial basic feasible solution to the linear programming problem. There are many problems for which such a feasible solution is readily available[1] but there are other problems encountered in practice which do not provide any knowledge of the system of constraints. The problem may have inconsistency due to errors in obtaining or recording the numerical values and moreover it may also have redundancies. In fact before we proceed to obtain a basic feasible solution, it is necessary to find out if the problem has inconsistency or redundancy which however, might be quite time consuming.

It is therefore important that a general mathematical technique must be developed to solve linear programming problems without any prior knowledge of the system of constraints.

Two methods are currently used to achieve this end.

(a) The two-phase method, developed by Dantzig, Orden and Wolfe.

(b) The method of penalties, due to A Charnes.

12.2. The Two-Phase Method [117]

The linear programming problem is first expressed in the standard form and the simplex method is used in two phases. In phase I the simplex algorithm is applied to an auxiliary linear programming problem and determines whether the original problem is feasible and if feasible finds a basic feasible solution. The basic feasible solution thus obtained then acts as an initial basic feasible solution to the original problem and the simplex algorithm is started in Phase II, which finally finds a solution to the original problem.

1. For example, if slack variables are added to every constraint to convert less than or equal to type inequalities into equations, we immediately see that these slack variables constitute the initial basic variables. If the constraints are equations or greater than or equal to type inequalities, it may not be possible to get a basic feasible solution directly.

It has several important features:

(i) No assumption on the original system of constraints are made; the system may be redundant, inconsistent or not solvable in nonnegative numbers.
(ii) An initial basic feasible solution for Phase I can be easily obtained.
(iii) The end product of Phase I is a basic feasible solution (if it exists) to the original problem so that the simplex algorithm can be initiated in Phase II.

Outline of the Procedure

Step 1: Express the linear programming problem in standard form by introducing slack or surplus variables wherever is necessary. Multiply, if necessary certain constraints of the system by –1, so that the constants on the right hand side are all nonnegative.

The problem then becomes

$$\text{Minimize} \quad z = c^TX \tag{12.1}$$

$$\text{Subject to} \quad AX = b \tag{12.2}$$

$$X \geq 0 \tag{12.3}$$

where $b^T = (b_1, b_2, \dots b_m) \geq 0$, $c^T = (c_1, c_3 \dots c_n)$
$X^T = (x_1, x_2 \dots x_n)$ and $A = (a_{ij})$ is an m x n matrix.

Step 2 (Phase I): The system of equations (12.2) is augmented by introducing additional variables $W_i \geq 0$, called the artificial variables so that the constraints become

$$AX + IW = b \tag{12.4}$$

where $W^T = (w_1, w_2, \dots w_m)$

The simplex algorithm is then applied to find a solution to the auxiliary linear programming problem

$$\text{Minimize} \quad z^* = e^TW \tag{12.5}$$

$$\text{Subject to} \quad AX + IW = b$$

$$X, W \geq 0 \tag{12.6}$$

where $e^T = (1, 1, \dots 1)$ is an m-component vector having unity as a value for each component.

In this case ($X = 0$, $W = b$) can be taken as initial basic feasible solution and the simplex algorithm can readily be started. Since $W_i \geq 0$, the problem cannot have an unbounded solution and the process terminates as soon as an optimal solution (X_0, W_0) to the problem is found.

Three cases may arise

(a) Min $z^* > 0$. In this case no feasible solution exists for the original problem because if $X \geq 0$ is a feasible solution, then (X, $W = 0$) is feasible to the auxiliary problem with all $W_i = 0$, which contradicts that Min $z^* > 0$.

(b) Min z* = 0 and all the artificial variables are nonbasic. In such a case, there exists an optimal solution (X_0, W_0) with $W_0 = 0$ and X_0 is a basic feasible solution to the original problem.

(c) Min z* = 0 and at least one artificial variable is basic at zero level. Thus a basic feasible solution of the problem is obtained with these artificial variables as a part of the basic set of variables. This will occur whenever the original system has redundancies and often when it has degenerate solutions.

Step 3 (Phase II): In cases (b) and (c) all the columns corresponding to the nonbasic artificial variables are deleted from the final table in Phase I and z* is replaced by z, the objective function of the original problem. The tableau so obtained forms the starting tableau of Phase II. The simplex algorithm is applied to this table taking care that the artificial variables (if any at zero level) never become positive. (Since this would destroy feasibility). To ensure this the usual simplex computation is slightly modified (See 12.6). The algorithm terminates as soon as either an optimal solution is obtained or there is an indication of the existence of an unbounded solution.

12.3. Examples

12.3.1. Minimize $z = x_1 + x_2$

Subject to
$$7x_1 + x_2 \geq 7$$
$$x_1 + 2x_2 \geq 4$$
$$x_1, x_2 \geq 0$$

Introducing the surplus variables x_3, x_4 to the constraints, we reduce the problem to the standard form:

Minimize $z = x_1 + x_2$

Subject to
$$7x_1 + x_2 - x_3 = 7$$
$$x_1 + 2x_2 - x_4 = 4$$
$$x_j \geq 0, j = 1,2,3,4.$$

We then add artificial variables $w_1 \geq 0$, $w_2 \geq 0$ to the equality constraints above and replace the objective function by $w_1 + w_2$ and the problem thus obtained is solved by the simplex algorithm in Phase I.

Phase I

Minimize $z^* = w_1 + w_2$

Subject to
$$7x_1 + x_2 - x_3 + w_1 = 7$$
$$x_1 + 2x_2 - x_4 + w_2 = 4$$
$$x_j \geq 0, j = 1,2,3,4, w_1, w_2 \geq 0.$$

Tableau 1

c_B	a_B	a_1	a_2	a_3	a_4	a_5	a_6	b	b/a_{i1}
1	a_5	7*	1	−1	0	1	0	7	1 →
1	a_6	1	2	0	−1	0	1	4	4
	c	0	0	0	0	1	1		
	$z^*_j - c_j$	8	3	−1	−1	0	0		
		↑							

Tableau 2

c_B	a_B	a_1	a_2	a_3	a_4	a_5	a_6	b	b/a_{i2}
0	a_1	1	1/7	−1/7	0	1/7	0	1	7
1	a_6	0	13/7*	1/7	−1	−1/7	1	3	21/13 →
	$z^*_j - c_j$	0	13/7	1/7	−1	−8/7	0		

Tableau 3

c_B	a_B	a_1	a_2	a_3	a_4	a_5	a_6	b
0	a_1	1	0	−2/13	1/13	2/13	−1/13	10/13
0	a_2	0	1	1/13	−7/13	−1/13	7/13	21/13
	$z^*_j - c_j$	0	0	0	0	−1	−1	

Since all $z_j^* - c_j \leq 0$, in tableau 3, an optimal basic solution to the auxiliary problem has been attained. Furthermore, no artificial variable appears in the basis. Hence,

$$x_1 = 10/13, x_2 = 21/13$$

forms a basic feasible solution to the original problem.

Now, deleting the columns corresponding to nonbasic artificial variables in tableau 3 and replacing the objective function by the original objective function, the problem is solved by the simplex algorithm.

Phase II

Tableau 4

c_B	a_B	a_1	a_2	a_3	a_4	b
1	a_1	1	0	–2/13	1/13	10/13
1	a_2	0	1	1/13	–7/13	21/13
	c	1	1	0	0	
	$z^*_j - c_j$	0	0	–1/13	–6/13	

Since all $z_j - c_j \leq 0$, an optimal basic solution to the original linear program has been obtained, which is

$$x_1 = 10/13,\ x_2 = 21/13 \text{ and Min } z = 31/13.$$

12.3.2 Maximize $z = 3x_1 + 5x_2$

$$x_1 + 2x_2 \leq 1$$
$$4x_1 + x_2 \geq 6$$
$$x_1, x_2 \geq 0$$

The linear program is first reduced to the auxiliary problem

Minimize $z^* = w$

Subject to

$$x_1 + 2x_2 + x_3 = 1$$
$$4x_1 + x_2 - x_4 + w = 6$$
$$x_j \geq 0,\ j = 1,2,3,4,\ w \geq 0$$

where x_3, x_4 are the slack and surplus variables and w is an artificial variable. Note that since x_3 can be taken as a basic variable, the artificial variable w is added only to the second constraint.

Phase I

Tableau 1

c_B	a_B	a_1	a_2	a_3	a_4	a_5	b	b/a_{i1}
0	a_3	1*	2	1	0	0	1	1 →
1	a_5	4	1	0	–1	1	6	3/2
	c	0	0	0	0	1		
	$z^*_j - c_j$	4 ↑	1	0	–1	0		

Tableau 2

c_B	a_B	a_1	a_2	a_3	a_4	a_5	b
0	a_1	1	2	1	0	0	1
1	a_5	0	−7	−4	−1	1	2
	c	0	0	0	0	1	
	$z^*_j - c_j$	0	−7	−4	−1	0	

Since all $z^*_j - c_j \leq 0$, an optimal basic solution to the auxiliary problem is obtained but the artificial variable w is in the basis at a positive level. This shows that the original linear programming problem does not have a feasible solution.

12.4. The Method of Penalties [71]

Another intuitive method to find a basic feasible solution to a linear programming problem is the method of penalties. The method is commonly known as Charnes' M-technique or Big M-method.

As in the two-phase method, here also artificial variables are added to the constraints which enable us to find an initial basic feasible solution to the augmented constraints. It is clear that if a new basic feasible solution to the augmented constraints is found where all the artificial variables have the value zero, we have a basic feasible solution to the original problem.

To achieve this, each artificial variable is multiplied by a penalty M per unit, an arbitrarily large positive number (negative in the case of a maximization problem) and their sum is added to the objective function of the original problem. The slack variables (if there is any) and the artificial variables form an initial basic feasible solution and the simplex algorithm can be readily applied to optimize the augmented problem. If there are no more artificial variables in the basis, we have a basic feasible solution to the original problem.

The simplex algorithm is then continued until an optimal solution to the original problem is found.

Outline of the Procedure

Consider the general linear programming problem taking care that the constants on the right hand side are all nonnegative, multiplying, if necessary, by −1 through out the constraint. We thus have the problem,

$$\text{Minimize} \quad z = \sum_{j=1}^{n} c_j x_j$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij}x_j \le b_i, \quad i = 1,2,..m_1$$

$$\sum_{j=1}^{n} a_{ij}x_j \ge b_i, \quad i = m_1+1,..m_1+m_2.$$

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = m_1+m_2+1...m$$

$$x_j \ge 0, \quad j = 1,2..n$$

where $b_i \ge 0$, for all $i = 1,2...m$. (12.7)

Step 1: Express the linear programming problem in the standard form by introducing slack or surplus variables wherever is necessary. The problem then reduces to,

$$\text{Minimize} \quad z = \sum_{j=1}^{n} c_j x_j$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij}x_j + x_{n+i} = b_i, \quad i = 1,2...m_1,$$

$$\sum_{j=1}^{n} a_{ij}x_j + x_{n+i} = b_i, \quad i = m_1,+1,...m_1+m_2,$$

$$\sum_{j=1}^{n} a_{ij}x_j \quad = b_i, \quad i = m_1+m_2+1,...m$$

$$x_j \quad \ge 0, \quad j = 1,2...n \tag{12.8}$$

where $b_i \ge 0$, for all $i = 1,2...m$,

$x_{n+i} \ge 0$, $i = 1,2..m$, are slack variables

and $x_{n+i} \ge 0$, $i = m_1 + 1,...m_1 + m_2$ are surplus variables.

Step 2: Introduce nonnegative artificial variables to the left hand side of the constraints except where we have slack variables, since a slack variable itself can be taken as a basic variable. Assign an arbitrarily large positive number M per unit to each of the artificial variables and their sum is added to the original objective function.

The problem then becomes,

$$\text{Minimize} \quad z = \sum_{j=1}^{n} c_j x_j + M \sum_{i=m_1+1}^{m} w_i$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij}x_j + x_{n+i} = b_i, \quad i = 1,2,..m_1,$$

$$\sum_{j=1}^{n} a_{ij}x_j - x_{n+i} + w_i = b_i, \qquad i = m_1 + 1 \ldots m_1 + m_2$$

$$\sum_{j=1}^{n} a_{ij}x_j \qquad + w_i = b_i, \qquad i = m_1 + m_2 + 1, \ldots, m,$$

$$\begin{aligned} x_j &\geq 0, & j &= 1, \ldots n, \\ x_{n+i} &\geq 0, & i &= 1,2 \ldots m_1 + m_2 \\ w_i &\geq 0, & i &= m_1 + 1, \ldots m \end{aligned} \qquad (12.9)$$

where $b_1 \geq 0$, $i = 1,2 \ldots m$

and M is an unspecified large positive number.

Step 3: The simplex algorithm is applied to the augmented problem (12.9). The slack variables along with the artificial variables form an initial basic feasible solution to start the computation.

The following three cases may arise:

Case (i): There are no more artificial variables in the basis. We then have a basic feasible solution to the original problem and we proceed to step 4.

Case (ii): At least one artificial variable is in the basis at a zero level and the coefficient of M in each $z_j - c_j$ is negative or zero. In such a case, the current basic feasible solution is a degenerate one to the original problem. If no $z_j - c_j$ is strictly positive, we have an optimal solution to our problem otherwise we proceed to step 4.

Case (iii): At least one artificial variable is in the basis at a positive level and the coefficient of M in each $z_j - c_j$ is negative or zero. In this case, the original problem has no feasible solution, for if there were a feasible solution to the problem, the corresponding artificial vector could be removed from the basis with an improvement in the value of z.

Step 4: Application of the simplex algorithm is continued until either an optimal basic feasible solution is obtained or there is an indication of the existence of an unbounded solution to the original problem.

Note: Whenever an artificial vector leaves the basis, we drop that vector and omit all the entries corresponding to its column from the simplex tableau as we are not interested in its re-entry into the basis.

12.5. Examples: Penalty Method

12.5.1 Minimize $z = 3x_1 + 4x_2$

Subject to

$$\begin{aligned} 2x_1 - 3x_2 &\leq 6 \\ x_1 + 2x_2 &\geq 10 \\ x_1 + x_2 &\geq 6 \\ x_1, x_2 &\geq 0. \end{aligned}$$

The augmented problem is given by

$$\text{Minimize} \quad z^* = 3x_1 + 4x_2 + Mx_6 + Mx_7.$$

$$\text{Subject to} \quad 2x_1 - 3x_2 + x_3 = 6$$

$$x_1 + 2x_2 - x_4 + x_6 = 10$$

$$x_1 + x_2 - x_5 + x_7 = 6$$

$$x_j \geq 0,\ j = 1,2,\ldots 7.$$

where x_3 is a slack variable, x_4, x_5 are surplus variables and x_6, x_7 are artificial variables.

Tableau 1

c_B	a_B	a_1	a_2	a_3	a_4	a_5	a_6	a_7	b	b/a_{i2}
0	a_3	2	–3	1	0	0	0	0	6	–
M	a_6	1	2*	0	–1	0	1	0	10	5 →
M	a_7	1	1	0	0	–1	0	1	6	6
	c	3	4	0	0	0	M	M		
	$z^*_j - c_j$	2M–3	3M–4	0	–M	–M	0	0		
			↑							

Tableau 2

c_B	a_B	a_1	a_2	a_3	a_4	a_5	a_7	b	b/a_{i1}
0	a_3	7/2	0	1	–3/2	0	0	21	6
4	a_2	1/2	1	0	–1/2	0	0	5	10
M	a_7	1/2*	0	0	1/2	–1	1	1	2 →
	c	3	4	0	0	0	M		
	$z^*_j - c_j$	M/2–1	0	0	M/2–2	–M	0		
		↑							

Tableau 3

c_B	a_B	a_1	a_2	a_3	a_4	a_5	b
0	a_3	0	0	1	–5	7	14
4	a_2	0	1	0	–1	1	4
3	a_1	1	0	0	1	–2	2
	c	3	4	0	0	0	
	$z^*_j - c_j$	0	0	0	–1	–2	

Since all $z^*_j - c_j \leq 0$ and there is no artificial variable in the basis, an optimal solution to the original problem is obtained which is given by $x_1 = 2$, $x_2 = 4$ and minimum $z = 22$.

12.5.2 Maximize $z = 2x_1 + 3x_2$

Subject to

$$x_1 + 2x_2 \leq 1$$
$$4x_1 + 3x_3 \geq 6$$
$$x_1, x_2 \geq 0.$$

The augmented problem is given by

Minimize $z^* = -2x_1 - 3x_2 + Mx_5$

Subject to

$$x_1 + 2x_2 + x_3 = 1$$
$$4x_1 + 3x_2 - x_4 + x_5 = 6$$
$$x_j \geq 0, j = 1, 2, \ldots 5.$$

where x_3, x_4 are slack and surplus variables respectively and x_5 is an artificial variable.

Tableau 1

c_B	a_B	a_1	a_2	a_3	a_4	a_5	b	b/a_{i1}
0	a_3	1^*	2	1	0	0	1	1→
M	a_5	4	3	0	–1	1	6	3/2
	c	–2	–3	0	0	M		
	$z^*_j - c_j$	4M+2	3M+3	0	–M	0		
		↑						

Tableau 2

c_B	a_B	a_1	a_2	a_3	a_4	a_5	b
–2	a_1	1	2	1	0	0	1
M	a_5	0	–5	–4	–1	1	2
	c	–2	–3	0	0	M	
	$z^*_j - c_j$	0	–5M–1	–4M–2	–M	0	

From tableau 2, we find that all $z^*_j - c_j \leq 0$, but the artificial variable x_5 appears in the optimal basis at a positive level. Hence the original problem has no feasible solution.

12.6. Inconsistency and Redundancy

The method of artificial variables which enables us to obtain an initial basic feasible solution to a linear programming problem without any assumption about the system of constraints, also provides us information about the inconsistency and redundancy of the problem.

Suppose that the two-phase method is used. If at the end of phase 1, all the artificial variables are nonbasic, we have a basic feasible solution of the original problem, the basis containing only column vectors of A and consequently r(A) = m. The constraints are therefore consistent and there is no redundancy in the system.

If at the end of phase 1, the optimal value of the auxiliary problem is positive, the original problem is not feasible and hence the constraints are inconsistent.

If the optimal value of the auxiliary problem is zero and at least one of the artificial variables appears in the basis at a zero level, we have a feasible solution of the original problem and therefore the original constraints are consistent.

If we further find that corresponding to the artificial basic variable w_r which is at a zero level there exists at least one $\alpha_{ij} \neq 0$ (positive or negative) say $\alpha_{rs} \neq 0$, we can replace the vector a_r, by the vector a_s. The new solution thus obtained will again be a basic feasible solution with $\bar{x}_s = 0$ and the values of nonzero basic variables as well as the value of the objective function will remain unchanged. If this process can be continued until all the artificial basic variables are removed from the basis, we will have a degenerate basic feasible solution involving only the original variables. Consequently r(A) = m and there is no redundancy in the system.

If however, it is not possible to remove all the artificial vectors from the basis by the procedure above, we must reach a stage where corresponding to each artificial basic vector a_i, α_{ij} are zero for every j. Suppose that k artificial vectors remain in the basis at a zero level. Every column vector of A can then be expressed as a linear combination of the (m – k) vectors of A in the basis. It is therefore clear that r(A) = m – k and k of the original constraints are redundant.

(Analogous conclusions may be derived from the method of penalties also).

12.7. Exercises

1. Show how the simplex method can be used to find a nonnegative solution of a system of linear equations AX = b.
2. Explain how the simplex method indicates whether a linear programming problem is inconsistent.
3. When does the simplex method indicate that the system of constraints in a linear programming problem is redundant?
4. Solve the following problems by the two-phase methods.

(a) Maximize $8x_1 + 5x_2$
Subject to
$$x_1 + x_2 \le 5$$
$$4x_1 + x_2 \ge 4$$
$$2x_1 + 3x_2 \ge 3$$
$$x_1, x_2 \ge 0$$

(b) Maximize $3x_1 + 4x_2$
Subject to
$$3x_1 + x_2 \le 9$$
$$-2x_1 + x_2 \le 4$$
$$x_1 + 2x_2 \ge 2$$
$$x_1, x_2 \ge 0$$

(c) Minimize $5x_1 + 4x_2$
Subject to
$$x_1 + 2x_2 \le 6$$
$$2x_1 + x_2 \le 5$$
$$4x_1 + x_2 \ge 2$$
$$x_1 + x_2 \ge 1$$
$$x_1, x_2 \ge 0$$

(d) Minimize $5x_1 - 2x_2$
Subject to
$$x_1 - x_2 + x_3 \le 2$$
$$3x_1 - x_2 - x_3 \ge 3$$
$$x_1, x_2, x_3 \ge 0$$

(e) Maximize $3x_1 + 2x_2 + 3x_3$
Subject to
$$2x_1 + x_2 + x_3 \le 2$$
$$3x_1 + 4x_2 + 2x_3 \ge 8$$
$$x_1, x_2, x_3 \ge 0$$

5. Using the two-phase method, show that the following linear programming problem is inconsistent.

Minimize $x_1 + x_2$

Subject to $x_1 + 2x_2 \leq 2$

$3x_1 + 4x_2 \geq 12$

$x_1, x_2 \geq 0$

6. Solve the following linear programming problems by the penalty method.

(a) Minimize $x_1 - 3x_2$

Subject to $3x_1 + x_2 \leq 3$

$x_1 + 2x_2 \geq 2$

$x_1 \leq 4$

$x_1, x_2 \geq 0.$

(b) Minimize $2x_1 + x_2 - x_3$

Subject to $2x_1 + x_2 - x_3 \geq 6$

$3x_1 + x_2 + x_3 \geq 4$

$x_2 + x_3 \leq 12$

$x_1, x_2, x_3 \geq 0$

(c) Minimize $3x_1 + 4x_2 + 8x_3$

Subject to $x_1 + 2x_2 \geq 5$

$x_2 + x_3 \geq 2$

$x_1, x_2 \geq 0$

(d) The problem in Q4(b).

7. Using the simplex method find the inverse of the matrix

(a) $\begin{bmatrix} 4 & 1 & 2 \\ 0 & 1 & 0 \\ 8 & 4 & 5 \end{bmatrix}$ (b) $\begin{bmatrix} 5 & 1 & 7 \\ 3 & 4 & 8 \\ 2 & 6 & 8 \end{bmatrix}$

8. Apply the simplex method to find a solution of the system of equations

(a) $x_1 + x_2 + 4x_3 + 2x_4 = 9$

$x_1 + x_2 + 2x_3 - x_4 = 3$

$2x_1 + 2x_2 + 4x_3 + x_4 = 12$

$x_1, x_2, x_3, x_4 \geq 0$

(b) $3x_1 - 2x_2 - x_4 = 1$

$x_1 - x_3 + 4x_4 = 3$

$2x_1 - x_2 = 3$

$x_1, x_2, x_3, x_4 \geq 0$

Chapter 13

Degeneracy in Linear Programming

13.1. Introduction

A degenerate basic feasible solution to a linear programming problem is one in which one or more basic variables are equal to zero. We know that the simplex algorithm is an iterative procedure which moves from one basis to another and finds an optimal solution at a basic feasible solution. The assumption of nondeneracy in the simplex method is necessary to show that for each successive admissible basis, the associated value of the objective function is better than the preceding one so that no basis is repeated. Consequently, since there are only a finite number of basis, the simplex procedure reaches an optimal solution (or indicates that there is an unbounded solution) in a finite number of steps. In the presence of degeneracy, the proof of convergence breaks down.

A basic variable becomes zero, when the minimum calculated by the simplex exit-criterion is not unique or when it enters the basis to replace a variable already zero. For example, if a_k were the vector to enter the basis and if

$$\theta_0 = \min_{i/a_{ik}>0} \frac{x_{i0}}{\alpha_{ik}} = \frac{x_{r0}}{\alpha_{rk}} \tag{13.1}$$

the vector a_r would leave the basis.

If the minimum in (13.1) occurs for x_{s0} also, we get a degenerate basic feasible solution in the next iteration. If x_{r0} is already zero, we again get a degenerate solution. In the presence of degeneracy, there is no improvement in the value of the objective function.

It is therefore possible that the same sequence of bases is selected repeatedly without reaching an optimal solution and an endless cycle starts. Since cycling in the simplex algorithm is only possible under degeneracy and degeneracy is quite a frequent phenomenon, it might be thought that there would be many cases of cycling. But in actual practice, cycling never seems to occur except in the specially constructed examples of Hoffman [232] and Beale [38].

In the absence of cycling, degeneracy is not a difficulty in itself. Although, the value of the objective function remains the same for a number of iterations, in practice, an optimal solution is reached in a finite number of steps. For this reason, most instruction codes for electronic computers do not include rules which guarantee convergence. Theoretically, however, it is at least possible for a problem to cycle and therefore it is desirable to develop a procedure that will ensure that cycling will never occur.

Two best known procedures for the resolution of degeneracy are the perturbation method, developed by Charnes [67] and the lexicographic method developed by Dantzig, Orden and Wolfe [117]

We present here the method suggested by Charnes.

13.2. Charnes' Perturbation Method

We note that degeneracy in a linear programming problem occurs when 'b', the requirement vector of the problem cannot be expressed as a positive linear combination of the basis vectors of some bases formed from the columns of A. It is then expected that if 'b' were perturbed, it might be possible to express the perturbed 'b' as a positive linear combination of each basis vector of every feasible basis and from the solution of this nondegenerate perturbed problem, we may then be able to get a solution of our original problem.

Consider the linear programming problem (11.1), where the constraints are

$$x_1a_1 + x_2a_2 + \dots + x_na_n = b, \tag{13.2}$$

a_j being the j the column of A.

We now perturb the vector 'b' and rewrite the constraints as,

$$x_1a_1 + x_2a_2 b+ \dots + x_na_n = b + \epsilon a_1 + \epsilon^2 a_2 + \dots + \epsilon^n a_n = b(\epsilon) \tag{13.3}$$

where ϵ is a sufficiently small positive number which is adjusted so as to make (13.3) nondegenerate. We will see later that it is not necessary to know the value of ϵ precisely. Only the knowledge of existence of such an ϵ is sufficient for our purpose. The numerical value of ϵ is never involved in our computation. Once a solution to the perturbed problem has been obtained, by setting $\epsilon = 0$, we have the corresponding solution of the original problem.

Let the vectors $a_1, a_2, a_3 \dots a_m$ form an admissible basis B, that is, the basis that yields a feasible solution (the problem can be so arranged that the basis vectors are the first m vectors). Then a basic feasible solution to (13.2) is given by,

$$X_0 = B^{-1}b \geq 0 \tag{13.4}$$

and a solution to (13.3) by

$$X_0(\epsilon) = B^{-1}b(\epsilon)$$

or

$$X_0(\epsilon) = X_0 + \epsilon B^{-1}a_1 + \epsilon B^{-1}a_2 + \dots + \epsilon^n B^{-1}a_n \tag{13.5}$$

Hence x_{i0} (ϵ) can be written as,

$$x_{i0}(\epsilon) = x_{i0} + \sum_{j=1}^{n} \epsilon^j \alpha_{ij}, \qquad i = 1, 2, ..m \tag{13.6}$$

Since B consists of vectors $a_1, a_2, ..a_m$, $B^{-1}a_j$ is an identity vector with the unit element in the jth position for j = 1,2, .. m. Hence

$$x_{i0}(\epsilon) = x_{i0} + \epsilon^i + \sum_{j=m+1}^{n} \epsilon^j \alpha_{ij}, \qquad i = 1,2,..m \tag{13.7}$$

It can be shown that

Lemma 13.1

There exists a range of values of ϵ, $0 < \epsilon < \epsilon_0$ for which an ϵ-polynomial, $f(\epsilon) = p_0 + p_1\epsilon + ... + p_m\epsilon^m$ is positive, if and only if the leading term in $f(\epsilon)$ is positive.

Proof: Exercise

Hence by taking $\epsilon > 0$ but sufficiently small, in (13.7), we can make all $x_{i0}(\epsilon) > 0$, i = 1,2,...m.

Since $z_j - c_j$ are the same for both the perturbed and the original problems and are independent of ϵ, the same vector may be introduced in the basic set. Let the vector a_k be introduced in the basis and if all $\alpha_{ik} \leq 0$, then there is an unbounded solution to the perturbed problem and to the original problem. If at least one $\alpha_{ik} > 0$, we select the vector to be eliminated from the basis by the usual exist criterion

$$\theta_0 = \operatorname*{Min}_i \left\{ \frac{x_{i0}(\epsilon)}{\alpha_{ik}}, \quad \alpha_{ik} > 0 \right\} \tag{13.8}$$

Obviously, $\theta_0 > 0$, since each $x_{i0}(\epsilon) > 0$. If this minimum is unique, we have a nondegenerate basic feasible solution to the perturbed problem.

We now rewrite θ_0 as

$$\theta_0 = \operatorname*{Min}_i \left\{ \frac{x_{i0} + \sum_{j=1}^{n} \epsilon^j \alpha_{ij}}{\alpha_{ik}}, \quad \alpha_{ik} > 0 \right\} \tag{13.9}$$

and show by applying Lemma 13.2 given below, that a unique minimum in (13.9) can really be obtained, so that degeneracy does not appear in the next iteration.

Lemma 13.2

If two polynomials

$$f(\epsilon) = \sum_{i-0}^{m} p_i \epsilon^i \quad \text{and} \quad g(\epsilon) = \sum_{i=0}^{m} q_i \epsilon^i \quad \text{are such that}$$

$$p_i = q_i, \; i = 0, 1,2.. k-1$$

$$p_k < q_k$$

and p_i, q_i are arbitrary for i > k, then for some $\epsilon_0 > 0$, $f(\epsilon) < g(\epsilon)$, for all $0 < \epsilon < \epsilon_0$

Proof: Exercise

The procedure for finding the minimum may now be stated as follows:

We first compare the ratios x_{i0}/α_{ik}. If the minimum is obtained for unique i, say for i = r, then for $\in < \in_0$, the minimum in (13.9) occurs at i = r and the vector a_r is uniquely determined to be eliminated. If however, there are ties for some set of indices, we compare the ratios α_{i1}/α_{ik} (i.e. the coefficients of $\in^j$, for j = 1 in (13.9)) for those indices. If the minimum occurs at i = s (s belongs to the tied set), then a_s is the vector to be eliminated. If there are still ties, we compare the ratios for the tied set for j = 2 and repeat the process.

For example, if we have

$$\theta_0 = \frac{x_{s0}}{\alpha_{sk}} = \frac{x_{t0}}{\alpha_{tk}}$$

we compute $$\text{Min}\ \left\{\frac{\alpha_{s1}}{\alpha_{sk}}, \frac{\alpha_{t1}}{\alpha_{tk}}\right\} \tag{13.10}$$

If minimum in (13.10) is α_{s1}/α_{sk}, then the vector a_s is eliminated and if minimum is given by α_{t1}/α_{tk}, then a_t is eliminated.

If however, $\alpha_{s1}/\alpha_{sk} = \alpha_{t1}/\alpha_{tk}$, we compare the ratios (the coefficients of $\in^j$ for j = 2 in (13.9))

$$\frac{\alpha_{s2}}{\alpha_{sk}} \text{ and } \frac{\alpha_{t2}}{\alpha_{tk}}$$

and find the vector to be eliminated from the basis. If still there is a tie, we proceed to compare the ratios for j = 3 and so on.

It is certain that eventually a unique minimum must be obtained. If it is not, it would imply that two or more rows of the inverse of the basis are proportional to each other. This however, is not possible as the basis matrix is of rank m. The procedure therefore, ensures that a new basic feasible solution is obtained with all basic variables strictly positive.

The above procedure is applied in every iteration and thus degeneracy can never appear in the perturbed problem.

Moreover, since,

$$\hat{z}(\in) = z(\in) - \theta_0(z_j - c_j) < z(\in), \ \ \theta_0 > 0.$$

in each iteration, no basis can be repeated and an optimal solution is obtained to the perturbed problem in a finite number of steps. Setting $\in = 0$, we then have an optimal solution to the original problem since $z_j - c_j$ does not depend on the value of $\in$.

It should be noted from the above analysis that the numerical value of $\in$ is never involved in our calculations. The information required to determine θ_0 is the values of the basic variables of the original problem and the coefficients of $\in^j$ in

(13.9). This information is all contained in the simplex tableau of the original problem. The $\in$-perturbation was introduced only to develop a procedure for preventing cycling and this we find, is done without ever using $\in$ explicitly.

To apply the perturbation technique therefore, only the simplex tableau of the original problem is needed.

13.3. Example

Consider the problem suggested by Beale.

$$\text{Minimize} \qquad z = -\frac{3}{4}x_1 + 150x_2 - \frac{1}{50}x_3 + 6x_4$$

$$\text{Subject to} \qquad \frac{1}{4}x_1 - 60x_2 - \frac{1}{25}x_3 + 9x_4 + x_5 \qquad = 0$$

$$\frac{1}{2}x_1 - 90x_2 - \frac{1}{50}x_3 + 3x_4 + x_6 \qquad = 0$$

$$x_3 + x_7 \qquad = 1$$

$$x_j \geq 0, \qquad i = 1,2,..7.$$

It can be shown that the problem cycles, if we use the simple rule that we select the vector to be eliminated, the one whose row index is the smallest among those tied. However, this cycling can be avoided by applying the perturbation technique as is shown below.

Rearranging the problem so that the basis vectors appear in the beginning of the tableau, we have

Tableau 1

c_B	a_B	a_1	a_2	a_3	a_4	a_5	a_6	a_7	x_B
0	a_1	1	0	0	1/4	–60	–1/25	9	0
0	a_2	0	1	0	1/2*	–90	–1/50	3	0
0	a_3	0	0	1	0	0	1	0	1
	c	0	0	0	–3/4	150	–1/50	6	
	z_j-c_j	0	0	0	3/4	–150	1/50	–6	
					↑				

To determine which vector is to depart from the basis, we compute

$$\underset{i}{\text{Min}}\left\{\frac{x_{Bi}}{a_{i4}},\ a_{i4} > 0\right\} = \frac{x_{B1}}{a_{14}} = \frac{x_{B2}}{a_{24}} = 0$$

Thus there is a tie between the basis vectors a_1 and a_2.

We therefore compute

$$\text{Min}\left\{\frac{a_{11}}{a_{14}},\ \frac{a_{21}}{a_{24}}\right\} = \text{Min}\left\{\frac{1}{1/4},\ \frac{0}{1/2}\right\} = 0$$

Since minimum occurs for a_{21}/a_{24}, the vector a_2 is to be eliminated.

Tableau 2

c_B	a_B	a_1	a_2	a_3	a_4	a_5	a_6	a_7	x_B
0	a_1	1	−1/2	0	0	−15	−3/100	15/2	0
−3/4	a_4	0	2	0	1	−180	−1/25	6	0
0	a_3	0	0	1	0	0	1^*	0	1
	c	0	0	0	−3/4	150	−1/50	6	
	$z_j - c_j$	0	−3/2	0	0	−15	1/20 ↑	−21/2	

It is now obvious that the vector a_6 is to enter and a_3 is to leave the basis.

Tableau 3

c_B	a_B	a_1	a_2	a_3	a_4	a_5	a_6	a_7	x_B
0	a_1	1	−1/2	3/100	0	−15	0	15/2	3/100
−3/4	a_4	0	2	1/25	1	−180	0	6	1/25
−1/50	a_6	0	0	1	0	0	1	0	1
	c	0	0	0	−3/4	150	−1/50	6	
	$z_j - c_j$	0	−3/2	−1/20	0	−15	0	−21/2	

The optimality condition is satisfied in the tableau 3 and an optimal basic solution to the problem is

$$x_1 = 1/25,\ x_2 = 0,\ x_3 = 1,\ x_4 = 0$$

and Min z = −1/20.

13.4. Exercises

1. Show that in Beale's problem (see example), the phenomenon of cycling occurs if in case of a tie in the simplex exist criterion, the vector to be eliminated in selected whose row index is the smallest among those tied.
2. Prove Lemma 13.1 and Lemma 13.2.
3. Solve the following linear programming problems.

(a) Minimize $z = 2x_1 - 3x_2 - 5x_3$

Subject to

$$2x_1 - x_2 + 2x_3 \geq 2$$
$$x_1 + 3x_2 \leq 5$$
$$-4x_1 + 3x_3 \leq 3$$
$$x_1, x_2, x_3 \geq 0$$

(b) Maximize $z = x_1 + 2x_2$

$$x_1 + 4x_2 \leq 8$$
$$3x_1 + 4x_2 \leq 12$$
$$-x_1 + 4x_2 \leq 8$$
$$x_1, x_2 \geq 0$$

(c) Maximize $z = 3x_1 + 5x_2$

Subject to

$$8x_1 + 3x_2 \leq 12$$
$$x_1 + x_2 \leq 2$$
$$2x_1 + 5x_2 \leq 10$$
$$x_1, x_2 \geq 0$$

Chapter 14

The Revised Simplex Method

14.1. Introduction

Soon after the simplex method came into practice, it was realized that the method requires to compute and record many numbers in each iteration which are not all needed in the subsequent steps. The simplex procedure may therefore be very time consuming even when solved on a computer. It may be noted that to reach a decision at any iteration, of the simplex method the following quantities are needed for a given basis.

(i) $z_j - c_j$, to determine the vector to be introduced into the basis or to test optimality for the current solution.

(ii) The column corresponding to the variable entering the basic set to determine the vector to be eliminated from the basis

(iii) The values of the basic variables.

It should be observed that in each iteration the information above can be obtained from the original date consisting of A, b and c and B^{-1}, the inverse of the current basis B only. This fact has led to the development of a computational procedure to solve the general linear programming problem known as the revised simplex method [115, 117].

14.2. Outline of the Procedure

Consider the problem,

$$\begin{aligned} &\text{Maximize} && z = c^T X \\ &\text{Subject to} && AX = b \\ & && X \geq 0 \end{aligned} \tag{14.1}$$

where c is a $n \times 1$, $b \geq 0$, is $m \times 1$ and A is $m \times n$ matrices.

The general approach of the revised simplex method is the same as in the original simplex method. The differences lie in the manner of calculations which are made to move from one iteration to the next. We shall first consider the case when an initial basis is known and then the case when the initial basic variables are artificial.

14.2.1 Case 1: Initial basis is known

In the revised simplex method, the objective function is added to the set of constraints in the form of a linear equation, thus forming an adjoined problem. In this case therefore, we deal with an (m + 1) dimensional basis instead of an m-dimensional basis as in the ordinary simplex method.

The problem becomes,

$$\text{Maximize} \quad z$$

$$\text{Subject to} \quad z - \sum_{j=1}^{n} c_j x_j = 0 \tag{14.2}$$

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = 1,2,..m$$

$$x_j \geq 0, \quad j = 1,2,..n$$

z unrestricted.

The problem can be restated as

$$\text{Maximize} \quad e_1^T \hat{X} \tag{14.3}$$

$$\text{Subject to} \quad \hat{A}\hat{X} = \hat{b}$$

where $\hat{A} = (e_1 \hat{a}_1 \hat{a}_2 ... \hat{a}_n)$ an $(m + 1) \times (n + 1)$ matrix

$\hat{a}_j^T = \left[-c_j, a_j^T\right]$, a_j being the jth column of A

$\hat{b}^T = \left[0, b^T\right]$,

$e_1^T = \left[1, 0^T\right]$ and $\hat{X}^T = \left[z, X^T\right], X \geq 0.$

Let B be an initial basis such that $BX_B = b$, $X_B \geq 0$ and thus $X_B = B^{-1}b$ is an initial basic feasible solution of the original problem. It can be easily seen that corresponding to every basis B of the original problem, we have a basis $\hat{B}$ for the adjoined problem and is given by

$$\hat{B} = \begin{bmatrix} 1 & -c^T B \\ 0 & B \end{bmatrix} \tag{14.4}$$

From the inversion formula for partitioned matrices, it follows that

$$\hat{B}^{-1} = \begin{bmatrix} 1 & c_B^T B^{-1} \\ 0 & B^{-1} \end{bmatrix} \tag{14.5}$$

and $\quad \hat{X}_B = \hat{B}^{-1}\hat{b}$

$$= \begin{bmatrix} 1 & c_B^T B^{-1} \\ 0 & B^{-1} \end{bmatrix} \begin{bmatrix} 0 \\ b \end{bmatrix}$$

$$= \begin{bmatrix} c_B^T B^{-1} b \\ B^{-1} b \end{bmatrix} = \begin{bmatrix} z \\ X_B \end{bmatrix} \tag{14.6}$$

Now, to see whether the solution is optimal or not, we require to know the values of $z_j - c_j$. We note that

$$\hat{\alpha}_j = \hat{B}^{-1} \hat{a}_j$$

$$= \begin{bmatrix} 1 & c_B^T B^{-1} \\ 0 & B^{-1} \end{bmatrix} \begin{bmatrix} -c_j \\ a_j \end{bmatrix}$$

$$= \begin{bmatrix} c_B^T B^{-1} a_j - c_j \\ B^{-1} a_j \end{bmatrix} = \begin{bmatrix} z_j - c_j \\ \alpha_j \end{bmatrix} \tag{14.7}$$

Thus $z_j - c_j$ are obtained by multiplying the first row of $\hat{B}^{-1}$ by the $\hat{a}_j$ and α_j are obtained by multiplying the last m rows of $\hat{B}^{-1}$ by $\hat{a}_j$.

The results obtained in (14.6) and (14.7) can however be obtained by single operation

$$\hat{B}^{-1}\left(\hat{A}, \hat{b}\right) = \begin{bmatrix} 1 & z_1 - c_1 & z_1 - c_2 & \dots & z_n - c_n & z \\ 0 & \alpha_1 & \alpha_2 & & \alpha_n & X_B \end{bmatrix} \tag{14.8}$$

If $z_j - c_j$ satisfy the optimality condition we have an optimal solution, if not, the vector $\hat{a}_k$ to be introduced in the basis is determined from

$$z_k - c_k = \text{Min } (z_j - c_j),\ z_j - c_j < 0. \tag{14.9}$$

The vector $\hat{a}_r$, to be removed from the basis is determined from

$$\underset{i}{\text{Min}} \left\{ \frac{x_{Bi}}{\alpha_{ik}}, \quad a_{ik} > 0 \right\} = \frac{x_{Br}}{\alpha_{rk}} \tag{14.10}$$

[If there is a tie, it may be resolved by any standard method.]

We now need to perform the transformation to get the new basic feasible solution and check for its optimality. As already discussed, in the revised simplex method only $\hat{B}^{-1}$ and $\hat{X}_B$ are required to be transformed. This is most conveniently done by the method of product form of the inverse [See Chapter 4].

14.2.2 Case 2. Initial basis consists of artificial variables

Suppose that the original matrix does not contain any unit vector and the initial basis consists of the artificial vectors only, forming an identity matrix. The two-phase method is then used to eliminate the artificial variables from the basis. In Phase I, the artificial variables are driven to zero and in Phase II an optimal solution to the original problem is obtained.

Suppose that the variables are renumbered such that $x_2, x_3, \ldots x_{m+1}$ are the artificial variables and the original variables are $x_{m+x}, x_{m+3}, \ldots x_{m+n+1}$.

In Phase-I instead of minimizing the sum of the artificial variables, we equivalently maximize

$$z^* = \sum_{i=2}^{m+1}(-x_i) \tag{14.11}$$

and setting

$a_{ij} = a_{i+1,\, j+m+1}$; $b_i = b_{i+1}$, $i = 1,2,..m$; $j = 1,2,..n$ and replacing z^* by x_1 we have the problem,

$$\begin{array}{llll}
 & \text{Maximize } x_1 & & \\
\text{Subject to} & x_1 + x_2 + \ldots + x_{m+1} & = 0. & \\
 & x_2 \qquad + \displaystyle\sum_{j=m+2}^{m+n+1} a_{2j}x_j & = b_2 & \\
 & \quad \ddots & & \\
 & \qquad x_{m+1} + \displaystyle\sum_{j=m+2}^{m+n+1} a_{m+j}x_j & = b_{m+1} & (14.12) \\
 & & x_j \geq 0, \quad j = 2,3,\ldots m+n+1 &
\end{array}$$

where $b_i \geq 0$, $i = 2,.. m+1$.

Note that the variable x_1 should always remain in the basis.

It is clear that x_1 will never have a positive value, since $\sum_{i=2}^{m+1} x_i \geq 0$

It max $x_1 < 0$, there is no feasible solution to the original problem and if max $x_1 = 0$, we have a basic feasible solution and we move to Phase II.

In Phase II, we are to maximize

$z = \sum_{j=m+2}^{m+n+1} c_j x_j$, where c_j of (14.1) is now denoted by c_{j+m+1}, $j = 1,2,..n$

$$\begin{array}{lll}
\text{Subject to} & & \\
x_2 \qquad + \displaystyle\sum_{j=m+2}^{m+n+1} a_{2j}x_j & = b_2 & \\
\quad \ddots & \vdots & \\
x_{m+1} + \displaystyle\sum_{j=m+2}^{m+n+1} a_{m+1}x_j & = b_{m+1} & (14.13) \\
x_j \geq 0, & j = 2 \ldots m+n+1 &
\end{array}$$

where $b_i \geq 0$, $i = 2,...m+1$.

We however, assume that at the optimum solution in Phase I, one or more artificial variables may remain in the basis at a zero level and therefore we must take care that these artificial variables remain zero in all subsequent iterations. For this purpose, we insert

$$x_2 + x_3 + .. + x_{m+1} = 0 \tag{14.14}$$

in the set of constraints. Moreover, at the end of Phase I, $x_1 = 0$ and hence it can be added to (14.14), since the simplex method will always maintain it at a zero level.

Now, setting $z = x_0$, $c_j = -a_{0j}$, we can write the set of constraints in a more symmetric form

$$\begin{array}{llll}
x_0 & & + \sum_{j=m+2}^{m+n+1} a_{0j}x_j & = 0 \\
& x_1 + x_2 + ... + x_{m+1} & & = 0 \\
& \quad x_2 & + \sum_{j=m+2}^{m+n+1} a_{2j}x_j & = b_2 \\
& \quad\ \ddots & & \\
& \qquad x_{m+1} & + \sum_{j=m+2}^{m+n+1} a_{m+1j}x_j & = b_{m+1}
\end{array} \tag{14.15}$$

Subtracting the sum of the m last equations from the second we have

$$\begin{array}{llll}
x_0 & & + \sum_{j=m+2}^{m+n+1} a_{0j}x_j & = 0 \\
& x_1 & - \sum_{i=2}^{m+1} \sum_{j=m+2}^{m+n+1} a_{ij}x_j & = - \sum_{i=2}^{m+1} b_i \\
& \quad x_2 & + \sum_{j=m+2}^{m+n+1} a_{2j}x_j & = b_2 \\
& \quad\ \ddots & & \\
& \qquad x_{m+1} & + \sum_{j=m+2}^{m+n+1} a_{m+1j}x_j & = b_{m+1}
\end{array} \tag{14.16}$$

and find that a unit basis of order (m + 2) is given by the first (m + 2) columns of (14.16)

We now set

$$a_{1j} = \sum_{i=2}^{m+1} (-a_{ij}), \qquad j = (m + 2)...(m + n + 1)$$

$$b_1 = \sum_{i=2}^{m+1} (-b_i) < 0 \tag{14.17}$$

$$b_0 = 0$$

and the system of equations can finally be expressed as

$$
\begin{aligned}
x_0 \qquad\qquad & + \sum_{j=m+2}^{m+n+1} a_{0j} x_j = b_0 \\
x_1 \qquad & + \sum_{j=m+2}^{m+n+1} a_{1j} x_j = b_1 \\
\ddots \quad & \\
x_{m+1} & + \sum_{j=m+2}^{m+n+1} a_{m+1j} x_j = b_{m+1}
\end{aligned}
\tag{14.18}
$$

The set of equations (14.18) can now be conveniently used in Phase I as well as m Phase II.

In Phase I, we

$$
\begin{aligned}
&\text{Maximize} \quad x_1 \\
&\text{Subject to the set of equations (14.18)} \\
&x_j \geq 0, j = 2,.. (m + n + 1).
\end{aligned}
\tag{14.19}
$$

and in Phase II, we

$$
\begin{aligned}
&\text{Maximize} \quad x_0 \\
&\text{Subject to the set of equations (14.18)} \\
&x_j \geq 0, j = 1,2,.. (m = n + 1)
\end{aligned}
\tag{14.20}
$$

where $b_0 = 0$, $b_1 < 0$ and $b_i \geq 0$, $i = 2,..(m + 1)$

The matrix of the system of equations (14.18) is given by

$$
\hat{A} = \left[\begin{array}{cc|c|cc}
1 & 0 & 0 & a_{0,m+2} & a_{0,m+n+1} \\
0 & 1 & & a_{1,m+2} & a_{1,m+n+1} \\
\hline
\multicolumn{2}{c|}{0} & I_m & \multicolumn{2}{c}{A}
\end{array}\right]
\tag{14.21}
$$

and the inverse of the basis matrices obtained from $\hat{A}$ for Phase I and Phase II enable us to derive the criteria to proceed with the iterative process to solve the problem.

Computation: Phase I

During Phase I, the variables x_0 and x_1 are never to leave the basis. The basis matrix $\hat{B}$ therefore always contains the first two columns of $\hat{A}$.

Setting $c^*_j = -a_{1j}$, we have the basis matrix as,

$$\hat{B} = \left[\begin{array}{cc|c} 1 & 0 & -c_B \\ 0 & 1 & -c^*_B \\ \hline \multicolumn{2}{c|}{0} & B \end{array}\right] \tag{14.22}$$

where B is the basis of the original problem and c^*_B, c_B represent the price vectors for the problems in Phase I and Phase II respectively.

By the formula of inversion of matrices by partition we have,

$$\hat{B}^{-1} = \left[\begin{array}{cc|c} 1 & 0 & c_B B^{-1} \\ 0 & 1 & c^*_B B^{-1} \\ \hline \multicolumn{2}{c|}{0} & B^{-1} \end{array}\right] \tag{14.23}$$

We note that the product of the second row of $\hat{B}^{-1}$ with $\hat{a}_j$ give $z_j^* - c_j^*$ from which we obtain the entering vector $\hat{a}_k$ by the usual criterion

$$z_k^* - c_k^* = \underset{j}{\text{Min}}\left[z_j^* - c_j^*;\ z_j^* - c_j^* < 0\right] \tag{14.24}$$

The product of the last m rows of $\hat{B}^{-1}$ with $\hat{a}_k$ give α_k and X_B is obtained from the product of the same row with $\hat{b}$ from which we find the vector to be removed from the basis by the simplex exit criterion

$$\frac{x_{B_r}}{\alpha_{rk}} = \underset{i}{\text{Min}}\left\{\frac{x_{Bi}}{\alpha_{ik}},\quad \alpha_{ik} > 0\right\},\qquad i = 2, \ldots m+1 \tag{14.25}$$

Now, to move to the next iteration, we only need to transform $\hat{B}^{-1}$. This can be efficiently done by the method of 'product form of the inverse'. (See Chapter 4).

Phase II: During Phase II, only the first column of $\hat{A}$ is never to leave the basis and x_1 is treated as any other artificial variable

The basis matrix $\hat{B}$ is given by,

$$\hat{B} = \left[\begin{array}{c|c} 1 & -c_B \\ \hline 0 & B_2 \end{array}\right] \tag{14.26}$$

where the set of constraints from B_2 includes the m original constraints and the second equation of (14.18). x_1 is now required to be nonnegative and the second equation of (14.18) ensures that x_1 and all other artificial variables remain at a zero level throughout the calculation. The basis matrix B_2 is of order (m + 1) and always includes at least one of the artificial vectors because otherwise it is not possible for $\hat{B}$ to have an (m + 2) dimensional basis.

$\hat{B}^{-1}$ is then determined as

$$\hat{B}^{-1} = \left[\begin{array}{c|c} 1 & c_B B_2^{-1} \\ \hline 0 & B_2^{-1} \end{array}\right] \tag{14.27}$$

As in Phase I, $z_j - c_j$ are obtained from the product of the first row of $\hat{B}^{-1}$ with $\hat{a}_j$ and the entering vector $\hat{a}_k$ is then obtained by the criterion,

$$z_k - c_k = \underset{j}{\text{Min}}\ [z_j - c_j;\ z_j - c_j < 0]$$

By multiplying the (m + 1) last rows of $\hat{B}^{-1}$ by $\hat{a}_k$ we get α_k and by multiplying it by $\hat{b}$, x_B is obtained. The variable which leaves the basis is then obtained by the usual simplex criterion. At each iteration therefore, we only need to transform the adjoined basis and as in phase I, this can be done by the method of 'product form of the inverse'

14.3. Example

$$\begin{array}{ll} \text{Minimize} & z = x_1 + 2x_2 \\ \text{Subject to} & 2x_1 + 5x_2 \geq 6 \\ & x_1 + x_2 \geq 2 \\ & x_1, x_2 \geq 0 \end{array}$$

Introducing the surplus variables x_3, x_4 and the artificial variables x_{a_1}, x_{a_2}, the problem can be converted to

$$\begin{array}{ll} \text{Maximize} & x_{a_0} = -x_1 - 2x \\ \text{Subject to} & x_{a_1} + 2x_1 + 5x_2 - x_3 = 6 \\ & x_{a_2} + x_1 + x_2 - x_4 = 2 \\ & x_1, x_2, x_3, x_4, x_{a_1}, x_{a_2} \geq 0 \end{array}$$

As indicated in (14.16) section 14.2.2, after subtracting the sum of the two equations from $x_{a_0} + x_{a_1} + x_{a_2} = 0$, the problem can be rewritten as:

$$\begin{array}{l} x_0 + x_1 + 2x_2 = 0 \\ x_{a_0} - 3x_1 - 6x_2 + x_3 + x_4 = -8 \\ x_{a_1} + 2x_1 + 5x_2 - x_3 = 6 \\ x_{a_2} + x_1 + x_2 - x_4 = 2 \\ \qquad x_j \geq 0, j = 1, 2, 3, 4 \\ \qquad x_{a_i} \geq 0, i = 1, 2 \end{array}$$

where in Phase I, we maximize x_{a_0} (= z^*) where x_0 and x_{a_0} are of arbitrary sign and in Phase II, we maximize x_0 (= z) where x_0 is of arbitrary sign and $x_{a_0} \geq 0$

The matrix $(\hat{A}, \hat{b})$, excluding the first four columns which comprise the basis matrix $\hat{B}$, is given below

Table 1

$\hat{A}$				$\hat{b}$
1	2	0	0	0
−3	−6	1	1	−8
2	5	−1	0	6
1	1	0	−1	2

Table 2

Variables in Basis	$\hat{B}^{-1}$				$\hat{B}\hat{b}$	$\hat{B}\hat{a}_2 = Y_2$
x_0	1	0	0	0	0	2
x_{a_0}	0	1	0	0	−8	−6
x_{a_1}	0	0	1	0	6	5
x_{a_2}	0	0	0	1	2	1

We now compute $z_j^* - c_j^*$, to determine which vector is to be entered in the basis. $z_j^* - c_j^*$ are the products of the second row $\hat{B}^{-1}$ with $\hat{A}$, i.e.

$$(0\ 1\ 0\ 0)\begin{pmatrix} 1 & 2 & 0 & 0 \\ -3 & -6 & 1 & 1 \\ 2 & 5 & -1 & 0 \\ 1 & 1 & 0 & -1 \end{pmatrix} = (-3,\ \ -6,\ \ 1,\ \ 1)$$

Thus $\hat{a}_2$ enters the basis.

To compute Y_2 (in conformity with the notations of the product form of inverse formula (4.21), we now use Y in place of α used in (14.8), we multiply $\hat{B}^{-1}$ by $\hat{a}_2$, and X_B = $\hat{B}^{-1}\hat{b} = \hat{b}$ in this iteration. These are recorded in the last two columns of Table 2. From the last two rows of these two columns, we determine the vector to leave the basis, i.e. from

$$\min\left(6/5, \frac{2}{1}\right) = 6/5.$$

Hence the variable x_{a_1}, i.e. the third column of $\hat{B}^{-1}$ is to be replaced by $\hat{a}_2$. $\hat{B}^{-1}$ is then transformed to $\hat{B}_2^{-1}$ by the formula (4.21) in Chapter 4. Let β_j be the jth column of $\hat{B}^{-1}$ and β_r is to be replaced by β_k. Then for the next table, we have

$$\beta_j^2 = \beta_j + \beta_{rj}(v_k - e_k) \text{ and}$$

$$X_B^2 = X_B + X_{B_r}(v_k - e_k).$$

where
$$v_k^T = \left(-\frac{y_{1k}}{y_{rk}}, -\frac{y_{2k}}{y_{rk}}, \ldots -\frac{y_{r-1,k}}{y_{rk}}, \frac{2}{y_{rk}}, -\frac{y_{r+1,k}}{y_{rk}}, \ldots, \frac{y_{mk}}{y_{rk}}\right)$$

and e_k is the kth unit vector

In our case,
$$v^T = \left(-\frac{2}{5}, \frac{6}{5}, \frac{1}{5}, -\frac{1}{5}\right)$$

Thus
$$\beta_3^2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + 1\left\{\begin{bmatrix} -2/5 \\ 6/5 \\ 1/5 \\ -1/5 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}\right\} = \begin{bmatrix} -2/5 \\ 6/5 \\ 1/5 \\ -1/5 \end{bmatrix}$$

and
$$X_B^2 = \begin{bmatrix} 0 \\ -8 \\ 6 \\ 2 \end{bmatrix} + 6\left\{\begin{bmatrix} -2/5 \\ 6/5 \\ 1/5 \\ -1/5 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}\right\} = \begin{bmatrix} -12/5 \\ -4/5 \\ 6/5 \\ 4/5 \end{bmatrix}$$

Table 3

Variables in Basis			$\hat{B}_2^{-1}$		X_B^2	Y^2
x_0	1	0	−2/5	0	−12/5	1/5
x_{a_0}	0	1	6/5	0	−4/5	−3/5
x_2	0	0	1/5	0	6/5	2/5
x_{a_1}	0	0	−1/5	1	4/5	3/5

Now from the product of the second row of $\hat{B}_2^{-1}$, with $\hat{A}$ we find

$$z_j^* - c_j^* = (-3/5, 0, -1/5, 1)$$

Hence $\hat{a}_1$, enters the basis

The vector Y_1 is obtained by multiplying $\hat{B}_2^{-1}$ by $\hat{a}_1$ and is recorded in the last column of Table 3

Hence X_{2a} leaves the basis

Table 4

Variables in Basis	$\hat{B}_3^{-1}$				X_B
x_0	1	0	–1/3	–1/3	–8/3
x_{a_0}	0	1	1	1	0
x_2	0	0	1/3	–2/3	2/3
x_1	0	0	–1/3	5/3	4/3

Since $x_{a_0} = 0$, all the artificial variables are zero and Phase 1 ends.

Now, to initiate Phase II, we take the product of the first row of $\hat{B}_3^{-1}$ with $\hat{A}$ and obtain

$$z_j - c_j = (0, 0, 1/3, 1/3)$$

Since all $z_j - c_j \geq 0$, the solution obtained in table 4 is optimal.

Hence the solution of the problem is

$$x_1 = 4/3, x_2 = 2/3 \text{ and Min } z = 8/3$$

14.4. Exercises

1. Compare the simplex and the revised simplex methods.
2. Discuss the revised simplex method for the linear program where the initial basis is not known and it is not necessary to add an artificial variable to every constraint to get an identity matrix.
3. Solve the following linear programming problems by the revised simplex method.

(a) Maximize $z = x_1 + x_2$

Subject to
$$2x_1 + 3x_2 \leq 8$$
$$x_1 + 4x_2 \leq 6$$
$$x_1, x_2 \geq 0$$

(b) Maximize $z = x_1 + 2x_2$

Subject to
$$x_1 + 6x_2 \leq 3$$
$$4x_1 + 3x_2 \leq 6$$
$$x_1, x_2 \geq 0$$

(c) Maximize $z = 3x_1 + 6x_2 - 2x_3$

Subject to
$$2x_1 + 2x_2 - x_3 \leq 2$$
$$4x_1 + x_2 \leq 4$$
$$x_1, x_2, x_3 \geq 0$$

(d) Minimize $z = x_1 + 3x_2$

Subject to
$$3x_1 + 2x_2 \geq 2$$
$$x_1 + x_2 \geq 1$$
$$x_1, x_2 \geq 0$$

(e) Maximize $z = 2x_1 + 2x_2 + x_3$

Subject to
$$3x_1 + 4x_2 - x_3 \leq 8$$
$$2x_1 + x_2 + .75x_3 \leq 3$$
$$3x_1 + x_2 + .25x_3 \leq 2$$
$$x_1, x_2, x_3 \geq 0$$

(f) Minimize $z = x_1 + 2x_2$

Subject to
$$3x_1 + 4x_2 \geq 6$$
$$2x_1 + x_2 \leq 3$$
$$x_1 + 3x_2 = 3$$
$$x_1, x_2 \geq 0$$

CHAPTER 15

Duality in Linear Programming

15.1. The concept of duality plays an important role in linear programming from a theoretical as well as a practical point of view. Associated with every linear programming problem is another closely related linear programming problem called the dual. The original problem is called the primal problem. The relations between these two problems are such that it is possible to use the optimal basic feasible solution of one problem to obtain an optimal solution for the other readily. This fact is important because situations may arise where it is more convenient to use the dual to solve a linear programming problem than the primal. The notion of duality was first introduced by J.von Neumann [488]. Subsequently Gale, Kuhn, and Tucker [185], Dantzig and Orden [116] and Goldman and Tucker [203] presented duality theorems and properties of dual linear programs.

We shall now proceed to establish duality relations between two canonical linear programs. From this we shall get a proof for the standard and the general program.

15.2. Cannonical Dual Programs and Duality Theorems

Consider the problems

$$\begin{array}{ll} \text{Maximize} & f(X) = c^T X \\ \text{Subject to} & AX \leq b \\ & X \geq 0 \end{array} \tag{15.1}$$

and

$$\begin{array}{ll} \text{Minimize} & g(y) = b^T Y \\ \text{Subject to} & A^T Y \geq c \\ & Y \geq 0 \end{array} \tag{15.2}$$

where A is an $m \times n$ matrix

The problem (15.1) is called the primal problem and the problem (15.2) the dual.

The relations between the primal and the dual problems will now be shown through a member of theorems.

Theorem 15.1. (Weak Duality Theorem)

If X and Y are feasible solutions to the primal and the dual problems respectively, then

$$f(X) \leq g(Y)$$

Proof: From the constraints of (15.1) and (15.2) we have

$$f(X) = c^T X \leq Y^T AX \leq Y^T b = b^T Y = g(Y)$$

Corollary 15.1. If X_0 and Y_0 are feasible solutions to the primal and the dual problems respectively such that $c^T X_0 = b^T Y_0$, then X_0 and Y_0 are optimal solutions to the primal and the dual problems respectively.

Proof: Let X be any feasible solution to the primal problem. By theorem 15.1, we then have

$$c^T X \leq b^T Y_0 = c^T X_0$$

and hence X_0 is an optimal solution to the primal problems. Similarly, it can be shown that Y_0 is an optimal solution to the dual.

Theorem 15.2. (Duality Theorem)

If either of the primal or the dual problems has an optimal solution, so does the other and their optimal values are equal.

Proof: Let X_0 be an optimal solution of the primal problem (15.1)

Consider the system of inequalities

$$\begin{aligned} A^T Y &\geq c \\ b^T Y &\leq c^T X_0 \\ Y &\geq 0 \end{aligned} \tag{15.3}$$

If (15.3) has a solution Y_0, then by theorem 15.1 $b^T Y_0 = c^T X_0$ and by corollary 15.1, Y_0 is then an optimal solution of the primal and the optimal values of the primal and the dual problems are equal.

Suppose that the system of inequalities (15.3) does not have a solution. It then follows from theorem 7.11 that there exists a (Z, θ) satisfying the system of inequalitites

$$\begin{aligned} &AZ - b\theta \leq 0 \\ &c^T Z - \theta\, c^T x_0 > 0 \\ &Z, \theta \geq 0. \end{aligned} \tag{15.4}$$

where θ is a single element.

Now, θ cannot be zero for if θ = 0, we have

$$\begin{aligned} AZ &\leq 0 \\ c^T Z &> 0 \\ Z &\geq 0 \end{aligned} \tag{15.5}$$

and then $(x_0 + t_z)$ is feasible for the primal problem for all $t \geq 0$, and

$$f(X_0 + t_z) = c^T(X_0 + t_z) = c^T X_0 + tc^T Z \to +\infty \text{ as } t \to +\infty, \text{ since } c^T Z > 0,$$

which contradicts the assumption that X_0 is an optimal solution of the primal problem.

Since $\theta > 0$, we note from (15.4) that z/θ is feasible to the primal problem and $c^TZ/\theta > c^TX_0$, which again contradicts that X_0 is an optimal solution of the primal problem. Thus, the system (15.4) has no solution.

Hence the system of inequalities (15.3) must have a solution Y_0 and Y_0, therefore is an optimal solution of the dual problem and their optimal values are equal.

Similarly, it can be proved that if Y_0 is an optimal solution of the dual problem (15.2), then there exists an X_0 such that X_0 is an optimal solution of the primal problem (15.1) and their optimal values are equal.

Theorem 15.3 (Unboundedness Theorem)

(a) If the primal problem is feasible while the dual is not, the objective function of the primal problem is unbounded.

(b) If the dual problem is feasible, while the primal is not, the objective function of the dual problem is unbounded.

Proof: (a) Let X_0 be a feasible solution to the primal problem (15.1). Since the dual problem is infeasible, the inequalities

$$A^TY \geq c$$
$$Y \geq 0$$

have no solution.

Hence there exists (theorem 7.11) a solution of

$$AZ \leq 0$$
$$c^TZ > 0$$
$$Z \geq 0$$

and then $X_0 + \mu Z$ is feasible for the primal problem for all $\mu \geq 0$ and

$$f(X_0 + \mu Z) = c^T(X_0 + \mu Z) = c^TX_0 + \mu c^TZ \to +\infty \text{ as } \mu \to +\infty \text{ since } c^TZ > 0.$$

(b) The proof goes exactly in the same way as in (a).

Corollary 15.2

(a) If the primal problem is feasible and f(X) is bounded above, then the dual problem is feasible

(b) If the dual problem is feasible and g(Y) is bounded below, then the primal problem is feasible.

(c) If the primal and the dual problems are both feasible, then f(X) is bounded above and g(Y) is bounded below.

Proof: (a), (b): If one problem is feasible and the other is not, then by theorem 15.3, the objective function of the feasible problem is unbounded contradicting the assumption.

(c): Suppose that the primal is unbounded. This can however, only happen if there exists a solution of

$$\begin{aligned} AZ &\leq 0 \\ c^T Z &> 0 \\ Z &\geq 0 \end{aligned}$$

in which case for any feasible solution X of the primal problem, X +tZ is also feasible for all $t \geq 0$ and $f(X + tZ) = c^T(X + tZ) = c^T X + tc^T Z \to +\infty$ for $t \to +\infty$, since $c^T Z > 0$. By theorem 7.11, it then follows that there does not exist a solution of the system

$$\begin{aligned} A^T Y &\geq c \\ Y &\geq 0 \end{aligned}$$

which implies that the dual problem is infeasible, contradicting the assumption.

Theorem 15.4. (Existence Theorem)

If both the primal and the dual problems are feasible, then both have optimal solutions and their optimal values are equal.

Proof: By corollary 15.2 (c), f(X) is bounded above on the constraint set of the primal and g(Y) is bounded below on the constraint set of the dual problem. Since the set

$$S = \{X |\ AX \leq b, X \geq 0\}$$

is a nonempty, closed bounded polyhedral convex set the linear function $c^T X$ attains its maximum value on S, that is, there exists an X_0 feasible for the primal problem such that max $c^T X = c^T X_0$ and X_0 is an optimal solution of the primal problem.

By theorem 15.2 then, there exists a Y_0 optimal for the dual problem and $c^T X_0 = b^T Y_0$.

Alternatively, theorem 15.4 can be proved as follows:

Consider the system of inequalities

$$\begin{aligned} AX &\leq b \\ A^T Y &\geq c \\ c^T X - b^T Y &\geq 0 \\ X, Y &\geq 0 \end{aligned} \tag{15.6}$$

If (15.6) has a solution (X_0, Y_0), then by theorem 15.1 $c^T X_0 = b^T Y_0$ and by corollary 15.1, X_0 and Y_0 are then optimal solutions to the primal and the dual problems respectively and their optimal values are equal.

Suppose (15.6) does not have a solution. It then follows by theorem 7.11, that there exists a (W, Z, θ) satisfying the system of inequalities.

$$\begin{aligned} A^T W - c\theta &\geq 0 \\ -AZ + b\theta &\geq 0 \\ b^T W - c^T Z &< 0 \\ W, Z, \theta &\geq 0 \end{aligned} \tag{15.7}$$

where θ is a single element

It can be seen easily that θ must be positive.

If $\theta = 0$, from the constraints of (15.1), (15.2), and (15.7) we have

$$b^TW \geq X^TA^TW \geq 0$$
$$c^TZ \leq Y^TAZ \leq 0$$

so that $b^TW - c^TZ \geq 0$ contradicting the third inequality of (15.7)

If $\theta > 0$, from (15.7) we get

$$A^TW/\theta \geq c$$
$$A\,Z/\theta \leq b$$
$$W/\theta,\ Z/\theta \geq 0$$

so that z/θ and W/θ are feasible solutions of the primal and the dual problems respectively. Hence by theorem 15.1,

$$b^TW/\theta - c^TZ/\theta \geq 0$$
$$\text{or} \quad b^TW - c^TZ \geq 0$$

which again contradicts the third inequality of (15.7).

Hence (15.7) is inconsistent and we must have a solution of (15.6) which implies, that both the primal and dual problems have optimal solutions and their optimal values are equal.

The possible status of the dual problems may be summarized as follows:

Primal; Max f(X)

		Feasible	Infeasible
Dual: Min g(Y)	Feasible	$f(X) \leq g(Y)$ Max f(x) = Min g(Y)	Min G(Y) = $\to -\infty$
	Infeasible	Max f(X) $\to +\infty$	Possible

Table 15.1

The following example shows that both the primal and the dual problems can be infeasible

Example : Consider the primal problem:

Maximize $\quad f(X) = 4x_1 - 3x_2$

Subject to $\quad = x_1 - x_2 \leq 1$

$$-x_1 + x_2 \leq -2$$
$$x_1, x_2 \geq 0$$

It is clear that the problem is infeasible. It's dual problem is

Minimize $\quad g(Y) = y_1 - 2y_2$

Subject to $\quad y_1 - y_2 \geq 4$

$$-y_1+y_2 \geq -3$$
$$y_1, y_2 \geq 0$$

and it is also infeasible.

Theorem 15.5: The dual of the dual is the primal.

Proof: Suppose that the primal problem is

$$\begin{aligned} \text{Maximize} \quad & f(X) = c^T X \\ \text{Subject to} \quad & AX \leq b, \\ & X \geq 0. \end{aligned} \tag{15.8}$$

The dual to this problem is then given by

$$\begin{aligned} \text{Minimize} \quad & g(Y) = b^T Y. \\ \text{Subject to} \quad & A^T Y \geq c. \\ & Y \geq 0. \end{aligned} \tag{15.9}$$

Now, the problem (15.9) is also a linear programming problem and we are interested in finding its dual. The problem (15.9) can easily be expressed as:

$$\begin{aligned} \text{Maximize} \quad & -g(Y) = -b^T Y \\ \text{Subject to} \quad & -A^T Y \leq -c \\ & Y \geq 0 \end{aligned} \tag{15.10}$$

Which is exactly in the form of (15.8) and hence the associated dual problem is given by

$$\begin{aligned} \text{Minimize} \quad & h(W) = -c^T W \\ \text{Subject to} \quad & -AW \geq -b \\ & W \geq 0 \end{aligned}$$

or

$$\begin{aligned} \text{Maximize} \quad & H(W) = c^T W \\ \text{Subject to} \quad & AW \leq b \\ & W \geq 0 \end{aligned} \tag{15.11}$$

The problem (15.11) which is the dual of dual problem (15.9) is just the primal problem, we had started with.

15.3. Equivalent Dual Forms

(a) Cannonical form (Symmetric form).

The duality relations between a pair of linear programming problems in cannonical form have already been established. If the problem

$$\begin{aligned} \text{Maximize} \quad & z = c^T X. \\ \text{Subject to} \quad & AX \leq b \\ & X \geq 0 \end{aligned} \tag{15.12}$$

is considered as the primal problem, then its dual is given by,

$$\begin{aligned} \text{Minimize} \quad & v = b^T Y \\ \text{Subject to} \quad & A^T Y \geq c \\ & Y \geq 0 \end{aligned} \tag{15.13}$$

The cannonical form of the dual problems is remarkable because of its symmetry and is referred to as symmetric dual programs. To see more clearly the connection between the pair of dual programs, they may be symbolized in the table below

	Primal					
	Variables	$x_1 \geq 0$	$x_2 \geq 0$	$x_n \geq 0$	Relation	Constants
Dual	$y_1 \geq 0$	a_{11}	a_{12}	a_{1n}	$\leq$	b_1
	$y_2 \geq 0$	a_{21}	a_{22}	a_{2n}	$\leq$	b_2
	$y_m \geq 0$	a_{m1}	a_{m2}	a_{mn}	$\leq$	b_m
	Relation	$\geq$	$\geq$	$\geq$		$\geq$ max z
	Constants	c_1	c_2	c_n	$\leq$ min v	

Table 15.2

(The primal problem reads across and the dual problem down)

Dual problems för the standard and the general linear programs may conveniently be obtained by first transforming them to their equivalent cannonical forms and then obtaining their dual.

(b) Standard form: The standard form of linear programming problem

$$\begin{aligned} &\text{Maximize} \quad && z = c^T X \\ &\text{Selected to} \quad && AX = b \\ & && X \geq 0 \end{aligned} \tag{15.14}$$

can be written in the cannonical form as

$$\begin{aligned} &\text{Maximize} \quad && z = c^T X \\ &\text{Subject to} \quad && AX \leq b \\ & && -AX \leq -b \\ & && X \geq 0 \end{aligned}$$

Its dual is therefore given by

$$\begin{aligned} &\text{Minimize} \quad && v = b^T Y \\ &\text{Subject to} \quad && A^T Y \geq c \\ & && Y \text{ unrestricted.} \end{aligned} \tag{15.15}$$

(c) General form

Consider the linear programming problem in general form

$$\text{Maximize } z = \sum_{j \in N} c_j x_j$$

$$\text{Subject to} \quad \sum_{j \in N} a_{ij} x_j \leq b_i, \; i \in M_1$$

$$\sum_{j\in N} a_{ij}x_j = b_i,\ i \in M_2 \tag{15.16}$$

$$x_j \geq 0,\ \ j \in N_1$$
$$x_j \ \text{unrestricted},\ j \in N_2.$$

where $M = \{1, 2,..m\}, M_1 \subset M, M_2 = M - M_1$
and $N = \{1, 2,\ldots n\}, N_1 \subset N, N_2 = N - N_1$

The problem can be rewritten as

$$\text{Maximize} \quad z = \sum_{j\in N_1} c_j x_j + \sum_{j\in N_2} c_j x_j^1 - \sum_{j\in N_2} c_j x_j^2$$

$$\text{Subject to} \quad \sum_{j\in N_1} a_{ij}x_j + \sum_{j\in N_2} a_{ij}x_j^1 - \sum_{j\in N_2} a_{ij}x_j^2 \leq b_i, \quad i \in M_1$$

$$\sum_{j\in N_1} a_{ij}x_j + \sum_{j\in N_2} a_{ij}x_j^1 - \sum_{j\in N_2} a_{ij}x_j^2 \leq b_i, \quad i \in M_2$$

$$-\sum_{j\in N_1} a_{ij}x_j - \sum_{j\in N_2} a_{ij}x_j^1 + \sum_{j\in N_2} a_{ij}x_j^2 \leq -b_i, \quad i \in M_2$$

$$x_j \geq 0, j \in N_1,\ x_j \geq 0, j \in N_2,\ \ x_j^2 \geq 0,\ j \in N_2$$

which is now in the cannonical form and its dual therefore is given by

$$\text{Minimize} \quad v = \sum_{i\in M_1} b_i y_i + \sum_{i\in M_2} b_i y_i^1 - \sum_{i\in M_2} b_i y_i^2$$

$$\text{Subject to} \quad \sum_{i\in M_1} a_{ij}y_i + \sum_{i\in M_2} a_{ij}y_i^1 - \sum_{i\in M_2} a_{ij}y_i^2 \geq c_j, \quad j \in N_1$$

$$\sum_{i\in M_1} a_{ij}y_i + \sum_{i\in M_2} a_{ij}y_i^1 - \sum_{i\in M_2} a_{ij}y_i^2 \geq c_j, \quad j \in N_2$$

$$-\sum_{i\in M_1} a_{ij}y_i - \sum_{i\in M_2} a_{ij}y_i^1 + \sum_{i\in M_2} a_{ij}y_i^2 \geq -c_j, \quad j \in N_2$$

$$y_i \geq 0,\ \ i \in M_1,\ \ y_i^1 \geq 0,\ \ i \in M_2,\ \ y_i^2 \geq 0,\ \ i \in M_2$$

which can be expressed as,

$$\text{Minimize} \quad v = \sum_{i\in M} b_i y_i$$

$$\text{Subject to} \quad \sum_{i\in M} a_{ij}y_i \geq c_j,\ \ j \in N_1 \tag{15.17}$$

$$\sum_{i\in M} a_{ij}y_i \geq c_j,\ \ j \in N_2$$

$$y_i \geq 0, i \in M_1,\ y_i \ \text{unrestricted for}\ i \in M_2$$

$$M_1 + M_2 = M = \{1,2..m\}, N_1 + N_2 = N\{1,3..n\}$$

The correspondence between a pair of dual programs may now be summarized as below:

Primal	Dual
Objective function: (Max $z = \Sigma_j c_j x_j$)	Constant terms c_j.
Constant terms: b_i	Objective function: Min $v = \Sigma_i b_i y_i$
Coefficient matrix: $A = (a_{ij})$	Transposed coefficient matrix : $A^T = (a_{ji})$
Constraints:	Variables:
i th inequality: $\Sigma_j a_{ij} x_j \le b_i$	i th variable $y_i \ge 0$
i th equaliity: $\Sigma_j a_{ij} x_j = b_i$	i th variable y_i unrestricted in sign
Variables:	Constraints:
(a) $x_j \ge 0$	(a) j the inequality: $\Sigma_i a_{ij} y_j \ge c_j$
(b) x_j unrestricted in sign	(b) j th equality: $\Sigma_i a_{ij} y_j = c_j$

15.4. Other Important Results

The following theorem can be interpreted as establishing a necessary and sufficient condition for optimality of feasible solutions to the pair of symmetric dual problems.

Theorem 15.6: (Weak theorem of Complementary Slackness)

Let X_0 and Y_0 be feasible solutions to the symmetric dual problems Max $\{c^T X \mid AX \le b, X \ge 0\}$ and Min $\{b^T Y \mid A^T Y \ge c, Y \ge 0\}$ respectively. Then a necessary and sufficient condition for X_0 and Y_0 to be optimal solution is that

$$Y^T_0 (b - AX_0) = 0 \qquad (15.18)$$

and $$X^T_0 (A^T Y_0 - c) = 0 \qquad (15.19)$$

Proof: Since X_0 and Y_0 are feasible solutions to the primal and the dual problems respectively, we have,

$$\alpha = Y^T_0 (b - AX_0) \ge 0 \qquad (15.20)$$

$$\beta = X^T_0 (A^T Y_0 - c) \ge 0 \qquad (15.21)$$

and $$\alpha + \beta = b^T Y_0 - c^T X_0 \ge 0 \qquad (15.22)$$

If X_0 and Y_0 are optimal solutions, then by the duality theorem, we must have

$$c^T X_0 = b^T Y_0 \qquad (15.23)$$

which means that $\alpha + \beta = 0$ and since $\alpha \ge 0$, $\beta \ge 0$, we must have $\alpha = 0$ and $\beta = 0$ and therefore (15.18) and (15.19) are true.

Now, let the conditions (15.18), (15.19) hold true. This means

$$\alpha = 0, \qquad \beta = 0$$

and therefore $\quad \alpha + \beta = 0$

i.e. $\quad b^T Y_0 - c^T X_0 = 0$

or $$c^T X_0 = b^T Y_0 \qquad (15.24)$$

Hence, by corollary 15.1, X_0 and Y_0 are optimal solutions to the primal and the dual problems respectively.

Corollary 15.3: For optimal solutions of the primal and dual systems, (i) whenever the ith variable is strictly positive in either system, the ith relation of its dual is an equality (ii) if the ith relation of either system is a strict inequality then the ith variable of its dual vanishes.

Proof: Consider the relation (15.18). Since each term in the summation

$$Y_0^T(b - AX_0) = \sum_{i=1}^{m} y_i^0 \left(b_i - \sum_{j}^{n} a_{ij} x_j^0 \right) = 0$$

is nonnegative, it follows that

$$y_i^0 \left(b_i - \sum_{j}^{n} a_{ij} x_j^0 \right) = 0 \quad \text{for } i = 1, 2, \ldots m. \tag{15.25}$$

This means that for each i = 1,2,.....m

$$y_i^0 > 0 \quad \text{implies} \quad \sum_{j=1}^{n} a_{ij}^0 x_j^0 = b_i \tag{15.26}$$

and

$$\sum_{j=1}^{n} a_{ij} x_j^0 < b_i \quad \text{implies} \quad y_i^0 = 0 \tag{15.27}$$

Similarly, from the relation (15.19), we find that

$$x_j^0 \left(\sum_{i=1}^{m} a_{ij} y_i^0 - c_j \right) = 0 \quad \text{for } j = 1, 2, \ldots n. \tag{15.28}$$

and hence for each $j = 1, 2, \ldots m.$

$$x_j^0 > 0 \quad \text{implies} \quad \sum_{i=1}^{m} a_{ij} y_i^0 = c_j \tag{15.29}$$

and

$$\sum_{i=1}^{m} a_{ij} y_i^0 > c_j \quad \text{implies} \quad x_j^0 = 0 \tag{15.30}$$

Thus the relations (15.26) and (15.29) prove case (i) and the relations (15.27) and (15.30) prove case (ii)

The possibility of the case, where both the terms in the product (15.25) or (15.28) are zero at the same time for a pair of optimal dual solutions cannot be ruled out. The following theorem will however, show that this eventuality cannot hold simultaneously for all pairs of optimal dual solutions.

Theorem 15.7: (Strong Theorem of Complementary Slackness).

Let X, Y be feasible solutions to the symmetric dual programs Max $\{c^TX \mid AX \leq b, X \geq 0\}$ and Min $\{b^TY \mid A^TY \geq c, Y \geq 0\}$ respectively. Then there exists at least one pair of optimal solution X_0, Y_0 satisfying the relations.

$$(b - AX_0) + Y_0 > 0 \tag{15.31}$$
$$(A^TY_0 - c) + X_0 > 0$$

First we prove the following lemma.

Lemma 15.1: The system of linear homogenous inequalities.

$$-AX + tb \geq 0 \tag{15.32}$$
$$A^TY - tc \geq 0$$
$$c^TX - b^TY \geq 0$$
$$X \geq 0,\ Y \geq 0, \text{ and } t \geq 0 \text{ is a single element.}$$

possesses at least one solution $\left(\bar{X}, \bar{Y}, \bar{t}\right)$ such that

$$-A\bar{X} + \bar{t}b + \bar{Y} > 0$$
$$A^T\bar{Y} - \bar{t}c + \bar{X} > 0 \tag{15.33}$$
$$c^T\bar{X} - b^T\bar{Y} + \bar{t} > 0$$
$$\bar{X}, \bar{Y}, \bar{t} \geq 0$$

Proof: The system of linear inequalities (15.32) can be expressed as

$$A_1X_1 \geq 0 \tag{15.34}$$
$$X_1 \geq 0$$

where

$$A_1 = \begin{bmatrix} 0 & A^T & -c \\ -A & 0 & b \\ c^T & -b^T & 0 \end{bmatrix}$$

is a skew – systematic matrix and

$$X_1^T = \left[X^TY^Tt\right]$$

By Theorem 7.16, the system possesses atleast one solution $\bar{X}_1 \geq 0$, such that

$$A_1, \bar{X}_1 + \bar{X}_1 > 0$$

Thus, the system of inequalities (15.32) must possess at least one solution $\left(\bar{X}, \bar{Y}, \bar{t}\right)$ such that

$$\begin{bmatrix} 0 & A^T & -c \\ -A & 0 & b \\ c^T & -b^T & 0 \end{bmatrix} \begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{t} \end{bmatrix} + \begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{t} \end{bmatrix} > 0 \tag{15.35}$$

$$\bar{X}, \bar{Y}, \bar{t} \geq 0.$$

and hence the result.

Proof of theorem 15.7. Let X and Y be feasible solutions of the primal and the dual problems respectively and thus

$$AX \geq b, X \geq 0$$
$$A^TY \geq c, Y \geq 0$$

By lemma 15.1, there exists a solution $(\bar{X}, \bar{Y}, \bar{t})$ to (15.32) satisfying (15.33). It can be easily seen that $\bar{t} > 0$. If $\bar{t} = 0$ then form (15.32) we have

$$A\bar{X} \leq 0, \bar{X} \geq 0$$

$$A^T\bar{Y} \geq 0, \bar{Y} \geq 0.$$

we then have $\quad c^T\bar{X} \leq T^TA\bar{X} \leq 0,$

and $\quad b^T\bar{Y} \geq X^TA^T\bar{Y} \geq 0.$

Hence $c^T\,\bar{X} - b^T\,\bar{Y} \leq 0$, which contradicts the last strict inequality in (15.33).

Therefore $\bar{t}$ must be positive.

Then $X_0 = \bar{X}/\bar{t}, Y_0 = \bar{Y}/\bar{t}, t_0 = 1$ is a solution of (15.32) and thus (X_0, Y_0) are feasible solutions of the primal and the dual problems respectively with $c^TX_0 \geq b^TY_0$. From corollary 15.1, it follows that $c^TX_0 = b^TY_0$, and hence (X_0, Y_0) constitute optimal solutions of the dual programs. The relations (15.33) show that (X_0, Y_0) satisfy the relations (15.31).

Thus the theorem states that there exists at least one pair of optimal solutions to the dual programs, for which if one of the two terms of the sum in (15.31), is zero, the other is strictly positive.

For example, if in (15.31)

$$b_i - \sum_{j=1}^{n} a_{ij}x_{0j} = 0, \quad \text{then } y_{0i} > 0$$

and if $\quad y_{i0} = 0, \quad \text{then } \sum_{j=1}^{n} a_{ij}x_{0j} < b_i$

Similarly, if

$$\sum_{i=1}^{m} a_{ij}y_{0i} - c_j = 0, \quad \text{then } x_{0j} > 0$$

and if $\quad x_{0j} = 0, \quad \text{then } \sum_{i=1}^{m} a_{ij}y_{0i} > c_j$

15.5. Lagrange Multipliers and Duality

In the theory differential calculus, Lagrange's method of minimizing a function $z = c^TX$ subject to the constraints $AX = b$, consists of introducing Lagrange multipliers U and look for the minimum of the unconstrained function

$$c^TX + U^T (b - AX)$$

called the Lagrangian function associated with the problem.

One may therefore feel that the method of Lagrange multipliers can perhaps be applied to solve any linear programming problem but this procedure does not work because of the additional nonnegativity constraints $X \geq 0$. We can however consider the Kuhn–Tucker theorem [291] which generalizes Lagrange's classical multiplier method to the determination of a solution of an optimization problem under inequality constraints and nonnegative variables. It has been shown that if the problem is

$$\begin{array}{ll} \text{Maximize} & f(x) \\ \text{Subject to} & g_i(X) \geq 0,\ i = 1, 2, \ldots m \\ & X \geq 0 \end{array}$$

where f(X) and all the $g_i(X)$ are differentiable concave functions then under certain conditions, a solution of the saddle value problem corresponding to the lagrangian function of the problem, constitutes optimal solutions to the problem and its dual. The Lagrange multipliers are the dual variables.

Since linear programming problems satisfy the Kuhn–Tucker conditions, this theory is applicable to linear programming also.

Consider the problem

$$\begin{array}{ll} \text{Maximize} & c^TX \\ \text{Subject to} & AX \leq b \\ & X \geq 0 \end{array} \tag{15.36}$$

The Lagrangian function associated with this problem is given by.

$$\phi\ (X,U) = c^TX + U^T (b - AX), \text{ for } X \geq 0,\ U \geq 0 \tag{15.37}$$

Definition: Saddle value problem: The saddle value problem corresponding to the Lagrangian function ϕ (X,U) is to find vectors $X_0 \geq 0$, $U_0 \geq 0$ which satisfy

$$\phi\ (X, U_0) \leq \phi\ (X_0, U_0) \leq \phi\ (X_0, U) \quad \text{for all } X \geq 0,\ U \geq 0 \tag{15.38}$$

(X_0, U_0) is called the saddle point of ϕ (X, U)

Theorem 15.8: For a pair of dual problems, a necessary and sufficient condition for the two vectors $X_0 \geq 0$ and $U_0 \geq 0$ to constitute dual optimal solutions is that (X_0, U_0) be a saddle point of the Lagrangian function ϕ (X, U). The common optimal value of the problem is then $\phi\ (X_0, U_0)$.

Proof:

Let X_0, U_0 be optimal solutions of the pair of dual problems respectively,

$$\begin{array}{llll} \text{Maximize} & c^TX & \text{Minimize} & b^TU \\ \text{Subject} & AX \le b \text{ and} & \text{Sunject to} & A^TU \ge c \\ & X \ge 0 & & U \ge 0 \end{array} \tag{15.39}$$

and then the Lagrangian function is given by

$$\phi(X,U) = c^TX + U^T(b - AX) \tag{15.40}$$

Now, $\phi(X_0, U_0) = c^TX_0 + U_0^T(b - AX_0)$

$= c^TX_0$ by theorem 15.6

$= b^TU_0 + X_0^T(c - A^TU_0)$

$= b^TU_0$ by theorem 15.6

Thus $\phi(X_0, U_0) = c^TX_0 = b^TU_0$

Moreover for all X, $U \ge 0$.

$$\phi(X, U_0) = c^TX + U_0^T(b - AX) = b^TU_0 + X^T(c - A^TU_0) \le b^TU_0 = \phi(X_0, U_0)$$

and $\phi(X_0, U) = c^TX_0 + U^T(b - AX_0) \ge c^TX_0 = \phi(X_0, U_0)$

Hence $\phi(X, U_0) \le \phi(X_0, U_0) \le \phi(X_0, U)$ for all X, $U \ge 0$.

Conversely, let there exist X_0, $U_0 \ge 0$ which satisfy (15.38). From the first part of (15.38) we have

$$\phi(X, U_0) \le \phi(X_0, U_0), X \ge 0$$

or $[c^TX + U_0^T(b - AX)] - [c^TX_0 + U_0^T(b - AX_0)] \le 0.$

or $(c^T - U_0^TA)(X - X_0) \le 0$

Taking successively $X = X_0 + e_j$, $j = 1, 2, \ldots n$

where e_j is the j the unit vector,

we have $c_j - a_j^T U_0 \le 0, j = 1, 2, \ldots n$

which implies that U_0 is a feasible solution to the dual problem.

Similarly, considering the second part of (15.38), it can be shown that X_0 is feasible to the primal problem.

Further, by putting X = 0, U = 0 in (15.38) we find

$$c^TX_0 \ge b^TU_0$$

By theorem 15.1, then $c^TX_0 = b^TU_0$ and hence X_0 and U_0 are optimal solutions to the primal and dual problems respectively and their optimal values are equal.

15.6. Duality in the Simplex Method

An important aspect of the simplex method is that from the optimal simplex tableau for a linear programming problem, an optimal solution to the dual can readily be obtained. Since the dual of dual is the primal, it is convenient to solve the dual problem to get a solution to the primal (original) whenever solution of the dual is easier than that of the primal.

15.6.1. Optimal solution of the dual

Consider the problem,

$$\text{Maximize} \quad z = c^T X \tag{15.41}$$
$$\text{Subject to} \quad AX \le b$$
$$X \ge 0$$

To initiate the simplex algorithm, we introduce slack vector $X_s \ge 0$ to convert the inequality constraints into equations and we have the problem

$$\text{Maximize} \quad z = c^T X \tag{15.42}$$
$$\text{Subject to} \quad AX + IX_s = b$$
$$X, X_s \ge 0$$

Let $X_0 = [^{X}{}_{0B}]$ be an optimal solution of the problem where B is the optimal basis and the optimal value of the objective function is Max $z = c_B{}^T X_B$, where c_B is the price vector associated with X_B.

According to the optimality criterion, we must have,

$$z_j - c_j \ge 0 \text{ for all } j$$

i.e. $$c_B^T B^{-1} a_j - c_j \ge 0 \quad \text{for} \quad j = 1, 2, \ldots\ldots n \tag{15.43}$$

and $$c_B^T B^{-1} e_i \ge 0 \quad \text{for} \quad i = 1, 2, \ldots\ldots m \tag{15.44}$$

Hence $$Y_0 = \left(c_B^T B^{-1}\right)^T \tag{15.45}$$

is feasible to the dual problem

$$\text{Maximize} \quad b^T Y \tag{15.46}$$
$$\text{Subject to} \quad A^T Y \ge c$$
$$Y \ge 0$$

and the value of the objective function is

$$b^T Y_0 = Y_0^T b = c_B^T B^{-1} b = c_B^T X_B = \text{Max } z$$

By corollary 15.1, then Y_0 is an optimal solution of the dual problem. Note that $Y_0^T = c_B^T B^{-1}$ is found in the $z_j - c_j$ row of the optimal simplex tableau under the columns corresponding to the slack variables.

Thus, to find an optimal solution of the problem (15.46) we obtain an optimal solution of its dual (15.41) and from the optimal simplex tableau, an optimal solution of the original problem (15.46) is easily found.

Now, suppose that the problem to be solved is

$$\text{Maximize} \quad b^T Y$$
$$\text{Subject to} \quad A^T Y \le c \tag{15.47}$$
$$Y \text{ unrestricted}$$

Its dual is

$$\begin{aligned} \text{Minimize} \quad & c^T X \\ \text{Subject to} \quad & AX = b \\ & X \geq 0 \end{aligned} \tag{15.48}$$

To solve the dual problem, we consider the equivalent problem

$$\begin{aligned} \text{Minimize} \quad & c^T X + M_a^T X^a \\ & AX + X^a = b \\ & X, X^a \geq 0 \end{aligned} \tag{15.49}$$

where $X^a = \left(x_1^a, x_2^a, \ldots x_m^a\right)^T$ is an artificial vector and $M_a = (M, M, \ldots M)^T$ is an m-vector with each component M, a large positive number.

For an optimal solution of the problem, we must have

$$\left(c_B^T B^{-1}\right) a_j - c_j \leq 0, \quad j = 1, 2, \ldots n \tag{15.50}$$

and
$$\left(c_B^T B^{-1}\right) e_i - M \leq 0, \quad i = 1, 2, \ldots m \tag{15.51}$$

The above inequalities imply that $Y_0^T = c_B^T B^{-1}$ is a feasible solution to the problem. Moreover, it is an optimal solution.

15.7. Example

Consider the problem

$$\begin{aligned} \text{Maximize} \quad & 2y_1 + 6y_2 + 7y_3 \\ \text{Subject to} \quad & y_1 + 2y_2 + y_3 \leq 2 \\ & 2y_1 - 3y_2 + y_3 \leq -1 \\ & -y_1 + y_2 + y_3 \leq -1 \\ & y_1, y_2, y_3, y_4 \text{ unrestricted.} \end{aligned} \tag{15.32}$$

The dual problem is

$$\begin{aligned} \text{Minimize} \quad & z = 2x_1 + x_2 - x_3 + Mx_4 + Mx_5 + Mx_5 \\ \text{Subject to} \quad & x_1 - 2x_2 - x_3 + x_4 = 2 \\ & 2x_1 + 3x_2 + x_3 + x_5 = 6 \\ & x_1 + x_2 + x_3 + x_6 = 7 \\ & x_j \geq 0, j = 1, 2, \ldots 6. \end{aligned} \tag{15.53}$$

where x_4, x_5, x_6 are artificial variables and M is a large positive number.

Tableau 1

c_B	X_B	x_1	x_2	x_3	x_4	x_5	x_6	b
M	x_4	1*	2	–1	1	0	0	2 →
M	x_5	2	–3	1	0	1	0	6
M	x_6	1	1	1	0	0	1	7
	c_j	2	–1	–1	M	M	M	
	$z_j - c_j$	4M–2	1	M+1	0	0	0	
		↑						

Tableau 2

		x_1	x_2	x_3	x_4	x_5	x_6	
2	x_1	1	2	–1	1	0	0	2
M	x_5	0	–7	3*	–2	1	0	2 →
M	x_6	0	–1	2	–1	0	1	5
	$z_j - c_j$	0	–8M+5	5M–1	–4M+2	0	0	
				↑				

Tableau 3

c_B	X_B	x_1	x_2	x_3	x_4	x_5	x_6	b
2	x_1	1	–1/3	0	1/3	1/3	0	8/3
–1	x_3	0	–7/3	1	–2/3	1/3	0	2/3
M	x_6	0	11/3*	0	1/3	–2/3	1	11/3 →
	$z_j - c_j$	0	11/3M+8/3	0	–2/3M+4/3	–5/3M+1/3	0	
			↑					

Tableau 4

		x_1	x_2	x_3	x_4	x_5	x_6	
2	x_1	1	0	0	12/33	3/11	1/11	3
–1	x_3	0	0	1	–15/33	–1/11	7/11	3
–1	x_2	0	1	0	1/11	–2/11	3/11	1
	$z_j - c_j$	0	0	0	36/33–M	9/11–M	–8/11–M	

Since all $z_j - c_j \leq 0$, in Tableau 4, we have an optimal solution of the dual problem.

$x_1 = 3, x_2 = 1, x_3 = 3$ and the optimal value of $z = 2$,

From Tableau 4, we then find the optimal solution of the original problem as

$$y_1 = 36/33,\ y_2 = 9/11,\ y_3 = -8/11$$

and the optimal value is 2.

15.8. Applications

We shall now show how duality relations in linear programming can be used in proving some important results.

Theorem 15.9 (Farkas' Lemma)

Let A be an $m \times n$ matrix and p and X are n– vectors. Then $p^TX \geq 0$ holds for all X satisfying $AX \geq 0$ if and only if there exists an m– vector U such that

$$A^TU = p,\ U \geq 0$$

Proof: Suppose that $p^TX \geq 0$ holds for all X satisfying $AX \geq 0$

Consider the linear programming problem

$$\begin{aligned} &\text{Minimize} && p^TX \\ &\text{Subject to} && AX \geq 0 \end{aligned} \tag{15.54}$$

The problem (15.54) is feasible and since p^TX is always nonnegative on the constraints set, by letting $X = 0$ we find that the optimal value of the objective function is zero.

Now, the dual to the problem (15.54) is given by

$$\begin{aligned} &\text{Maximize} && 0 \\ &\text{Subject to} && A^TU = p \\ &&& U \geq 0 \end{aligned} \tag{15.55}$$

and by duality theorem, the problem (15.55) has a solution.

Thus there exists $U \geq 0$, such that $A^TU = p$.

Conversely, suppose that there exists an m-vector $U \geq 0$ such that $A^TU = p$. The problem (15.55) then has an optimal solution with the optimal value equal to zero since the U satisfying the constraints of (15.55) will give the value zero to the objective function. Hence, by duality theorem, there exists an X satisfying $AX \geq 0$ and Min $p^TX = 0$, which implies that $p^TX \geq 0$ for all X satisfying $AX \geq 0$.

Solution of the Linear Inequalities

Consider the problem of solving a system of linear inequalities

$$AX \leq b \tag{15.56}$$

By theorem 7.10, (15.56) has no solution if and only if the system of equations

$$A^TY = 0,\ b^TY = -1,\ Y \geq 0 \tag{15.57}$$

has a solution.

Consider the linear programming problem

$$\begin{aligned} &\text{Minimize} && \theta \\ &\text{Subject to} && A^T Y = 0 \\ &&& b^T Y - \theta = -1 \qquad (15.58) \\ &&& Y, \theta \geq 0 \end{aligned}$$

If the minimum value of θ is zero, it will mean that the problem (15.57) has a solution and therefore the system of inequalities (15.56) does not have a solution. If on the other hand, the minimum value of θ is positive then an optimal solution of the dual to (15.58) provides us with a solution of (15.56).

The dual to (15.58) is given by the problem

$$\begin{aligned} &\text{Maximize} && -\mu \\ &\text{Subject to} && AX + b\mu \leq 0 \qquad (15.59) \\ &&& -\mu \leq 1. \end{aligned}$$

where μ is a single element.

and $-\mu_0 = \max(-\mu) = \min \theta > 0$

Thus $-\mu_0 > 0$ and hence $X/(-\mu_0)$ solves the original system (15.56).

Thus to find a solution of (15.56), we first solve the linear programming problem (15.58) by the simplex method and from the optimal tableau, we note the solution of (15.58) and its dual. If the optimal value of (15.58) is positive, a solution of (15.56) can be obtained from the optimal solution of the dual problem.

Examples of duality also arise in Leontief input–output system [303] (which in fact was one of the inspirations for the development of linear programming), matrix games, mathematics and in many other fields.

15.9. Economic Interpretation of Duality

Let us now consider the economic interpretation of a pair of dual problems in linear programming. The basic ingredients of all economic problems are inputs, outputs and profit. In this context we consider the activity analysis problem discussed in chapter 1. The primal problem is

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i = 1, 2, \ldots m.$$

$$x_j \geq 0, \quad j = 1, 2, \ldots n$$

$$\text{Maximize} \quad \sum_{j=1}^{n} c_j x_j$$

where a_{ij} represents the quantity of resource i required to produce one unit of product j, b_i represents the availability of the resource i and c_j represents the profit from one unit of product j. x_j is the level of production of the jth product.

The primal problem is thus

Subject to

$$\sum_{j=1}^{n} \begin{pmatrix} \text{Production rate of j} \\ \text{at unit level from} \\ \text{resource i} \end{pmatrix} \times \begin{pmatrix} \text{level of} \\ \text{product j} \end{pmatrix} \leq \begin{pmatrix} \text{availability} \\ \text{of resource i} \end{pmatrix} \qquad i = 1, 2, ...m.$$

$$(\text{level of product j}) \geq 0, \qquad j = 1,2,\ldots n$$

Maximize (overall profit)

$$= \sum_{j=1}^{n} \begin{pmatrix} \text{Price per unit of} \\ \text{product j} \end{pmatrix} \times (\text{level of product j})$$

The corresponding dual problem is

Subject to

$$\sum_{i=1}^{n} \begin{pmatrix} \text{Production rate of j} \\ \text{at unit level from} \\ \text{resource i} \end{pmatrix} \times y_i \geq \begin{pmatrix} \text{Price per unit} \\ \text{of product j} \end{pmatrix}, \qquad j = 1,2,\ldots n$$

$$y_i \geq 0, \qquad i = 1,2,\ldots m$$

$$\text{Minimize (cost)} = \sum_{i=1}^{m} (\text{availability of resource i}) \times y_i$$

To interpret the dual variables y_i, we note that in a situation of equilibrium, the laws of economics for society require no profits. Hence, y_i should represent unit price of resource i. It is common to refer to y_i as the shadow price.

Thus the dual problem is

Subject to

$$\sum_{i=1}^{n} \begin{pmatrix} \text{Production rate of j} \\ \text{at unit level from} \\ \text{resource i} \end{pmatrix} \times \begin{pmatrix} \text{Unit cost of} \\ \text{resource i} \end{pmatrix} \geq \begin{pmatrix} \text{Price per unit} \\ \text{of product j} \end{pmatrix}, \; j = 1,2,\ldots n$$

$$(\text{Unit cost of resource i}) \geq 0, \qquad i = 1,2,\ldots m$$

$$\text{Minimize (cost)} = \sum_{i=1}^{n} \begin{pmatrix} \text{Availability of} \\ \text{resource i} \end{pmatrix} \times \begin{pmatrix} \text{Unit cost of} \\ \text{resource i} \end{pmatrix}$$

Thus, the two dual problems represents two opposing interests. The former considers the question of determining the level of production so that the overall profit is maximum and the latter wants to fix the prices of the resources such that the total cost of resources used for production is minimum. Note that the solution of one naturally solves the other problem.

15.9. Exercises

1. Find the dual of the problem

$$\begin{array}{ll} \text{Minimize} & z = 4x_1 + 2x_2 + 3x_3 \\ \text{Subject to} & 6x_1 + x_2 + 4x_3 \le 5 \\ & 7x_1 + 3x_2 + x_3 = 3 \\ & 5x_1 + 2x_2 + 3x_3 \ge 2 \\ & x_1 \text{ unrestricted, } x_2, x_3 \ge 0. \end{array}$$

2. Find optimal solution of the following problems by solving their duals.

(i) $$\begin{array}{ll} \text{Minimize} & z = x_1 + x_2 \\ \text{Subject to} & 7x_1 + 3x_2 \ge 12 \\ & x_1 + 2x_2 \ge 8 \\ & x_1, x_2 \ge 0 \end{array}$$

(ii) $$\begin{array}{ll} \text{Minimize} & z = x_1 - x_2 \\ \text{Subject to} & x_1 + x_2 \le -1 \\ & 2x_1 + x_2 \ge 2 \\ & x_1, x_2 \ge 0 \end{array}$$

(iii) $$\begin{array}{ll} \text{Maximize} & z = 4x_1 + 2x_2 \\ \text{Subject to} & x_1 - x_2 \ge 2 \\ & x_1 + x_2 \ge 3 \\ & x_1, x_2 \ge 0 \end{array}$$

(iv) $$\begin{array}{ll} \text{Minimize} & z = 6x_1 + 4x_2 + 3x_3 \\ \text{Subject to} & x_1 + x_2 \ge 2 \\ & x_1 + x_3 \ge 5 \\ & x_1, x_2, x_3 \ge 0. \end{array}$$

(v) $$\begin{array}{ll} \text{Minimize} & z = 45x_1 + 36x_2 + 60x_3 \\ & x_1 + x_2 + 4x_3 \ge 50 \\ & 2x_1 + x_2 + 2x_3 \ge 40 \\ & 5x_1 + 2x_2 + x_3 \ge 25 \\ & x_1, x_2, x_3 \ge 0 \end{array}$$

3. Obtain dual to the following linear program

$$\begin{array}{ll} \text{Maximize} & z = c^T X \\ \text{Subject to} & A_1 X \le b_1, \\ & A_2 X = b_2, \\ & A_3 X \ge b_3 \\ & X \ge 0 \end{array}$$

where A_1 is $m_1 \times n$, A_2 is $m_2 \times n$ and A_3 is $m_3 \times n$ matrices,

4. Solve the following system of inequalities by the simplex method.

$$\begin{array}{l} 4x_1 - 7x_2 + 5x_3 \le 1 \\ -2x_1 + 4x_2 - 3x_3 \le 3 \\ 2x_1 + 2x_2 - 3x_3 \ge 7 \\ -2x_1 + x_2 + x_3 \ge 4 \end{array}$$

5. Show the problem

   Maximize c^TX

   Subject to $AX = b$

   has a solution if and only if c is a linear combination of the columns of A.
6. Construct an example different from that given in section 15.2 to show that both the primal and dual problems may have no solution.
7. Consider the problem: Minimize c^TX subject to $AX = b$, $X \geq 0$ where A is an $n \times n$ symmetric matrix and $c = b$. Show that if there exists an $X_0 \geq 0$ such that $AX_0 = b$, then X_0 is an optimal solution of the problem.
8. Prove that the optimal dual solution is never unique if the optimal basic solution is degenerate and the optimal dual is not.

Chapter 16

Variants of the Simplex Method

16.1. Introduction

By simplex variants we mean the methods developed by varying the simplex algorithm so as to reduce the number of iterations in solving a linear programming problem. This is specially needed for problems having many constraints in order to reduce the cost of computation. It is also needed for problems involving a large number of variables, for the number iterations appears to increase with this number. This is in particular the case when the introduction of artificial vectors is required to obtain an initial basic feasible solution which involves considerable computational effort.

We present in this chapter two variants of the simplex method which rely on duality relations: *the dual simplex method* and *the primal-dual algorithm.*

16.2. The Dual Simplex Method

As we know the simplex algorithm needs to have a basic feasible solution of the problem to start with. It then calculates a sequence of basic feasible solutions which eventually lead to an optimal solution of the problem. Thus, having an initial basic feasible solution is an a priori condition for applying the simplex method. The search for an initial basic feasible solution involves substantial computational effort when the introduction of artificial variables is required because in that case, we first need to replace the artificial variables one by one.

To reduce this prolonged computations, we start with a basic solution to the original (called the primal) problem which is not feasible but satisfies the optimality criteria. By duality relations a basic feasible solution to the dual problem is easily obtained and such a primal basic solution is called a dual feasible basic solution. Now, starting with a dual feasible but primal infeasible basic solution, the simplex algorithm is applied to the dual problem to obtain a sequence of primal infeasible solutions satisfying the optimality criteria which ultimately lead to an optimal solution. This is precisely what Lemke's [300] dual simplex method does.

Consider the linear programming problem

$$\begin{aligned} \text{Minimize} \quad & z = c^T X \\ \text{Subject to} \quad & AX = b. \\ & X \geq 0 \end{aligned} \tag{16.1}$$

and its dual

$$\begin{aligned} \text{Maximize} \quad & v = bTW \\ \text{Subject to} \quad & A^T W \leq c \end{aligned} \tag{16.2}$$

where A is an m x n matrix.

Suppose that we have obtained from A, a basis matrix $B = (a_1, a_2, ... a_m)$, a_j being the jth column of A such that $X_B = B^{-1}b$ is a basic solution to the primal problem which satisfies the optimality criterion.

Hence $z_j - c_j \leq 0$, for all j

or $\quad A^T W \leq c$

where $W^T = c_B^T B^{-1}$

Therefore W is a feasible solution to the dual problem (if section (15.6.1) and X_B is then called a dual feasible basic solution. If X_B is feasible, i.e. $X_B \geq 0$, it is of course, an optimal solution to the primal.

If W is not an optimal solution to the dual problem, then by the duality theorem X_B is not feasible to the primal and at least one element of $X_B = B^{-1}b$ is negative. We shall now show that if $X_{B_i} < 0$ for at least one i (i = 1, 2..m), it is possible to find another feasible solution to the dual which is better than W.

Let β_i be the ith row of B^{-1} and consider

$$\hat{W}^T = W^T - \theta \beta_i \tag{16.4}$$

where θ is a scalar.

We then have

$$\begin{aligned} \hat{W}^T b &= W^T b - \theta \beta_i b. \\ &= W^T b - \theta X_{B_i} \end{aligned} \tag{16.5}$$

so that $$\hat{W}^T b > W^T b \text{ for all } \theta > 0 \text{ since } X_{B_i} < 0 \tag{16.6}$$

If now $\hat{W}$ can be shown to satisfy the constraints of the dual, we have a new solution $\hat{W}$ to the dual which is better than W.

Note that $\quad \beta_i a_j = \alpha_{ij}$, $\quad$ for a_j not in the basis

$$\beta_i a_r = \delta_{ir}, \quad i, r = 1, 2,..m, \; a_i, a_r \text{ are in the basis} \tag{16.7}$$

We have, $\quad \hat{W}^T a_j = (W^T - \theta \beta_i) a_j$,

$\quad = W^T a_j - \theta \beta_i a_j$, $\quad$ for all j

Thus, $\quad \hat{W}^T a_j = W^T a_j - \theta \alpha_{ij}$

$$= z_j - \theta\, \alpha_{ij} \text{ for j nonbasic}^1 \tag{16.8}$$

$$\hat{W}^T a_r = z_r = c_r, \text{ for r, i basic, } r \neq i \tag{16.9}$$

and
$$\hat{W}^T a_i = z_i - \theta = c_i - \theta,\ r = i. \tag{16.10}$$

If $\alpha_{ij} \geq 0$ for all j nonbasic it follows from (16.8), (16.9), (16.10) and (16.3) that $\hat{W}$ is a feasible solution to the dual problem for every $\theta > 0$ and from (16.5) we note that the dual problem is unbounded (when $\theta \to +\infty$) which implies that the primal problem has no feasible solution.

If $\alpha_{ij} < 0$, for at least one j nonbasic we note from (16.8), (16.9) and (16.10) that $\hat{W}$ remains a feasible solution to the dual problem if and only if

$$z_j - \theta\alpha_{ij} \leq c_j, \qquad \text{for all j nonbasic}$$

or
$$\theta \leq \operatorname*{Min}_j \left[\frac{z_j - c_j}{\alpha_{ij}} \middle| \alpha_{ij} < 0 \right]$$

$$= \frac{z_k - c_k}{\alpha_{ik}} \tag{16.11}$$

To have a maximum improvement in the solution to the dual, θ must be as large as possible and therefore θ is taken to be

$$\theta = \frac{z_k - c_k}{\alpha_{ik}} \tag{16.12}$$

and then from (16.8) we have
$$\hat{W}^T a_k = z_k - \theta\, \alpha_{ik}. \tag{16.13}$$
$$= c_k$$

Since $\alpha_{ik} \neq 0$, a new basis matrix B' is now formed by replacing a_i by a_k and it follows from (16.9) and (16.13) that a new solution $\hat{W}$ to the dual is obtained so that

$$\hat{W}^T B' = c_{B'}^T$$

or
$$\hat{W}^T = c_{B'}^T (B')^{-1} \tag{16.14}$$

Let
$$X_{B'} = (B')^{-1} b \tag{16.15}$$

be the new basic solution to the primal.

If $X_{B'} \geq 0$, it is an optimal solution of the primal problem. If at least one element of $X_{B'}$ is negative, we apply the preceding process again and obtain a new solution $\hat{\hat{W}}$ to the dual, better than $\hat{W}$.

If it is assumed that $\theta \neq 0$, then from (16.5), it follows that the value of the dual problem is strictly increasing and no basis is therefore repeated. Since the number of bases is finite, the process will terminate in a finite number of steps with a basis that solves the dual and thus provides a feasible and hence an optimal solution to the primal problem or with an indication that the dual problem is unbounded implying that the primal problem has no feasible solution.

Consider now the case when θ computed from (16.11) is zero. It occurs when in the previous iteration of the dual simplex algorithm, there is a tie in the computation of the entry criterion and degeneracy appears in the dual problem. To handle degeneracy in the dual, Lemke [300] has shown that a method similar to the degeneracy method for the primal may be established.

16.3. The Dual Simplex Algorithm

The dual simplex method may now be summarized as follows:

Step 1. Obtain a basis B of the primal problem such that $z_j - c_j \leq 0$, for all j nonbasic[1] (see Section 16.4).

Step 2. Check the basic solution $X_B = B^{-1}b$,

If $X_B \geq 0$, it is an optimal solution of the problem.

Otherwise some of the elements of X_B are negative. Let $X_{B_i} < 0$ for at least one i in the basis[1]

Step 3. Check α_{ij} for all i for which $X_{Bi} < 0$ and all j nonbasic.

(a) If $\alpha_{ij} \geq 0$ for at least one i and every j nonbasic, the primal problem has no feasible solution.

(b) If $\alpha_{ij} < 0$ for every i for which $X_{Bi} < 0$ and for at least one j nonbasic, select r by the relation

$$X_{Br} = \min_i \left[X_{Bi} \mid X_{Bi} < 0\right], \qquad \text{Exit criterion}$$

and the vector a_r is removed from the basis, and determine k by the relationship

$$\frac{z_k - c_k}{\alpha_{rk}} = \min_j \left[\frac{z_j - c_j}{\alpha_{rj}} \middle| \alpha_{rj} < 0\right], \qquad \begin{array}{c}\text{Entry criterion}\\ \text{for all j nonbasic}\end{array}$$

so that a_k is entered in the basis.

Step 4. Obtain the new basis matrix B' by removing the vector a_r and introducing a_k in the basic set. Calculate the new values of $X_{B'}$, z'_j and α'_{ij} and repeat the algorithm from step 2.

Note that the dual simplex method is carried out with the same tableau as the primal simplex method.

1. For brevity we write j nonbasic meaning that j belonging to the set of indices of nonbasic vectors. Similarly, for i in the basis.

16.4. Initial Dual - Feasible Basic Solution

The dual simplex algorithm is used to solve a linear programming problem when a basic solution to the problem with all $z_j - c_j \leq 0$ is known. In certain cases such as in parametric and post-optimization problems a dual-feasible basic solution is readily available but in general it is not easy to find such a basic solution to the problem. However, there may be situations, where we wish to use the dual simplex algorithm when a dual-feasible basic solution is not already known. In such cases the following procedures may be followed to find initial dual-feasible basic solution.

16.4.1. A particular case

Consider the problem

$$\text{Minimize } c^T X, \text{ subject to } AX \geq b, X \geq 0 \tag{16.15}$$

where c_j are all nonnegative.

Suppose a surplus variable is added to each constraint

Then the initial basis matrix is $B = -I$ and $X_B = -b$ with the associated costs equal to zero. Further, since $c_j \geq 0$ for every j, $z_j - c_j \leq 0$. Thus X_B is a dual feasible solution and the dual simplex algorithm can now be applied.

16.4.2. Dantzig's Method [100]

Dantzig has suggested a method of getting started on the dual simplex algorithm for a linear programming problem which involves the direct use of its dual.

Consider the problem (16.1), i.e.

$$\begin{aligned} &\text{Minimize} && z = c^T X \\ &\text{Subject to} && AX = b \\ & && X \geq 0 \end{aligned} \tag{16.16}$$

where $A = (a_{ij})$ is an m x n matrix, $n > m$ and $c_j \geq 0$, for all $j = 1,2,..n$.

Its dual is

$$\begin{aligned} &\text{Maximize} && v = b^T W \\ &\text{Subject} && A^T W \leq c \end{aligned} \tag{16.17}$$

The dual problem (16.17) is written in the standard form by adding a slack variable W_j^a to the jth constraint for each $j = 1, 2,..n$. The initial dual basis is then $\overline{B}_1 = I$ and $W_j^a = c_j$ is the initial dual solution. Starting with $\overline{B}_1$, a dual basis $\overline{B}$ $(n \times n)$ is obtained which contains the m column vectors a_i^T, transposed from rows of A. Without loss of generality (after rearranging and possible renumbering of the rows), $\overline{B}$ can be expressed as

$$\begin{bmatrix} a_1^T & 0, & 0,\cdots 0 \\ a_2^T & 0, & 0,\cdots 0 \\ \vdots & & \\ a_m^T & 0, & 0,\cdots 0 \\ a_{m+1}^T & 1, & 0,\cdots 0 \\ \vdots & & \\ a_n^T & 0, & 0,\cdots 1 \end{bmatrix}$$

where the first m rows are the vectors a_j^T transposed from a column of A and of (n − m) zeros respectively. The m vectors a_j, (j = 1, 2,..m) are necessarily independent $\left(\because r(\overline{B}) = n\right)$ and hence form a primal basis B. The dual solution associated with $\overline{B}$ satisfies the first m constraints with equality and it then follows that the primal basic solution associated with B is dual feasible.

16.4.3. Lemke's Method [300]

Suppose that the given problem is

$$\begin{aligned} &\text{Minimize} && z = c^T X \\ &\text{Subject to} && AX=b \\ &&& X \geq 0 \end{aligned} \tag{16.18}$$

Lemke has suggested the following procedure to obtain a dual feasible solution of the problem.

Step 1. Find m linearly independent columns of A and let the vector $\hat{b}$ be a linear combination of these vectors with each coefficient in the linear expression positive.

Step 2. Solve the problem Minimize $z = c^T x$, subject to $AX= \hat{b}$, $X \geq 0$, by the simplex method.

Step 3. In the optimal solution, replace $\hat{b}$ by b. This will give a basic solution to the original problem with all $z_j - c_j \leq 0$, since $z_j - c_j$ does not depend on the requirement vector b.

The dual simplex algorithm can then be applied.

However, both Dantzig's method and the method suggested by Lemke require a good deal of calculations and may not therefore be advantageous over the usual artificial variable technique to solve the problem.

16.4.4. A General Method: The Artificial Constraint Technique

Consider the linear programming problem in standard form (16.1) and suppose that a basis B is known.

The constraints of the problem can be written as

$$X_B + \sum_j B^{-1} a_j x_j = B^{-1} b = \bar{b}$$

Assuming that B corresponds to the first m vectors of A, we have from (11.4)

$$x_i + \sum_{j=m+1}^{n} \alpha_{ij} x_j = \bar{b}, \qquad i = 1, 2, \ldots m.$$

where $\left(\bar{b}_i\right) = \bar{b} = \bar{X}_B$ = value of the basic variables

Multiplying this expression by c_B^T on the left and subtracting it from

$$z = c_B^T X_B + \sum_{j=m+1}^{n} c_j x_j, \quad \text{we have}$$

$$z - c_B^T \bar{b} = \sum_{j=m+1}^{n} c_j x_j = \sum_{j=m+1}^{n} c_B^T \alpha_j x_j, \quad \alpha_j = \left(\alpha_{ij}\right), \quad i = 1, 2, \ldots m$$

Setting $c_B^T \bar{b} = c_B^T \bar{X}_B = \bar{z}$ = the value of the objective function for this basic solution we have

$$z - \bar{z} = \sum_{j=m+1}^{n} \left(c_j - z_j\right) x_j, \qquad \text{since by (11.5)}, \ \sum_{i=1}^{m} c_i \alpha_{ij} = z_j.$$

The problem can thus be expressed as

$$\text{Minimize} \quad z = \bar{z} - \sum_{j=m+1}^{n} \left(z_j - c_j\right) x_j \tag{16.19}$$

$$\text{Subject to} \quad x_i + \sum_{j=m+1}^{n} \alpha_{ij} x_j = \bar{b}_i, \qquad i = 1, 2, \ldots m \tag{16.20}$$

$$x_j \geq 0 \qquad j = 1, 2, \ldots n \tag{16.21}$$

Suppose that the basic solution is not feasible and that not all $z_j - c_j \leq 0$. We then introduce the artificial constraint

$$\sum_{j=m+1}^{n} x_j \leq M \tag{16.22}$$

where M is a large positive number, larger than any finite number. Now, introducing a nonnegative slack variable x_0, the artificial constraint (16.22) is transformed into an equality, so that

$$x_0 + \sum_{j=m+1}^{n} x_j = M, \quad x_0 \geq 0 \tag{16.23}$$

Consider now, the augmented problem

$$\text{Minimize} \quad z = \bar{z} - \sum_{j=m+1}^{n} (z_j - c_j)\, x_j \tag{16.24}$$

Subject to constraints (16.20), (16.21) and (16.23).

$$\begin{aligned} \text{Clearly} \quad & x_i = \bar{b}_i, && i = 1,2,\ldots m \\ & x_0 = M \\ & x_j = 0, && j = m+1,\ldots n \end{aligned} \tag{16.25}$$

constitute a basic solution to the augmented problem (16.24)

Now, change the basis by introducing the variable x_k $(m + 1 \leq k \leq n)$ in place of x_0, where k is defined by

$$z_k - c_k = \text{Max}\,(z_j - c_j) \tag{16.26}$$

Replacing x_k by its value $M - x_0 - \sum_{j=m+1}^{n} x_j$ in (16.19) and (16.20) we get the problem

$$\begin{aligned} \text{Minimize} \quad z = {} & [\bar{z} - M(z_k - c_k)] + (z_k - c_k)\, x_0 \\ & + \sum_{\substack{j=m+1 \\ j \neq k}}^{n} \left[(z_k - c_k) - (z_j - c_j)\right] x_j \end{aligned} \tag{16.27}$$

Subject to

$$-\alpha_{ik} x_0 + x_i + \sum_{\substack{j=m+1 \\ j \neq k}}^{n} (\alpha_{ij} - \alpha_{ik}) x_j = \bar{b}_i - \alpha_{ik} M, \quad i = 1,2,\ldots m.$$

$$x_0 + x_k + \sum_{\substack{j=m+1 \\ j \neq k}}^{n} x_j = M \tag{16.28}$$

$$x_j \geq 0, \qquad j = 0, 1, 2\ldots n.$$

$x_1, x_2, \ldots x_m, x_k$ can then be taken as basic variables and from (16.26), it follows that the basic solution obtained is dual feasible. The dual simplex algorithm may therefore be applied to the augmented problem.

The application of the dual simplex algorithm leads to the termination of the process with one of the following three results:

(a) The augmented problem has no feasible solution.

In this case the original problem also does not have a feasible solution since every feasible solution $(\bar{x}_1, \bar{x}_2, \ldots \bar{x}_n)^T$ to the original problem yields a feasible solution $(\bar{x}_0, \bar{x}_1, \ldots \bar{x}_n)^T$ to the augmented problem where

$$\bar{x}_0 = M - \sum_{j=m+1}^{n} \bar{x}_j$$

(b) The augmented problem has a minimal solution and x_0 is not a basic variable in this solution.

In this case at the minimal solution, the constraint (16.22) is satisfied with equality, i.e.

$$\sum_{j=m+1}^{n} x_j = M. \tag{16.29}$$

and the values of the basic variables are functions of M. Now two cases may arise.

(i) The minimal value $\bar{z}$ of z is an explicit function of M for every value of M greater than a fixed value. M_1. If $M \to +\infty$, then $z \to -\infty$, since z cannot tend towards $+\infty$, because there is a feasible solution of the original problem which yields a finite value of z. Hence the augmented problem has an unbounded solution and since the objective functions of the augmented and the original problems are the same and further every feasible solution of the augmented problem is a feasible solution of the original problem ($x_0 = 0$), the original problem has also an unbounded solution.

(ii) The minimal value $\bar{z}$ is independent of M for every value of M greater than M_1. When M_1 varies and is larger than M_1, the hyperplane

$$\sum_{j=m+1}^{n} x_j = M.$$

is displaced parallel to itself and the minimal vertex which is lying on this hyperplane ($x_0 = 0$) moves out to an infinite edge of the polytope represented by the set of feasible solutions of the augmented problem. Since $\bar{z}$ is not a function of M, the hyperplane $c^T X = \bar{z}$ contains this edge and therefore all the points on this edge are minimal feasible solutions. In particular, there exists a minimal basic feasible solution of the original problem represented by the origin of this infinite edge and this origin is obtained by decreasing M until one of the variables which is a function of M vanishes.

(c) The augmented problem has a minimal solution and x_0 is a basic variable in this solution.

It follows from the constraints (16.23) and (16.20) that the minimal basic solution is given by

$$\begin{bmatrix} 1 & e^T \\ 0 & B \end{bmatrix} \begin{bmatrix} x_0 \\ X_B \end{bmatrix} = \begin{bmatrix} M \\ \bar{b} \end{bmatrix} \tag{16.30}$$

where e is an m-vector of ones and

$$\bar{b}^T = (\bar{b}_1, \bar{b}_2, \ldots \bar{b}_m).$$

It should be noted from (16.30) that

$$x_0 + M + (\text{a constant independent of } M) > 0$$

and all the basic variables x_B are independent of M. Since $x_0 > 0$, the constraint (16.22) is a strict inequality, i.e.

$$\sum_{j=m+1}^{n} x_j < M$$

in this minimal basic solution.

Hence the values of X_B in the minimal basic solution constitute a minimal solution to the original problem.

16.5. Example

Consider the problem

Maximize $x_1 + 2x_2 + x_3$

Subject to

$$5x_1 - 2x_2 - 3x_3 \leq -4$$
$$3x_1 + 4x_2 + 6x_3 \leq 8$$
$$x_1 + x_2 - 9x_3 \leq -3$$
$$x_1, x_2, x_3 \geq 0$$

Introducing the slack variables x_4, x_5, x_6 we express the problem in the following equivalent form.

Minimize $z = -x_1 - 2x_2 - 3x_3$

Subject to

$$5x_1 - 2x_2 - 3x_3 + x_4 = -4$$
$$3x_1 + 4x_2 + 6x_3 + x_5 = 8$$
$$x_1 + x_2 - 9x_3 + x_6 = -3$$
$$x_j \geq 0, \; j = 1,2\ldots6.$$

The basic solution $x_4 = -4$, $x_5 = 8$, $x_6 = -3$ is not feasible and since $z_1 - c_1 = 1$, $z_2 - c_2 = 2$, $z_3 - c_3 = 1$, the optimal condition is not satisfied.

We, therefore introduce the artificial constraint

$$x_0 + x_1 + x_2 + x_3 = M$$

where $x_0 \geq 0$ and M is a large positive number.

Now, since $\underset{j}{\text{Max}}\,(z_j - c_j) = Z_2 - c_2 = 2$, we replace x_0 by x_2 in the basis by substituting

$$x_2 = M - x_0 - x_1 - x_3$$

in the problem

Thus, we have the augmented problem

$$\begin{aligned} \text{Minimize} \quad & z = -2M + 2x_0 + x_1 + x_3 \\ \text{Subject to} \quad & 2x_0 + 7x_1 - x_3 + x_4 = -4 + 2M \\ & -4x_0 - x_1 + 2x_3 + x_5 = 8 - 4M \\ & -x_0 - 10x_3 + x_6 = -3 - M \\ & x_0 + x_1 + x_2 + x_3 = M \\ & x_j \geq 0, j = 0,1,\ldots 6. \end{aligned}$$

and the basic solution $x_2 = M$, $x_4 = -4 + 2M$, $x_5 = 8 - 4M$, $x_6 = -3 - M$ is dual feasible and we apply the dual simplex algorithm

Tableau 1

c_B	X_B	x_0	x_1	x_2	x_3	x_4	x_5	x_6	$\bar{X}_B$	
0	x_4	2	7	0	–1	1	0	0	–4+2M.	
0	x_5	–4*	–1	0	2	0	1	0	8–4M	→
0	x_6	–1	0	0	–10	0	0	1	–3–M.	
0	x_2	1	1	1	1	0	0	0	M	
	c_j	2	1	0	1	0	0	0		
	$z_j - c_j$	–2	–1	0	–1	0	0	0		
		↑								

Tableau 2

0	x_4	0	13/2	0	0	1	1/2	0	0	
2	x_0	1	1/4	0	–1/2	0	–1/4	0	M–2	
0	x_6	0	1/4	0	–21/2*	0	–1/4	1	–5	→
0	x_2	0	3/4	1	3/2	0	1/4	0	2	
	$z_j - c_j$	0	–1/2	0	–2	0	–1/2	0		
					↑					

Tableau 3

c_B	X_B	x_0	x_1	x_2	x_3	x_4	x_5	x_6	$\bar{X}_B$
0	x_4	0	13/2	0	0	1	1/2	0	0
2	x_0	1	5/21	0	0	0	–5/21	–1/21	M–37/21
1	x_3	0	–1/42	0	1	0	1/42	–2/21	10/21
0	x_2	0	11/14	1	0	0	3/14	1/7	9/7
$z_j - c_j$		0	–23/42	0	0	0	–19/42	–4/21	

Thus an optimal solution of the augmented problem is obtained in tableau 3 and hence an optimal solution of the original problem is given by

$$x_1 = 0, x_2 = 9/7. x_3 = 10/21$$

and $\quad$ max. $z = 64/21$

16.6. The Primal - Dual Algorithm

We have seen that the dual simplex method develops a technique which eliminates the necessity of introducing artificial variables and works with an infeasible basic solution of the problem which satisfy the optimality criterion. The primal-dual algorithm developed by Dantzig, Ford and Fulkerson [112] does introduce artificial variables into the primal but instead of driving the artificial variables to zero as usual, the method works simultaneously on the problem (primal) and its dual. Associated with any feasible solution to the dual, a restricted primal problem is obtained from the auxiliary primal problem by dropping certain variables to satisfy the theorem of complementary slackness (theorem 15.6) and minimize the sum of artificial variables by the revised simplex method. If the optimal solution thus obtained is not feasible to the original primal problem, a new feasible solution to the dual is obtained and the whole process is repeated. After a finite number of iterations, the optimal solution of the restricted primal is either optimal to the original problem or it indicates that the primal problem has no feasible solution or is unbounded

Development of the Method

Let us assume that the primal problem is

$$\begin{aligned} \text{Minimize} \quad & z = c^T X \\ \text{Subject to} \quad & AX = b,\ b \geq 0 \\ & X \geq 0 \end{aligned} \tag{16.31}$$

Its dual then is

$$\begin{aligned} \text{Maximize} \quad & v = b^T W \\ & A^T W \leq c \end{aligned} \tag{16.32}$$

As in the ordinary simplex process, an auxiliary primal problem (Phase I problem) is obtained by adding artificial variables to the constraints of (16.31) and by replacing the objective function by the sum of the artificial variables. Instead of driving the artificial variables to zero as usual, the primal-dual method [112] deals with the primal and the dual problems simultaneously and reduces the number of iterations taken by the two-phase method.

The auxiliary primal problem is

$$\text{Minimize} \quad z^* = \sum_{i=1}^{m} x_{a_i}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij}x_j + x_{a_i} = b_i, \quad b_i \geq 0. \; i = 1, 2,\ldots m$$

$$x_j \geq 0, \; x_{a_i} \geq 0, \qquad i = 1, 2,\ldots m$$

$$j = 1, 2,\ldots n \qquad (16.33)$$

where x_{a_i} is an artificial variable

Its dual is given by,

$$\text{Maximize} \quad v = \sum_{i=1}^{m} b_i w_i$$

$$\text{Subject to} \quad \sum_{i=1}^{m} a_{ij} w_i \leq 0, \qquad j = 1, 2,\ldots n \qquad (16.34)$$

$$w_i \leq 1, \qquad i = 1, 2,\ldots m$$

To initiate the primal-dual algorithm, we need to have the knowledge of a feasible solution to the dual problem (16.32). In problems where $c \geq 0$, $W = 0$ is an obvious feasible solution but in general, a dual solution may not be easily available. We then make use of a simple device suggested by Beale [36] which leads to an immediate feasible solution to the dual. (See section 16.9)

Theorem 16.1

Any feasible solutions to (16.33) and to (16.32) are optimal solutions to the original primal and the dual problems respectively if

$$z^* = 0$$

and $$X^T(A^TW - c) = 0.$$

Proof: Since $x_{a_i} \geq 0$ (i = 1, 2,..m), the condition $z^* = 0$ implies that each $x_{a_i} = 0$ and hence the feasible solution to (16.33) (for which $z^* = 0$) is a feasible solution to the original primal problem (16.31).

For every feasible solutions X, W to the primal and dual problems respectively, We have from the constraints of (16.31) and (16.32).

$$b^T W = X^T A^T W \leq X^T c$$

The condition $X^T (A^T W - c) = 0$ then implies that

$$b^T W = c^T X$$

and hence X and W are optimal solutions to the primal and the dual problems respectively.

It may be noted that the condition $X^T(A^T W - c) = 0$ is the condition of the weak theorem of complementary slackness (theorem 15.6)

Let $\bar{w}_i$ $(i = 1, 2,...m)$ be a feasible solution to the dual problem (16.32). Associated with this solution, a restricted primal problem is obtained from the auxiliary primal (16.33) under the conditions

$$x_j = 0, \quad \text{if } \sum_{i=1}^{m} a_{ij}\bar{w}_i < c_j. \tag{16.35}$$

Let Q denote the set of indices j, for which

$$\sum_{i=1}^{m} a_{ij}\bar{w}_i - c_j = 0$$

i.e.

$$Q = \left\{ j \left| \sum_{i=1}^{m} a_{ij}\bar{w}_i - c_j = 0 \right. \right\} \subset \{1,2,...n\} \tag{16.36}$$

Thus if $\bar{w}_i$ are such that the jth constraint of (16.32) is satisfied with strict inequality, the variable x_j is dropped from the auxiliary primal problem (16.33), to form the restricted primal problem. The dual of the restricted primal is obtained from the auxiliary dual (16.34) by dropping the jth constraint for which

$$\sum_{i=1}^{m} a_{ij}\bar{w}_i - c_j < 0$$

The restricted primal problem is then solved by the revised simplex method. The values $x_{a_i} = b_i$ may be taken as the initial basic feasible solution and the method provides optimal solutions $\left(x_j^0, x_{a_i}^0\right)$ and w_i^0 and $(i = 1,2,....m, j = 1,2,...n)$ to the restricted primal and its dual respectively.

Theorem 16.2

If the minimal value of z* for the restricted primal problem is zero, then the solutions $\left(x_j^0, w_i^0\right)$ are optimal solutions to the original primal and the dual problems respectively.

Proof: The optimal solution to the restricted primal problem is obviously a feasible solution to the auxiliary primal problem. Now this feasible solution to the auxiliary primal and the feasible solution to the dual problem from which this restricted problem is derived satisfy the conditions of theorem 16.1. Hence they are optimal solutions to the original primal and the dual problems respectively.

From the optimal solutions to the restricted primal and its dual, we note that,

$$\sum_i a_{ij} w_i^0 \le 0 \text{ for } j \in Q \text{ and } \sum_i a_{ij} w_i^0 = 0, \text{ if } x_j^0 > 0 \tag{16.37}$$

$$w_i^0 \le 1,\ i = 1, 2, \dots m \text{ and } w_i^0 = 1, \quad \text{if } x_{a_i}^0 > 0 \tag{16.38}$$

Theorem 16.3

If for the restricted primal problem Min z* > 0, then either the original primal problem has no feasible solution or a new feasible solution to the dual problem with a strict increase in the value of its objective function can be obtained.

Proof: Let us define

$$w_i^\theta = \overline{w}_i + \theta w_i^0, \qquad i = 1,2,\dots\dots m \tag{16.39}$$

where θ is a scalar parameter

Now for all $j \in Q$, $\sum_i a_{ij}\overline{w}_i = c_j$ and $\sum_i a_{ij} w_i^0 \le 0$ by (16.37) and for

$j \notin Q, \sum_i a_{ij}\overline{w}_i < c_j$.

Two cases may now arise:

Case (a) For every $j \notin Q, \sum_i a_{ij} w_i^0 \le 0$

(b) There is at least one $j \notin Q$ for which

$\sum_i a_{ij} w_i^0 > 0$

In case (a),

$$\sum_i a_{ij} w_i^\theta = \sum_i a_{ij}\overline{w}_i + \theta \sum_i a_{ij} w_i^\theta \le c_j,$$

and hence w_i^θ (i = 1 ,2,..m) is feasible to the dual problem (16.32) for every $\theta \ge 0$.

The optimal solution $\left(x_j^0, x_{a_i}^0\right)$ to the restricted primal problem is obviously a feasible solution to the auxiliary primal (16.33) and w_i^0 is a feasible solution to its dual (16.34). Since they satisfy the conditions of the theorem of complementary slackness, as can be seen from (16.37), 16.38), and the definition of the associated restricted primal, they are optimal solutions to the auxiliary dual problems. Since Min z* > 0, it implies that the original primal problem has no feasible solution.

In case (b), it follows that

w_i^θ (i = 1, 2,...m) is a feasible solution to the dual problem for every $\theta, 0 < \theta \leq \theta_0$

$$\text{where } \theta_0 = \underset{j}{\text{Min}}\left[-\frac{\sum_i a_{ij}\bar{w}_i - c_j}{\sum_i a_{ij} w_i^0} \middle| \sum_i a_{ij} w_i^0 > 0 \right] \tag{16.40}$$

$$= -\frac{\sum_i a_{ir}^j \bar{w}_i - c_j}{\sum_i a_{ir} w_r^0}$$

The new value of v is

$$v^\theta = \sum_{i=1}^m b_i w_i^\theta = \sum_{i=1}^m b_i \left(w_i + \theta w_i^0\right)$$

$$= v + \theta \sum_{i=1}^m b_i w_i^0 \tag{16.41}$$

From the constraint of (16.33)

$$\sum_{i=1}^m b_i w_i^0 = \sum_j \sum_i \left(a_{ij} w_i^0\right) x_j^0 + \sum_{i=1}^m x_{a_i}^0 w_i^0$$

Since for $x_j^0 > 0, \sum_{i=1}^m a_{ij} w_i^0 = 0$ and for $x_{a_i}^0 > 0, w_i^0 = 1$, by (16.37) and (16.38), we have

$$\sum_{i=1}^m b_i w_i^0 = \sum_{i=1}^m x_{a_i}^0$$

Thus $\quad v^\theta = v + \theta z^*$

and hence $\quad v^\theta > v \quad (16.42)$

Thus, a new feasible solution w_i^θ to the dual problem is obtained with strict increase in the value of the maximizing function.

For the new feasible solution to the dual problem (16.32), we take

$$w_i^1 = \bar{w}_i + \theta_0 w_i^0, \qquad i = 1,2,\ldots m \tag{16.43}$$

in order to get the greatest possible in v and the new restricted primal associated with w_i^1 is formed from (16.33) by the conditions

$$\sum a_{ij} w_i^1 < c_j \quad \text{implies} \qquad x_j = 0 \tag{16.44}$$

Since for those j, for which $x_j > 0$ in the optimal solution of the restricted primal problem

$$\sum_i a_{ij} \overline{w}_i - c_j = 0,$$

and $\sum_i a_{ij} w_i^0 = 0,$ by (16.37),

we have from (16.43)

$$\sum_i a_{ij} w_i^1 - c_j = 0$$

We may therefore take the prior minimizing solution of the restricted primal as an initial basic feasible solution of the new restricted primal.

At each iteration there is a variable x_r in the new restricted primal for which $\sum_i a_{ir} w_i^0 > 0$ and hence can be introduced in place of one of the basic variables and under non-degeneracy assumption z* will be strictly decreased. Thus no basis can be repeated and the process terminates in a finite number of steps either with a basic feasible solution for which z* = 0 and an optimal solution of the original problem is obtained or arrives at a solution where min z* > 0 and $\sum_i a_{ij} w_i^0 \le 0$ for every j which implies that the original problem has no solution.

16.7. Summary of the Primal-Dual Algorithm

The iterative procedure for the primal-dual algorithm may now be summarized as follows:

Steps

1. Obtain a feasible solution w_i^1 (i = 1,2,..m) to the dual problem (16.32) [See section 16.9]
2. Let w_i^{k-1} (i = 1, 2,...m) be a feasible solution to the dual problem at the end of the (k–l)th iteration. Associated with this solution, obtain a restricted primal problem from the auxiliary primal (16.33) by the conditions
$$x_j = 0 \quad \text{if} \ \sum_i a_{ij} w_i^{k-1} < c_j$$
3. Solve the restricted primal problem by the revised simplex method. The optimal solution of the restricted primal problem at the (k–1)th iteration may be taken as an initial basic feasible solution for the new problem. The optimal tableau of the revised simplex method also provides an optimal solution to its dual problem. Thus the restricted primal problem to be solved is

$$\text{Minimize} \quad z^* = \sum_{i=1}^{m} x_{ai}$$

$$\text{Subject to} \quad \sum_{j \in Q_k} a_{ij} x_j + x_{a_i} = b_i, \quad i = 1, 2, \ldots m$$

$$x_j \geq 0 \quad \text{for} \quad j \in Q_k$$

$$x_{a_i} \geq 0$$

where

$$Q_k = \left\{ j \,\middle|\, \sum_i a_{ij} w_i^{k-1} - c_j = 0 \right\}$$

If Min $z^* = 0$, then stop: the optimal solution of the restricted primal is an optimal solution of the original primal problem and w_i^{k-1} $(i = 1,2,\ldots m)$ is an optimal solution of the original dual.

If Min $Z^* > 0$, go to step 4.

4. Evaluate

$$\sum_i a_{ij} w_i^k$$

(i) if $\sum_i a_{ij} w_i^k \leq 0$, for all $j \notin Q_k$.

the primal has no feasible solution and the process terminates.

(ii) if $\sum_i a_{ij} w_i^k > 0$, for at least one $j \notin Q_k$.

take
$$\theta^k = \underset{j}{\text{Min}} \left[-\frac{\sum_i a_{ij} w_i^{k-1} - c_j}{\sum_i a_{ij} w_i^k} \,\middle|\, \sum_i a_{ij} w_i^k > 0 \right]$$

and define
$$w_i^k = w_i^{k-1} + \theta^k w_i^k$$

as a new feasible solution to the dual problem.

5. Repeat the process from step 2 onwards.

16.8. Example

Consider the problem

$$\text{Minimize} \quad z = 6x_1 + 7x_2 + 4x_3 + x_4 + 3x_5$$

$$\text{Subject to} \quad 2x_1 + x_2 + x_3 + 2x_4 + x_5 - x_6 = 3$$

$$x_1 + 6x_2 - x_3 - 5x_4 + 2x_5 - x_7 = 6$$

$$x_j \geq 0, \; j = 1,2,\ldots 7.$$

Its dual problem is given by

$$\begin{array}{ll} \text{Maximize} & v = 3w_1 + 6w_2 \\ \text{Subject to} & 2w_1 + W_2 \leq 6 \\ & w_1+6w_2 \leq 7. \\ & w_1 - w_2 \leq 4 \\ & 2w_1 - 5w_2 \leq 1. \\ & w_1 + 2w_z \leq 3 \\ & -w_1 \leq 0 \\ & -w_2 \leq 0 \\ & w_1, w_2 \text{ unrestricted} \end{array}$$

An initial dual feasible solution is given by $W_1 = (w_1, w_2)^T = (0, 0)^T$ and thus the last two dual constraints are satisfied with equality. Hence $Q_1 = \{6, 7\}$.

Let x_8, x_9 be the artificial variables, so that the restricted primal problem is given by

$$\begin{array}{ll} \text{Minimize} & z^* = x_8 + x_9 \\ \text{Subject to} & -x_6 + x_8 = 3 \\ & -x_7 + x_9 = 6 \\ & x_6, x_7, x_8, x_9 \geq 0. \end{array}$$

The optimal solution to this restricted primal is clearly $x_6 = 0$, $x_7 = 0$, $x_8 = 3$, $x_9 = 6$ and Min $z^* = 9$.

The dual of the above restricted primal is then given by

$$\begin{array}{ll} \text{Maximize} & v = 3w_1 + 6w_2 \\ \text{Subject to} & -w_1 \leq 0 \\ & -w_2 \leq 0 \\ & w_1 \leq 1 \\ & w_2 \leq 1 \\ & w_1, w_2 \text{ unrestricted.} \end{array}$$

Since x_8, x_9 are positive, from the complementary slackness theorem, we note that for an optimal solution of the above dual problem, the last two of its constraints must be satisfied with equality and the optimal solution of the dual is W^0 $\left(w_1^0, w_2^0\right)^T = (1,1)^T$.

To obtain a new feasible solution, we compete $W^{0T} a_j$, for $j \notin Q_1$ and we have

$$W^{0T}a_1 = 3,\ W^{0T}a_2 = 7,\ W^{0T}a_3 = 0,\ W^{0T}a_4 = -3,\ W^{0T}a_5 = 3$$

and thus $\theta_0 = \text{Min}\left\{-\dfrac{0-6}{3}, -\dfrac{0-7}{7}, -\dfrac{0-3}{3}\right\} = 1.$

Thus $W^2 = (0,0)^T + (1, 1)^T = (1,1)^T$ is the new dual solution.

With this new solution W^2, we compute the new Q as $Q_2 = \{2, 5\}$ and the new restricted primal is given by

$$\begin{array}{ll} \text{Minimize} & x_8 + x_9 \\ \text{Subject to} & x_2 + x_5 + x_8 = 3 \\ & 6x_2 + 2x_5 + x_9 = 6 \\ & x_2, x_5, x_8, x_9 \geq 0 \end{array}$$

The optimal solution of the above restricted problem is given by

$$x_2 = 0,\ x_5 = 3,\ x_8 = 0,\ x_9 = 0 \text{ and } z^* = 0$$

Thus, we have an optimal solution to the original problem

$$x_5 = 3,\ x_j = 0, \text{ for } j = 1, 2, 3, 4, 6, 7$$

and the optimal solution to its dual is given by

$$w_1^0 = 1, w_2^0 = 1.$$

All the necessary operations in the primal-dual algorithm can be carried out in the tableau format. If the initial solution to the dual problem is not readily available, we add the artificial constraint to the problem and proceed with the simplex tableau.

16.9. The Initial Solution to the Dual Problem: The Artificial Constraint Technique

We need to have a feasible solution to the dual problem to initiate the primal-dual algorithm. If it is not easily available, we make use of a simple device suggested by Beale [36]

An additional constraint

$$x_0 + \sum_{j=1}^{n} x_j = b_0 \tag{16.45}$$

is added to the constraints of the primal problem, where the additional variable $x_0 \geq 0$ has an associated cost of zero and the constant $b_0 > 0$ is an unspecified arbitrarily large number.

The primal problem (16.31) then becomes

$$\begin{array}{lll} \text{Minimize} & z = \sum\limits_{j=1}^{n} c_j x_j. & \\ \text{Subject to} & x_0 + \sum\limits_{j=1}^{n} x_j = b_0. & \qquad (16.46) \\ & \sum\limits_j a_{ij} x_j = b_i, & i = 1, 2, \ldots m \\ & x_0, x_j \geq 0. & j = 1, 2, \ldots n \end{array}$$

The problem (16.46) is called the modified primal problem. The dual of (16.46), called the modified dual is

$$\text{Maximize} \quad \bar{v} = b_0 w_0 + \sum_{i=1}^{m} b_i w_i$$

$$\text{Subject to} \quad w_0 + \sum_{i=1}^{m} a_{ij} w_i \le c_j \tag{16.47}$$

$$w_0 \le 0$$

A feasible solution to the dual problem is now readily available; it can be easily seen that

$$w_0 = \underset{j}{\text{Min}}[0, c_j] \tag{16.48}$$

$$w_i = 0, \qquad i = 1, 2...m$$

constitutes a feasible solution to the dual problem (16.47)

The primal-dual algorithm may therefore be applied to the modified problems.

Two cases may arise.

(a) The dual problem (16.47) is unbounded which implies that the modified primal is infeasible. It then follows that the original primal problem is also infeasible because if the original primal is feasible, there must be a feasible solution to the modified primal:

(b) Suppose that x_j, (j = 0, 1, 2...n) and w_i^0, (i = 0, 1,...m) are optimal solutions to the pair of dual problems (16.46) and (16.47) respectively.

There are two possible cases

(i) $w_0^0 = 0$. This implies that x_j^0, (j = 1, 2,...n) and w_i^0 (i = 1, 2,...m) are feasible to the original primal and its dual problems respectively where z and v have the same value. Hence they are optimal solutions to the original problems.

(ii) $w_0^0 < O$. In that case Max $\bar{v} \to -\infty$ as $b_0 \to +\infty$ (since b_0 can be made arbitrarily large) and by duality, the objective function z of the modified primal has the same value at the optimum. Now, since x_j^0, (j = 1, 2,...n) is feasible to the original problem, its objective function, which is the same as that of the modified primal has no lower bound.

16.9. Exercises

1. Solve the following linear programming problems by the dual simplex method.

(i) Minimize $z = 2x_1 + 3x_2$

Subject to

$$3x_1 + 4x_2 \ge 6$$
$$x_1 + 3x_2 \ge 3$$
$$x_1 + x_2 \le 3$$
$$x_1, x_2 \ge 0$$

(ii) Minimize $z = 2x_1 + 2x_2 + 4x_3$

Subject to
$$2x_1 + 3x_2 + 5x_3 \geq 2$$
$$x_1 + 4x_2 + 6x_3 \leq 5$$
$$3x_1 + x_2 + 7x_3 \leq 3$$
$$x_1, x_2, x_3 \geq 0$$

(iii) Minimize $z = 2x_1 + 3x_3$

Subject to
$$2x_1 - x_2 - x_3 \geq 3$$
$$x_1 - x_2 + x_3 \geq 2$$
$$x_1, x_2, x_3 \geq 0$$

(iv) Minimize $z = x_1 + 2x_2 + 3x_3$

Subject to
$$x_1 - x_2 + x_3 \geq 4$$
$$x_1 + x_2 + 2x_3 \leq 8$$
$$x_1 - x_3 \geq 2$$

2. Solve the following linear programming problems by the dual simplex method using artificial constraint

(i) Maximize $z = x_1 - x_2 - 2x_3$

Subject to
$$6x_1 + 3x_2 + 4x_3 \leq 8$$
$$3x_1 - 5x_2 + 2x_3 \geq 4$$
$$9x_1 - x_2 - x_3 \geq 3$$
$$x_1, x_2, x_3, \geq 0$$

(ii) Maximize $z = x_1 - 4x_2$

Subject to
$$x_1 - x_2 \leq 1$$
$$-x_1 + x_2 \leq 1$$
$$x_1 \geq 2$$
$$2x_2 \geq 3$$
$$x_1, x_2 \geq 0$$

(iii) Minimize $z = 3x_1 - x_2$

Subject to
$$x_1 + x_2 \geq 4$$
$$x_1 - x_2 \geq -2$$
$$-x_1 + x_2 \geq 3/2$$
$$x_1, x_2, \geq 0$$

3. Solve the following linear programming problems by the primal dual algorithm.

(i) Minimize $z = -2x_1 - x_2$

Subject to
$$6x_1 + x_2 \geq 3$$
$$x_1 + 1.5x_2 \geq 3$$
$$x_1, x_2 \geq 0$$

(ii) Minimize $z = 2x_1 + x_2 - x_4$

Subject to
$$3x_1 + 2x_2 - x_3 - 2x_4 \geq 5$$
$$x_1 + x_2 + x_3 + x_4 \leq 6$$
$$x_j \geq 0, j = 1, 2, 3, 4$$

4. Using the primal-dual algorithm, show that the problem [105]

Minimize $z = -x_1$

Subject to
$$-x_1 + x_2 - x_3 = 1$$
$$x_1 - x_2 - x_4 = 1$$
$$x_1, x_2, x_3, x_4 \geq 0$$

has no feasible solution

5. When is the dual simplex method is preferred to the simplex method for finding a solution of a linear programming problem? When do we consider the primal-dual algorithm is better than the dual simplex method?

CHAPTER 17

Post-Optimization Problems: Sensitivity Analysis and Parametric Programming

17.1. Introduction

In a linear programming problem, we assume that all the coefficients of the problem are given constants. However, for many problems, these constants are either estimates or they vary over time or there are some errors in recording their numerical values. It may also happen that some variable of interest or some constraint was omitted from the problem. It is therefore important to find not only an optimal solution of the given problem but also to determine what happens to this optimal solution when certain changes are made in the system.

The changes may be discrete or continuous. The study of the effect of discrete changes is called sensitivity analysis and if the changes are continuous, it is known as parametric programming.

17.2. Sensitivity Analysis

In this section, we are concerned with the analysis that determines the range of a given element for which the original optimal solution remains optimal. In other words, we are interested in performing a sensitivity analysis of the optimal solution already obtained. If the change is beyond this range, in many cases, it is not necessary to solve the problem over again. Some additional work applied to the optimal tableau will take care of the effect of modification. In other cases however, there is no alternative but to solve the problem afresh.

Suppose that the problem under consideration is

$$\begin{aligned} \text{Minimize} \quad & z = c^T X \\ \text{Subject to} \quad & AX = b. \\ & X \geq 0 \end{aligned} \tag{17.1}$$

where $X \in R^n$ and A is an $m \times n$ matrix.

Let $X_0 = B^{-1}b$, be an optimal basic solution where B denotes the basis matrix.

Changes in the problem which are usually studied can be divided into the following five categories.

1. Changes in the cost vector (c)
2. Changes in the requirement vector (b)
3. Changes in the elements of the technology matrix (A)
4. Addition of a constraint
5. Addition of a variable

It is however possible to imagine a great variety of more complicated problems involving different combinations of the changes described above.

17.3. Changes in the Cost Vector

Let Δc_k be the amount to be added to c_k, the kth component of the cost vector c. It is clear that since a solution of the problem is independent of c_j's, any change in 'c_j' does not disturb the feasibility of the current solution. Hence $X_0 = B^{-1}b$, remains a basic feasible solution.

Now, to preserve optimality, we must have

$$z_k^+ - c_k^+ \leq 0, \quad \text{where + refers to the modified problem} \tag{17.2}$$

If the variable x_k is not in the final basis, we must have

$$z_k^+ - c_k^+ = z_k - (c_k + \Delta c_k) \leq 0$$

or
$$\Delta c_k \geq z_k - c_k \tag{17.3}$$

and Δc_k then has no upper bound. Hence for any change in c_k, satisfying (17.3), the current optimal solution remains optimal and the value of the objective function also does not change since $x_k = 0$.

In the case, when x_k is in the final basis, the evaluations of z_j for all nonbasic variables are affected by any change in c_k and we should have

$$z_j^+ - c_j^+ = \sum_{\substack{i \text{ in basis} \\ i \neq k}} \alpha_{ij} c_j + \alpha_{kj}(c_k + \Delta c_k) - c_j \leq 0 \quad \text{for all j not in the basis}$$

$$= z_j - c_j + \alpha_{kj} \Delta c_k \leq 0$$

or
$$\alpha_{kj} \Delta c_k \leq -(z_j - c_j)$$

Thus,
$$\Delta c_k \leq \frac{-(z_j - c_j)}{\alpha_{kj}}, \quad \text{if } \alpha_{kj} > 0$$

$$\Delta c_k \geq \frac{-(z_j - c_j)}{\alpha_{kj}}, \quad \text{if } \alpha_{kj} < 0$$

Hence
$$\underset{\alpha_{kj}<0}{\text{Max}} \frac{-(z_j - c_j)}{\alpha_{kj}} \leq \Delta c_k \leq \underset{\alpha_{kj}<0}{\text{Min}} \frac{-(z_j - c_j)}{\alpha_{kj}}. \tag{17.4}$$

for all j not in the basis.

Thus if (17.4) is satisfied, changes in c_k will not affect the original optimal basis or the values of the optimal solution. The only change will occur in the optimal value of the objective function z_0. The optimal value will now be given by $z_0 + \Delta c_k x_k$.

Any violation of (17.4) indicates that an improvement in the solution can be obtained by introducing the variable that violates, into the basis in the final optimal tableau of the original problem.

17.4. Changes in the Requirement Vector

Since optimality criterion does not depend on the requirement vector, any change in the requirement vector does not affect the optimality condition. It however affects the values of the basic variables and hence the value of the objective function. Thus if the magnitude of the change in the requirement vector be such that it preserves the feasibility of the optimal basis, then the original optimal basis remains optimal.

Let b_k, the kth element of the requirement vector b be changed to $b_k + \Delta b_k$.

The basic solution of the modified problem associated with the original optimal basis B is then given by

$$\bar{X}_0 = B^{-1}\bar{b}$$

where $\quad \bar{b} = [b_1, b_2, \ldots b_k + \Delta_{bk}, \ldots b_m]^T$

or $\quad \bar{X}_0 = B^{-1}b + B^{-1}\,[0, 0 \ldots 0\ \Delta_k, 0 \ldots 0]^T$

or $\quad \bar{x}_{i0} = x_{i0} + \beta_{ik}\,\Delta b_k, \text{ for all i in the basis} \qquad (17.5)$

where β_{ik} is the element in the i th row and k th column of B^{-1}.

For maintaining the feasibility, we must have

$$\bar{x}_{i0} \geq 0$$

or $\quad x_{i0} + \beta_{ik}\,\Delta b_k \geq 0. \qquad (17.6)$

for all i in the basis

or $\quad \Delta b_k \geq \dfrac{-x_{i0}}{\beta_{ik}}, \qquad \text{for } \beta_{ik} > 0$

and $\quad \Delta b_k \leq \dfrac{-x_{i0}}{\beta_{ik}}, \qquad \text{for } \beta_{ik} < 0$

Thus the range for Δb_k for which the optimal basis remains optimal is

$$\underset{\beta_{ik}>0}{\text{Max}} \frac{-x_{i0}}{\beta_{ik}} \leq \Delta b_k \leq \underset{\beta_{ik}<0}{\text{Min}} \frac{-x_{i0}}{\beta_{ik}}. \qquad (17.7)$$

The solution of the modified problem is given by (17.6) and the change in the value of the objective function is $\sum_{i \text{ in basis}} c_k \beta_{ik} \Delta b_k$.

Suppose that the Δb_k are such that (17.6) is violated by some variables so that these variables become infeasible for the modified problem.

To determine the new optimal solution, it is not necessary to solve the problem from the beginning. This can be done as follows.

We multiply by -1, those rows of the original optimal tableau for which the basic variables become infeasible. Next, we add artificial variables to these rows and replace the infeasible variables in the basis by the artificial variables. The two-phase method may now be applied to find a new optimal solution.

Application of the dual simplex method to find a new optimal solution will however be more rapid.

In the case where the number of the basic variables becoming negative is large, it is desirable to solve the problem from the beginning.

17.5. Changes in the Elements of the Technology Matrix

Let a_{rk} the (r, k)th element of $A = (a_{ij})$ be changed to $a_{rk} + \Delta a_{rk}$.

Two cases may arise.

Case (i): The vector a_k (the k th column of A) is a vector of the optimal basis.

Case (ii): The vector a_k is not a basis vector.

Case (i). Due to change in a_{rk} the optimal basis matrix B becomes

$$\overline{B} = B + \Delta a_{rk} D_{rk}, \tag{17.8}$$

where D_{rk} is an $m \times m$ matrix where all elements are zero except for the rkth element which is equal to unity.

Let us assume that $\overline{B}$ is nonsingular. Now to preserve feasibility we must have,

$$\overline{X}_0 = \overline{B}^{-1} b = (B + \Delta a_{rk} D_{rk})^{-1} b \geq 0$$

or

$$\overline{X}_0 = (I + B^{-1} \Delta a_{rk} D_{rk})^{-1} B^{-1} b$$

$$= (I + B^{-1} \Delta a_{rk} D_{rk})^{-1} X_0 \geq 0 \tag{17.9}$$

$B^{-1} \Delta a_{rk} D_{rk}$ can be written as

$$\begin{pmatrix} 0 & \cdots & \beta_{1r} \Delta a_{rk} & \cdots & 0 \\ 0 & \cdots & \beta_{2r} \Delta a_{rk} & \cdots & 0 \\ 0 & \cdots & \beta_{kr} \Delta a_{rk} & \cdots & 0 \\ 0 & \cdots & \beta_{mr} \Delta a_{rk} & \cdots & 0 \end{pmatrix}$$

where $B^{-1} = (\beta_{ij})$, $\quad i = 1, 2 \ldots m$

$\quad j = 1, 2, \ldots m$

and therefore

$$(I + B^{-1}\Delta a_{rk} D_{rk}) = \begin{pmatrix} 1 & \cdots & \beta_{1r}\Delta a_{rk} & \cdots & 0 \\ 0 & \cdots & \beta_{2r}\Delta a_{m} & \cdots & 0 \\ 0 & \cdots & 1+\beta_{kr}\Delta a_{rk} & \cdots & 0 \\ 0 & \cdots & \beta_{mr}\Delta a_{rk} & \cdots & 0 \end{pmatrix} \tag{17.10}$$

and

$$(I + B^{-1}\Delta a_{rk} D_{rk})^{-1} = \begin{pmatrix} 1 & \cdots & \dfrac{-\beta_{1r}\Delta a_{rk}}{1+\beta_{kT}\Delta a_{rk}} & \cdots & 0 \\ 0 & \cdots & \dfrac{1}{1+\beta_{kr}\Delta a_{rk}} & \cdots & 0 \\ 0 & \cdots & \dfrac{-\beta_{mr}\Delta a_{rk}}{1+\beta_{kr}\Delta a_{rk}} & \cdots & 1 \end{pmatrix} \tag{17.11}$$

Hence we have

$$\overline{X}_0 = \overline{B}^{-1} b = (I + B^{-1}\Delta a_{rk} D_{rk})^{-1} X_0.$$

$$= \begin{pmatrix} x_{10} & - & \dfrac{\beta_{1r}\Delta a_{rk}}{1+\beta_{kr}\Delta a_{rk}} x_{k0} \\ & & \vdots \\ & & \dfrac{1}{1+\beta_{kr}\Delta a_{rk}} x_{k0} \\ & & \vdots \\ x_{m0} & - & \dfrac{\beta_{mr}\Delta a_{rk}}{1+\beta_{kr}\Delta a_{rk}} x_{n0} \end{pmatrix} \tag{17.12}$$

Thus for existence and feasibility of $\overline{X}_0$, we must have

$$1 + \beta_{kr}\Delta a_{rk} > 0 \tag{17.13}$$

and

$$x_{i0} - \frac{\beta_{ir}\Delta a_{rk}}{1+\beta_{kk}\Delta a_{rk}} x_{k0} \geq 0, \quad \underline{i = 1, 2,...m} \tag{17.14}$$

Assuming that $1 + \beta_{kr}\Delta a_{rk} > 0$, we have for all $i \neq k$,

$$\Delta a_{rk} \leq \frac{x_{i0}}{\beta_{ir}x_{k0} - \beta_{kr}x_{i0}}, \qquad \text{for } \beta_{ir}x_{k0} - \beta_{kr}x_{i0} > 0$$

and

$$\Delta a_{rk} \geq \frac{x_{i0}}{\beta_{ir}x_{k0} - \beta_{kr}x_{i0}}, \qquad \text{for } \beta_{ir}x_{k0} - \beta_{kr}x_{i0} < 0$$

Hence for maintaining feasibility, we must have

$$\max_{i\neq k} \frac{x_{i0}}{[\beta_{ir}x_{k0}-\beta_{kr}x_{i0}]<0} \leq \Delta a_{rk} \leq \min_{i\neq k} \frac{x_{i0}}{[\beta_{ir}x_{k0}-\beta_{kr}x_{i0}]>0} \tag{17.15}$$

along with $1 + \beta_{kr}\,\Delta a_{rk} > 0$

Now, to preserve optimality, we must have

$$\bar{z}_j - c_j = c_B^T \bar{B}^{-1} a_j - c_j \leq 0, \quad \text{for all j not in the basis} \tag{17.16}$$

where c_B is the vector of costs associated with the optimal basic variables.

Thus, we must have

$$c_B^T \bar{\alpha}_j - c_j \leq 0, \qquad \text{for all j not in the basis} \tag{17.17}$$

where $\bar{\alpha}_j = \bar{B}^{-1} a_j = \left[I + B^{-1}\Delta a_{rk} D_{rk}\right]^{-1} B^{-1} a_j$

$$= \left[I + B^{-1}\Delta a_{rk} D_{rk}\right]^{-1} \alpha_j$$

As in (17.12), we then have

$$\bar{\alpha}_j = \begin{bmatrix} \alpha_{ij} - \dfrac{\beta_{1r}\Delta a_{rk}}{1+\beta_{kr}\Delta a_{rk}}\alpha_{kj} \\ \vdots \\ \dfrac{1}{1+\beta_{kr}\Delta a_{rk}}\alpha_{kj} \\ \vdots \\ \alpha_{mj} - \dfrac{\beta_{mr}\Delta a_{rk}}{1+\beta_{kr}\Delta a_{rk}}\alpha_{kj} \end{bmatrix} \tag{17.18}$$

From (17.16), (17.17) and (17.18), we obtain

$$\bar{z}_j = \sum_{i\neq k} c_i\alpha_{ij} - \sum_{i\neq k} c_i - \frac{\beta_{ir}\Delta a_{rk}}{1+\beta_{kr}\Delta a_{rk}}\alpha_{kj} + c_k \frac{1}{1+\beta_{kr}\Delta a_{rk}}\alpha_{kj}$$

$$= \sum_{\substack{\text{i in the}\\ \text{basis}}} c_i\alpha_{ij} - \sum_{\substack{\text{i in the}\\ \text{basis}}} c_i \frac{\beta_{ir}\Delta a_{rk}}{1+\beta_{kr}\Delta a_{rk}}\alpha_{kj} \tag{17.19}$$

Thus

$$\bar{z}_j - c_j = z_j - c_j - \sum_{\substack{\text{i in the}\\ \text{basis}}} c_i \frac{\beta_{ir}\Delta a_{rk}}{1+\beta_{kr}\Delta a_{rk}}\alpha_{kj} \tag{17.20}$$

Since by (17.13), $1 + \beta_{kr} \Delta a_{rk} > 0$, (17.20) reduces to

$$z_j - c_j \le \Delta a_{rk} \left[\sum_i c_i \beta_{ir} \alpha_{kj} - (z_j - c_j)\beta_{kr} \right]. \tag{17.21}$$

If the term in brackets in (17.21)

$$\sum_i c_i \beta_{ir} \alpha_{kj} - (z_j - c_j)\beta_{kr}$$

is positive, we have

$$\frac{z_j - c_j}{\sum_i c_i \beta_{ir} \alpha_{kj} - (z_j - c_j)\beta_{kr}} \le \Delta a_{rk} \tag{17.22}$$

and if it is negative,

$$\Delta a_{rk} \le \frac{z_j - c_j}{\sum_i c_i \beta_{ir} \alpha_{kj} - (z_j - c_j)\beta_{kr}} \tag{17.23}$$

Hence in order to maintain the optimality of the new solution, Δa_{rk} must satisfy

$$\max_j \left[\frac{z_j - c_j}{L_j}, L_j > 0 \right] \le \Delta a_{rk} \le \min_j \left[\frac{z_j - c_j}{L_j}, L_j < 0 \right] \tag{17.24}$$

where $\quad L_j = \alpha_{kj} \sum_i c_i \beta_{ir} - (z_j - c_j)\beta_{kr}$

Case (ii): If the change is in the element a_{rk} of the vector a_k not in the basis, the feasibility condition is not distributed. To preserve the optimality we must have,

$$\bar{z}_k - c_k \le 0$$

or $\quad c_B^T B^{-1} \bar{a}_k - c_k \le 0$

or $$c_B^T B^{-1}[a_k + \Delta a_{rk} e_r] - c_k \le 0 \tag{17.25}$$

where e_r is a unit vector with rth element unity.

Thus we have,

$$\bar{z}_k - c_k = z_k - c_k + \Delta a_{rk} \sum_i \beta_{ir} c_i \le 0 \tag{17.26}$$

and therefore $$\Delta a_{rk} \le \frac{-(z_k - c_k)}{\sum_i \beta_{ir} c_i}, \quad \text{if } \sum_i \beta_{ir} c_i > 0 \tag{17.27}$$

and $$\Delta a_{rk} \ge \frac{-(z_k - c_k)}{\sum_i \beta_{ir} c_i}, \quad \text{if } \sum_i \beta_{ir} c_i < 0$$

Hence, if the original solution has to remain optimal, even after a change in a_{rk} of a_k not in the basis, Δa_{rk} must satisfy,

$$\max\left\{\frac{-(z_k - c_k)}{\sum_i \beta_{ir} c_i} \middle| \sum_i \beta_{ir} c_i < 0\right\} \le \Delta a_{rk} \le \min\left\{\frac{-(z_k - c_k)}{\sum_i \beta_{ir} c_i} \middle| \sum_i \beta_{ir} c_i > 0\right\} \tag{17.28}$$

10.6. Addition of a Constraint

Suppose that an additional constraint is added to the problem after an optimal solution has already been obtained. If the optimal solution to the original problem satisfies the new constraint, it is obvious that it is also an optimal solution to the modified problem. If it does not satisfy the new constraint, a new optimal solution has to be found.

Let the new constraint added to the system be

$$\sum_{j=1}^{n} a_{m+1,j} x_j \quad \{\le \quad = \quad \ge) b_{m+1} \tag{17.29}$$

A basis matrix for the modified problem is then,

$$B^+ = \begin{bmatrix} B & 0 \\ U & \pm 1 \end{bmatrix} \tag{17.30}$$

where B is the optimal basis for the original problem, U is a row vector whose elements are the coefficients in the new constraint of the optimal basic variables in B and the last column is a null vector except for the last element which is ± 1, corresponding to the slack, surplus or artificial variable added to the new constraint.

The new solution is

$$X_B^+ = B^{+^{-1}} b^+ = \begin{pmatrix} B^{-1} & 0 \\ \mp UB^{-1} & \pm 1 \end{pmatrix} \begin{pmatrix} b \\ b_{m+1} \end{pmatrix}$$

$$= \begin{pmatrix} X_B \\ \mp UX_B \pm b_{m+1} \end{pmatrix} \tag{17.31}$$

If a slack or surplus variable is added, its coefficient in the objective function is zero and the optimality condition $z_j - c_j \le 0$, remains unchanged. But since the original optimal solution does not satisfy the new constraint, the slack or surplus variable must be negative. This implies that the solution satisfies the optimality condition but is not feasible. Dual simplex method can therefore, be applied to find an optimal solution to the modified problem.

If the new constraint is an equality, an artificial variable is added to it. If the value of the artificial variable in the basic solution is negative, its cost element is taken to be zero and dual simplex method may be applied. If however, the value of the artificial variable is positive, an arbitrarily large positive M is assigned as a cost per unit to this variable and the usual simplex methods is applied.

17.7. Addition of a Variable

Suppose that a nonnegative variable x_{n+1} is added to the original problem. This entails an addition of a column vector a_{n+1} to A and a cost element c_{n+1} to c.

It is obvious that the original optimal solution is feasible to the modified problem. It also remains optimal if $z_{n+1} - c_{n+1} \leq 0$, i.e. if $c_B B^{-1} a_{n+1} - c_{n+1} \leq 0$. If however, $z_{n+1} - c_{n+1} > 0$ the present solution can be improved by introducing a_{n+1} into the basis. The last (optimal) tableau of the original problem is then augmented by the vector a_{n+1} and the simplex algorithm may be applied to find an optimal solution to the modified problem.

17.8. Parametric Programming

As mentioned earlier, we are interested not only in the optimal solution of a linear programming problem but also in how it behaves as changes are made in the coefficients of the problem. When the changes are discrete, such a study is called sensitivity analysis and has been discussed in the previous sections. In the following sections, a similar type of analysis is presented, when the changes are continuous and is known as parametric programming.

17.9. Parametric Changes in the Cost Vector

Suppose that the cost vector c vary continuously as a linear function of a parameter θ, so that the cost vector now becomes

$$c = c^0 + \theta d \tag{17.32}$$

where c^0 and d are constant vectors.

The problem then becomes,

$$\text{Minimize} \quad z = (c^0 + \theta d)^T X \tag{17.33}$$
$$AX = b.$$
$$X \geq 0$$

Let us assume that our problem is nondegenerate and that there exists an optimal solution of the problem for some known value of θ. Without loss of generality, we take that value to be $\theta = \theta_0 = 0$. (It is always possible by a change of origin of θ)

Let $\quad X_0 = B^{-1}b$ be an optimal solution for $\theta = \theta_0 = 0$ (17.34)

where B is the corresponding basis matrix,

When θ varies, the solution remains feasible but the optimality conditions may be disturbed. The solution will remain optimal provided

$$z_j - c_j = \sum_{i=1}^{m} \alpha_{ij}\left(c_i^0 + \theta d_i\right) - \left(c_j^0 + \theta d_j\right)$$

$$= \left(\sum_{i=1}^{m} \alpha_{ij} c_i^0 - c_j^0\right) + \theta\left(\sum_{i=1}^{m} \alpha_{ij} d_i - d_j\right)$$

$$= \left(z_j^0 - c_j^0\right) + \theta(v_j - d_j) \le 0. \tag{17.35}$$

where $$v_j = \sum_{i=1}^{m} \alpha_{ij} d_i$$

Now,

if $v_j - d_j \le 0$ for all j, the solution X_0 remains optimal for all $\theta \ge 0$

if $v_j - d_j \ge 0$ for all j, the solution X_0 remains optimal for all $\theta \le 0$

Apart from these special cases, in general $v_j - d_j$ may have positive as well as negative values and in that case, the range of θ for which the current solution remains optimal is

$$\max_j \left[-\frac{\left(z_j^0 - c_j^0\right)}{v_j - d_j} \middle| v_j - d_j < 0 \right] \le \theta \le \min_j \left[-\frac{\left(z_j^0 - c_j^0\right)}{v_j - d_j} \middle| v_j - d_j > 0 \right]. \tag{17.36}$$

If $v_j - d_j > 0$ for at least one j, then for a sufficiently large increase in θ, $(z_j - c_j)$ will be positive for at least one j and the optimality criterion is violated. The critical value θ_1 of θ, beyond which X_0 is no longer an optimal solution is then given by,

$$\theta = \min_j \left[-\frac{\left(z_j^0 - c_j^0\right)}{v_j - d_j} \middle| v_j - d_j > 0 \right] = -\frac{z_k^0 - c_k^0}{v_k - d_k} \tag{17.37}$$

For $\theta > \theta_1$, we are to find a new optimal solution and a new range of θ, for which this solution remains optimal. This can be easily done by the simplex method.

For $\theta > \theta_1$, $z_k - c_k > 0$ and if $\alpha_{ik} \le 0$ for all i, the problem is unbounded for θ beyond θ_1. If $\alpha_{ik} > 0$, for at least one i, the vector a_k is introduced in the basis and the vector to be eliminated is given by the simplex exit criterion

$$\frac{x_{0r}}{\alpha_{rk}} \min_i \left[\frac{x_{0i}}{\alpha_{ik}} \middle| \alpha_{ik} > 0 \right] \tag{17.38}$$

so that a_r is eliminated from the basis and the basis matrix is changed to B' say.

Now, from (17.35) and (17.37), we have for $\theta = \theta_1$,

$$z'_j - c'_j = z_j^0 - c_j^0 + \theta_1(v_j - d_j)$$

$$= z_j^0 - c_j^0 - \frac{z_k^0 - c_k^0}{v_k - d_k}(v_j - d_j) < 0 \quad \text{for } j \ne k$$

$$\text{and} \quad z'_k - c'_k = 0 \tag{17.39}$$

After simplex transformation, we have for $\theta = \theta_1 + \theta'$, $\theta' \geq 0$,

$$z'_j - c_j = z'_j - c'_j + \theta' (v'_j - d_j), \;\; j \neq r$$

$$z'_r - c_r = -\theta' \frac{(v_k - d_k)}{\alpha_{rk}} \tag{17.40}$$

It is clear that $z'_j - c'_j < 0$ and $v_k - d_k > 0$, and therefore the new solution is optimal for $\theta \geq \theta_1$ if $v'_j - d_j \leq 0$ for all j. If $v'_j - d_j > 0$ for at least one j, there exists a critical value $\theta_2 > \theta_1$, beyond which the new solution is no longer optimal. The critical value θ_2 can be obtained from (17.40)

We can proceed in the same manner to obtain a new solution and a new critical value θ_3.

Thus a series of increasing sequence of critical values $\theta_0, \theta_1, \theta_2, \theta_3$ ... and a series of solutions X_0, X_1, X_2, X_3... may be determined so that X_i remains optimal for all values of θ in the interval $\theta_i \leq \theta \leq \theta_{i+1}$, i = 0, 1, 2, 3.

If θ varies in the negative direction, we can obtain in a similar way a decreasing sequence of critical values θ_i, (i = 0, –1, –2...) and a series of solutions X_i so that X_i remains optimal for all values of θ in the interval

$$\theta_{i-1} \leq \theta \leq \theta_i, \;\; i = 0, -1, -2, -3\ldots$$

17.10. Parametric Changes in the Requirement Vector

Suppose that the requirement vector b is to change continuously as a linear function of a parameter θ. The requirement vector then becomes

$$b = b_0 + \theta b', \; \theta \geq 0. \tag{17.41}$$

where b_0 and b' are constant vectors and θ is a parameter. We consider here the case where θ varies in the positive direction from zero. The case where θ decreases from zero is completely analogous.

Suppose that these exits an optimal solution of the problem for some known value of θ. Without loss of generality we take that value to be $\theta = \theta_0 = 0$.

$$\text{Let} \qquad X_0^B = B^{-1} b_0 \tag{17.42}$$

be an optimal solution for $\theta = \theta_0 = 0$, where B is the corresponding basis matrix. A basic solution of the modified problem is

$$X_\theta^B = B^{-1} b_0 + \theta B^{-1} b' \tag{17.43}$$

Since the calculation of $z_j - c_j$ is independent of the requirement vector, the optimality criterion remains satisfied whatever be the value of θ but the basic solution may cease to be feasible.

Hence, the basic solution will remain optimal as long as

$$\begin{aligned} X_\theta^B &= B^{-1} b_0 + \theta B^{-1} b' \\ &= X_0^B + \theta\, Q \geq 0. \end{aligned} \tag{17.44}$$

where $Q = B^{-1} b'$.

if $Q \geq 0$, the solution X_θ^B remains optimal for all $\theta \geq 0$.

if $Q \leq 0$, the solution X_θ^B remains optimal for all $\theta \leq 0$.

In general however, Q may have both positive and negative elements. If $Q_i < 0$ for at least one i, we cannot increase θ indefinitely without violating feasibility. There exists a critical value θ_1 of θ, beyond which X_θ^B is no longer feasible to the problem.

θ_1 is then given by,

$$\theta_1 = \underset{\substack{i \\ \text{i in the basis}}}{\text{Min}} \left[\frac{-x_{0i}^B}{Q_i} \middle| Q_i < 0 \right] = \frac{-x_{0r}^B}{Q_r} \tag{17.45}$$

If $\theta = \theta_1 + \in$, $\in > 0$, at least one component of X_θ^B becomes negative and X_θ^B ceases to be a feasible solution. Optimality criteria, however, remains satisfied and hence the dual simplex method may be applied to obtain a new optimal solution and a new critical value of θ. The process may be repeated to obtain subsequent critical values and optimal solutions in the corresponding ranges.

17.11. Exercises

1. Consider the problem

   Maximize $z = 5x_1 + 3x_2$

   Subject to $3x_1 + x_2 \leq 1$

   $x_1 + x_2 \leq 1$

   $x_1, x_2 \geq 0$

   Determine the ranges of the discrete changes in the cost coefficients c_1 and c_2 which maintains the optimality of the current solution

2. Consider the problem

   Minimize $z = x_1 + x_2 + x_3$

   Subject to $x_1 - x_2 - 2x_3 \geq 4$

   $2x_1 + 2x_2 - x_3 \geq 2$

   $-x_1 + 3x_2 + x_3 \geq 6$

   $x_1, x_2, x_3 \geq 0.$

   a) Obtain the optimal solution

   b) Find the ranges of the discrete changes in b_1 and b_2 separately so that the optimality of the solution obtained is maintained.

3. Consider the problem

   Minimize $z = x_1 + x_2 - 2x_3$

   Subject to $x_1 - x_2 + 4x_3 \geq 6$

   $-x_1 + 3x_2 + x_3 \leq 10$

   $x_1 + x_3 \leq 4$

   $x_1, x_2, x_3, \geq 0$

a) Obtain the optimal solution

b) Find the separate ranges of the discrete changes in b_1 and b_3 so that the optimality of the solution obtained is maintained.

4. Consider the problem

$$\begin{aligned} \text{Maximize} \quad & z = 2x_1 + x_2 + 4x_3 - x_4 \\ \text{Subject to} \quad & x_1 + 2x_2 + x_3 - 3x_4 \le 8 \\ & -x_2 + x_3 + 2x_4 \le 0 \\ & 2x_1 + 7x_2 - 5x_3 - 10x_4 \le 21 \\ & x_j \ge 0, j = 1, 2, 3, 4. \end{aligned}$$

The optimal simplex tableau for the problem is

c_B	a_B	a_1	a_2	a_3	a_4	a_5	a_6	a_7	X_B
2	a_1	1	0	3	1	1	2	0	8
1	a_2	0	1	–1	–2	0	–1	0	0
0	a_7	0	0	–4	2	–2	3	1	5
	$Z_j - c_j$	0	0	+1	+1	+2	+3	0	

Find the effect of the discrete parameter changes given below, on the optimal solution.

(a) c is changed to $[1, 2, 3, 4]^T$

(b) b is changed to $[3, -2, 4]^T$

5. Consider the problem

$$\begin{aligned} \text{Maximize} \quad & z = 10x_1 + 3x_2 + 6x_3 + 5x_4 \\ \text{Subject to} \quad & x_1 + 2x_2 + x_4 \le 6 \\ & 3x_1 + 2x_3 \le 5 \\ & x_2 + 4x_3 + 5x_4 \le 3 \\ & x_j \ge 0, j = 1, 2, 3, 4. \end{aligned}$$

Find the conditions on discrete changes in a_{11} and a_{23}, so that the new solution remains optimal.

6. Describe the role of duality theory in sensitivity analysis of a linear programming problem.

7. Solve the following linear programming problem with parametric objective function for all values of $\theta \ge 0$.

$$\begin{aligned} \text{Minimize} \quad & z = (2 - 2\theta)x_1] + (4 - \theta)x_2 + (1 - 3\theta)x_3 \\ \text{Subject to} \quad & 2x_1 + 4x_2 + 3x_3 \ge 6 \\ & 5x_1 + x_2 + 2x_3 \le 4 \\ & 2x_1 + 3x_2 + x_3 = 3 \\ & x_1, x_2, x_3 \ge 0 \end{aligned}$$

8. Study the variation in the optimal solution of the following parametric linear programming problem for all values of $\theta \geq 0$.

$$
\begin{aligned}
\text{Minimize} \quad & z = 2x_1 + 4x_2 + x_3 \\
\text{Subject to} \quad & 2x_1 + 4x_2 + 3x_3 \geq 6 + 2\theta \\
& 5x_1 + x_2 + 2x_3 \leq 4 - \theta \\
& 2x_1 + 3x_2 + x_3 = 3 + 3\theta \\
& x_1, x_2, x_3, \geq 0.
\end{aligned}
$$

CHAPTER 18

Bounded Variable Problems

18.1. In linear programming problems, often it happens that instead of being subject to the usual constraints of nonnegativity, the variables are constrained by lower and upper bounds: $l_j \le x_j \le u_j$.

The lower bound l_j could be caused by contractual obligations or policy restrictions and the upper bound u_j could be caused by capacity limitations, resource limitations or the size of the market.

18.2. Bounded from Below

Consider the linear programming problem

$$\text{Minimize} \quad z = \sum_{j=1}^{n} c_j x_j \tag{18.1a}$$

$$\text{Subject} \quad \sum_{j=1}^{n} a_{ij} x_j = b_i, \qquad i = 1, 2, \ldots m \tag{18.1b}$$

$$x_j \ge l_j, \qquad j = 1, 2, \ldots n \tag{18.1c}$$

l_j being an arbitrary constant, positive, negative or zero.

We define $\quad x_j = l_j + y_j, \quad j = 1,2\ldots n \qquad (18.2)$

and substitute them for x_j in the problem.

The problem is then reduced to an equivalent problem,

$$\text{Minimize} \quad z = \sum_{j=1}^{n} c_j y_j \tag{18.3}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} a_{ij} y_j = b'_i, \qquad i = 1, 2, \ldots m$$

$$y_j \ge 0. \qquad j = 1, 2, \ldots n.$$

where $b'_i = b_i - \sum_{j=1}^{n} a_{ij} l_j$

which is in the usual linear programming form. An optimal solution of the problem (18.3) will then yield an optimal solution of the problem (18.1).

We thus note that the lower bound conditions offer no difficulty.

The situation, however, is not as simple in the case of a problem with upper bounds.

18.3. Bounded from Above

Consider now the problem,

$$\text{Minimize} \quad z = \sum_{j=1}^{n} c_j x_j \tag{18.4a}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} a_{ij} x_j = b_i, \qquad i = 1, 2, \ldots m \tag{18.4b}$$

$$x_j \geq 0 \qquad j = 1, 2, \ldots n \tag{18.4c}$$

$$x_j \leq u_j \qquad j = 1, 2, \ldots n \tag{18.4d}$$

where $A = (a_{ij})$ is an $(m \times n)$ matrix, rank $(A) = m$ and it is assumed that all the variables are bounded above because if not, we can always bound a variable by a suitable large number.

The system (18.4a) to (18.4c) is referred to as the original system whereas the system (18.4a) to (18.4d) is called the capacitated or enlarged system.

If we now convert the upper bound inequalities into equations by introducing slack variables v_j, i.e. $x_j = u_j - v_j$, $v_j \geq 0$, $j = 1, 2, \ldots n$ and substitute them for x_j in the problem, it reduces to an equivalent problem

$$\text{Minimize} \quad z = -\sum_{j=1}^{n} c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^{n} a_{ij} v_j = b'_i, \qquad i = 1, 2, \ldots m$$

$$0 \leq v_j \leq u_j \qquad j = 1, 2, \ldots n$$

$$\text{where} \quad b'_i = \sum_{j=1}^{n} a_{ij} u_j - b_i, \qquad i = 1 \ldots m \tag{18.5}$$

which is of the same from as (18.4)

It seems, therefore, that the upperbound constraints must be included explicitly within the original system but this necessitates an enlargement of the basis matrix. The enlarged system then consists of $(m + n)$ constraints and the numbers of variables becomes $2n$ because slack variables are to be added in (18.4d). This makes the computation very expensive.

A more efficient technique has been developed by Charnes and Lemke [78] and Dantzig [101] which takes advantage of the simple structure of the upperbound constraints and slightly modifies the standard simplex procedure and makes it possible to solve the problem without explicit representation of the upperbound constraints in the original system.

In the standard simplex method, a basic solution of the original system is defined as a solution obtained by setting (n – m) nonbasic variables equal to zero and solving the resulting m × m square system for the basic variables. A solution is said to be a basic feasible solution if all the variables are nonnegative and constitutes an extreme point solution. From the set of all extreme point solutions, the simplex method finds an optimal solution to the problem. If the basic solution satisfies the constraints (18.4 c,d) it is a solution of (18.4) and is an extreme point of the polyhedron defined by (18.4 b,c,d). This procedure, however does not work since it does not furnish all the extreme point solutions. (because the problem is not in standard form).

It can be easily shown that

Theorem 18.1. A necessary and sufficient condition for $X \in K$, where $K = \{X : AX = b, 0 \leq X \leq U\}$, A is m × n, r (A) = m; to be an extreme point of K is that (n – m) components of X each have the value zero or u_j and the remaining m components are basic variables (i.e. associated with m independent column vectors of A) and have values between zero and u_j.

From theorem 18.1, it is clear that the bounded variable problem (18.4) can be solved by applying the simplex algorithm to the original system provided that a nonbasic variable is assigned either the value zero or the corresponding upper bound. We then define a basic solution for the bounded variable problem to be one for which (n – m) variables are set equal to either their lower or upper bounds (zero or u_j) and solves the resulting system of m equations (corresponding to the basic matrix). If the basic solution satisfies the lower and upper bound constraints of (18.4) it is called a basic feasible solution to the bounded variable problem.

18.4. The Optimality Criterion

Theorem 18.2. A basic feasible solution to the bounded variable problem is optimal if for nonbasic variables

$$z_j - c_j \leq 0, \qquad \text{if } x_j = 0 \tag{18.6}$$

and

$$z_j - c_j \geq 0, \qquad \text{if } x_j = u_j. \tag{18.7}$$

Proof: As in (16.19), we can write

$$z = z_0 - \sum_j (z_j - c_j)x_j, \qquad \text{for j not in the basis} \tag{18.8}$$

where z_0 is the value of the objective function for the basic variables

$$\text{Hence,}\quad \min z = \min\left[z_0 - \sum_{\substack{j\text{ not}\\ \text{in the basis}}} (z_j - c_j)x_j\right] \tag{18.9}$$

If the conditions (18.6) and (18.7) hold, then any changes in the nonbasic variables will increase the value of objective function. Hence (18.9) gives the minimum value.

18.5. Improving a Basic Feasible Solution

If the current basic feasible solution does not satisfy the optimality conditions (18.6) or (18.7), we can replace a current basic variable by a nonbasic variable and get a new improved solution.

Suppose that we have a basic feasible solution to the bounded variable problem (18.4) with basis B and that the optimality conditions (18.6) or (18.7) are not satisfied by some nonbasic variable x_k

Case 1. Let $x_k = 0$.

For the given solution with $x_k = 0$, we have $z_k - c_k > 0$.

Now, the basic variables can be expressed as

$$x_i = x_{i0} - \sum_{j\in J_1} \alpha_{ij} u_j - \alpha_{ik} x_k, \qquad i \in B \tag{18.10}$$

where J_1 is the set of indices of the nonbasic variables which have values equal to their upper bounds.

Let $\quad b'_i = x_{i0} - \sum_{j\in J_1} \alpha_{ij} u_j \quad$ and we then have

$$x_i = b'_i - \alpha_{ik} x_k, \qquad i \in B \tag{18.11}$$

If $\alpha_{ik} > 0$, an increase in x_k will decrease the value of x_i. We must however, restrict any increase to x_k so that it does not cause any x_i to become negative

For $\alpha_{ik} > 0$, we must therefore have

$$b'_i - \alpha_{ik} x_k \geq 0.$$

$$\text{or}\quad x_k \leq \frac{b'_i}{\alpha_{ik}}$$

$$\text{or}\quad x_k \leq \min_{x_{ik}>0} \frac{b'_i}{\alpha_{ik}} = \frac{b'_r}{\alpha_{rk}} \tag{18.12}$$

Similarly, for those for i which $\alpha_{ik} < 0$, an increase in x_k will also increase x_i but we must be ensured that the value assigned to x_k does not violate its upper bound constraint.

Hence for $\alpha_{ik} < 0$, we must have

$$b'_i - \alpha_{ik} x_k \le u_i$$

$$x_k \le \frac{u_i - b'_i}{-\alpha_{ik}}$$

$$x_k \le \operatorname*{Min}_{\alpha_{ik}<0} \frac{u_i - b'_i}{-\alpha_{ik}} = \frac{u_s - b'_s}{-\alpha_{sk}} \tag{18.13}$$

Moreover, we must have

$$x_k \le u_k \tag{18.14}$$

From (18.12), (18.13) and (18.14), we see that the maximum increase that can be given to x_k is given by

$$\max \quad x_k = \min_{i \in B} \left[\frac{b'_i}{\alpha_{ik}}, \text{ if } \alpha_{ik} > 0; \frac{u_i - b'_i}{-\alpha_{ik}}; \text{ if } \alpha_{ik} < 0, u_k \right]$$

i.e.

$$\max \quad x_k = \min \left[\frac{b'_r}{\alpha_{rk}}, \frac{u_s - b'_s}{-\alpha_{sk}}, u_k \right] \tag{18.15}$$

If max x_k is b'_r/α_{rk}, then x_k replaces x_r in the basis and x_r becomes a nonbasic variable with its value equal to zero. If max x_k is $u_s - b'_s/\alpha_{sk}$, then x_k replaces x_s in the basis and x_s becomes a nonbasic variable at its upper bound u_s and if it is u_k, x_k stays nonbasic but its value changes to u_k.

Case 2. Let $x_k = u_k$ and $z_k - c_k < 0$. Since x_k is at its upper bound, it can only be reduced which makes an improvement in our solution

The basic variables in this case can be expressed as

$$x_i = x_{i0} - \sum_{\substack{j \in J_1 \\ j \ne k}} \alpha_{ij} u_j - \alpha_{ik} x_k, \qquad i \in B \tag{18.16}$$

Let $\bar{b}_i = x_{i0} - \sum_{\substack{j \in J_1 \\ j \ne k}} \alpha_{ij} u_j$ and we then have (18.17)

$$x_i = \bar{b}_i - \alpha_{ik} x_k, \qquad i \in B \tag{18.18}$$

For $\alpha_{ik} > 0$, a decrease in x_k will increase x_i and we must ensure that the value of x_k is not decreased to the extent that the basic variable exceeds its upper bound.

We therefore, must have

$$x_i = \bar{b}_i - \alpha_{ik} x_k \le u_i \qquad \text{for } \alpha_{ik} > 0.$$

or

$$x_k \ge \max_{\alpha_{ik}>0} \frac{\bar{b}_i - u_i}{\alpha_{ik}} = \frac{\bar{b}_p - u_p}{\alpha_{pk}} \tag{18.19}$$

On the other hand, if $\alpha_{ik} < 0$, a decrease in x_k will also decrease x_i and we must ensure that x_i does not become negative.

Thus we must have,

$$x_i = \bar{b}_i - \alpha_{ik} x_k \geq 0. \qquad \text{for } \alpha_{ik} > 0.$$

or $$x_k \geq \max_{\alpha_{ik}<0} \frac{\bar{b}_i}{\alpha_{ik}} = \frac{\bar{b}_q}{\alpha_{qk}} \tag{18.20}$$

and finally we must have

$$x_k \geq 0 \tag{18.21}$$

We can therefore assign to x_k, the value given by

$$\text{Min } x_k = \text{Max}\left[\frac{\bar{b}_p - u_p}{\alpha_{pk}}, \frac{\bar{b}_q}{\alpha_{qk}}, 0\right] \tag{18.22}$$

If min x_k is $b_p - u_p/\alpha_{pk}$, then x_k replaces x_p in the basis and the new value of x_p is u_p. If min x_k is $\bar{b}_q/\alpha_{qk}$, then x_k replaces x_q in the basis and the new value of x_q is zero and if min x_k is zero, then the basis remains unchanged but the value of x_k changes to zero.

After an appropriate change in the nonbasic variable x_k, the above procedure is repeated until the optimality conditions (18.6), (18.7) are satisfied.

Note that in replacement operations indicated above the pivot could be negative.

18.6. Example

Consider the problem

$$\begin{aligned}
\text{Maximize} \quad & 3x_1 + 5x_2 + 2x_3 \\
\text{Subject to} \quad & x_1 + 2x_2 + 2x_3 \leq 10 \\
& 2x_1 + 4x_2 + 3x_3 \leq 15 \\
& 0 \leq x_1 \leq 4, 0 \leq x_2 \leq 3, 0 \leq x_3 \leq 3
\end{aligned}$$

Introducing slack variables x_4, x_5, the problem is reduced to

$$\begin{aligned}
\text{Minimize} \quad & z = -3x_1 - 5x_2 - 2x_3 \\
\text{Subject to} \quad & x_1 + 2x_2 + 2x_3 + x_4 = 10 \\
& 2x_1 + 4x_2 + 3x_3 + x_5 = 15 \\
& 0 \leq x_1 \leq 4, 0 \leq x_2 \leq 3, 0 \leq x_3 \leq 3, \\
& x_4, x_5 \geq 0
\end{aligned}$$

Tableau 1: Initial Tableau

c_B	a_B	a_1	a_2	a_3	a_4	a_5	X_B
0	a_4	1	2	2	1	0	10
0	a_5	2	4	3	0	1	15
	c_j	–3	–5	–2	0	0	
	$z_j - c_j$	3	5 ↑	2	0	0	

Thus x_2 is to be increased from its value of zero. Since Min (10/2, 15/4, 3) = 3, x_2 stays nonbasic with $x_2 = 3$ (upper bound)

The new values of basic variables are.

$$x^1_4 = x^0_4 - \alpha_{12}u_2 = 10 - 2 \times 3 = 4$$
$$x^1_5 = x^0_5 - \alpha_{22}u_2 = 15 - 4 \times 3 = 3$$

By putting $x'_2 = 3 - x_2$, the new tableau becomes

Tableau 2

c_B	a_B	a_1	a'_2	a_3	a_4	a_5	X_B
0	a_4	1	–2	2	1	0	4
0	a_5	2*	–4	3	0	1	3
	c_j	–3	–5	–2	0	0	
	$z_j - c_j$	3 ↑	–5	2	0	0	

a_1 enters the basis

Since Min (4/1, 3/2, 4) = 3/2,

a_5 leaves the basis

Tableau 3

c_B	a_B	a_1	a'_2	a_3	a_4	a_5	X_B
0	a_4	0	0	1/2	1	–1/2	5/2
–3	a_1	1	–2*	3/2	0	1/2	3/2
	c_j	–3	–5	–2	0	0	
	$z_j - c_j$	0	1 ↑	–5/2	0	–3/2	

a'_2 enters the basis

$$\text{Since Min}\left[\infty, \frac{4-3/2}{-(-2)} = \frac{5}{4}, 3\right] = \frac{5}{4}$$

a_1 leaves the basis

The table then becomes

Tableau 4

c_B	a_B	a_1	a'_2	a_3	a_4	a_5	X_B
0	a_4	0	0	1/2	1	–1/2	5/2
5	a'_2	–1/2	1	–3/4	0	–1/4	–3/4
	c_j	–3	–5	–2	0	0	
	$z_j - c_j$	1/2	0	–7/4	0	–5/4	

The new values of basic variables are

$x_4 = 5/2 - 0 \times 4 = 5/2$

$x_2 = -3/4 - (-1/2) \times 4 = 5/4.$

Also, the nonbasic variable x_1 at upper bound is put at zero level by substituting $x_1 = 4 - x'_1$ and we obtain the final tableau as

Tableau 5

c_B	a_B	a'_1	a'_2	a_3	a_4	a_5	X_B
0	a_4	0	0	1/2	1	–1/2	5/2
5	a'_2	1/2	1	–3/4	0	–1/4	5/4
	c_j	–3	–5	–2	0	0	
	$z_j - c_j$	–1/2	0	–7/4	0	–5/4	

which shows that an optimal solution has been attained.

The optimal solution is

$x_1 = 4 - x'_1 = 4 - 0 = 4$

$x_2 = 3 - x'_2 = 3 - 5/4 = 7/4$

$x_3 = 0$

and Min $z = -83/4$

18.7. Exercises

Solve the following problems by the bounded variable algorithm

1. Maximize $z = x_1 + x_2$

   Subject to
$$\begin{aligned} 6x_1 - 3x_2 &\le 9 \\ x_1 + 2x_2 &\le 18 \\ -2x_1 + x_2 &\le 6. \\ 0 \le x_1 &\le 3 \\ 0 \le x_2 &\le 8 \end{aligned}$$

2. Maximize $z = 5x_1 + 2x_2$

   Subject to
$$\begin{aligned} 3x_1 + 2x_2 &\le 9 \\ x_1 + 2x_2 &\le 5 \\ 5x_1 + 2x_2 &\le 10 \\ 0 \le x_1 &\le 6 \\ -4 \le x_2 &\le 5 \end{aligned}$$

3. Maximize $z = 3x_1 + x_3$

   Subject to
$$\begin{aligned} 2x_1 - x_2 &\le 0 \\ -x_1 + 2x_3 &\le 10 \\ x_1 + x_2 + x_3 &\le 10 \\ x_1 &\ge 0 \\ 0 \le x_2 &\le 8 \\ 0 \le x_3 &\le 4 \end{aligned}$$

4. Maximize $z = 3x_1 + 4x_2 + 4x_3$

   Subject to
$$\begin{aligned} 2x_1 + x_2 - x_3 &\le 10 \\ -x_1 + 2x_2 + x_3 &\le 6 \\ x_1 - x_3 &\le 4 \\ 3x_1 - x_2 + 2x_3 &\le 15 \\ 0 \le x_1 &\le 4 \\ 0 \le x_2 &\le 8 \\ 0 \le x_3 &\le 4. \end{aligned}$$

CHAPTER 19

Transportation Problems

19.1. Introduction

One of the earliest and most important applications of linear programming has been the formulation and solution of the transportation problem as a linear programming problem. The classical transportation problem can be described as follows. A certain commodity is available in fixed quantities at a number of origins (sources), a specified amount of which are sent to satisfy its demand at each of a number of destinations. It is assumed that the total demand is equal to the total supply and the transportation cost is proportional to the amount transported. The problem is to determine an optimal schedule of shipments.

In 1939, Kantorovich [264, 265] had shown that a class of transportation type problems arise very frequently in practical situations but the method of solution given was however not complete. The standard form of the problem was first formulated along with a constructive method of solution by Hitchcock [230] in 1941 and later, the problem was discussed in detail by Koopmans [282]. The classical transportation problem is therefore often referred to as the Hitchcock–Koopmans transportation problem. The linear programming formulation and the method of its solution were first given by Dantzig [97].

An intuitive presentation of Dantzig's method is due to Charness and Cooper [68] and is called the "stepping stone method". Other methods of solution which are based on the theory of graphs were given by Ford and Fulkerson [170] and Flood [167]. The method to be described in this chapter was developed by Dantzig by specializing the general simplex method.

19.2. The Mathematical Formulation

Suppose that there are m origins (warehouses) which contain various amounts of a (homogeneous) commodity which has to be shipped to n destinations (retail outlets). Let a_i be the quantity of the commodity available at the origin i and b_j be the quantity required at the destination j. Let c_{ij} be the cost of shipping a unit quantity from the origin i to the destination j. It is assumed that the total demand is equal to the total supply that is,

$$\sum_{i}^{m} a_i = \sum_{j}^{n} b_j \tag{19.1}$$

The condition (19.1) implies that the system is in balance and the problem is then called a *balanced transportation problem.*

The problem is to determine the number of units to be shipped from the origin i to the destination j, so that the total demand at the destinations is completely satisfied and the cost of transportation is minimum.

Let $x_{ij} \geq 0$ be the quantity shipped from the origin i to the destination j. The mathematical formulation of the problem then is,

$$\text{Minimize} \quad z = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \tag{19.2}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} x_{ij} = a_i, \quad a_i > 0, \qquad i = 1, 2,\ldots m \tag{19.3}$$

$$\sum_{i=1}^{m} x_{ij} = b_j, \quad b_j > 0, \qquad j = 1, 2,\ldots n \tag{19.4}$$

$$x_{ij} > 0, \qquad i = 1, 2,\ldots m$$
$$j = 1, 2,\ldots n \tag{19.5}$$

We note that the problem is a linear programming problem with (m + n) equations in mn variables.

Theorem 19.1. The transportation problem always has a feasible solution and is bounded.

Proof : Since $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$, it can be easily verified that $x_{ij} = \dfrac{a_i b_j}{\sum_i a_i}$, $i = 1, 2,\ldots m$; $j = 1, 2,\ldots n$ is a feasible solution to the problem.

Moreover, for every component x_{ij} of a feasible solution, we have

$$0 \leq x_{ij} \leq \min\,[a_i, b_j].$$

The feasible region of the problem therefore is closed, bounded and nonempty and hence there always exists an optimal solution to the problem.

Writing the equations (19.3), (19.4) as,

$$
\begin{array}{llllllll}
x_{11} + x_{12} + .. + x_{1n} & & & & & & & = a_1 \\
 & & x_{21} + x_{22} + .. + x_{2n} & & & & & = a_2 \\
 & & & & x_{m1} + x_{m2} + .. + x_{mn} & & & = a_m \\
x_{11} & & + x_{21} & & + x_{m1} & & & = b_1 \\
 & x_{12} & & x_{22} & & x_{m2} & & = b_2 \\
 & x_{1n} & & x_{2n} & & x_{mn} & & = b_n
\end{array} \quad (19.6)
$$

we see that the transportation problem can be expressed as a standard linear programming problem

$$
\begin{array}{ll}
\text{Minimize} & c^T X \\
\text{Subject to} & AX = b \\
 & X \geq 0
\end{array} \quad (19.7)
$$

where $c = (c_{11}, c_{12}, \ldots c_n, c_{1n}, c_2, c_{2n}, \ldots c_{m1}, \ldots c_{mn})^T$

$b = (a_1 a_2 \ldots a_m, b_1 .. b_n)^T$

$X = (x_{11}, x_{12}, .. x_{1n}, x_{21}, .. x_{2n} \ldots x_{m1}, .. x_{mn})^T$

and

$$
A = \begin{bmatrix}
1_n & 0 & 0 & 0 \\
0 & 1_n & 0 & 0 \\
0 & 0 & 1_n & 0 \\
0 & 0 & 0 & 1_n \\
I_n & I_n & I_n & I_n
\end{bmatrix} \quad (19.8)
$$

an $(m + n) \times mn$ matrix, in which 1 is a sum vector and the subscript n indicates that it has n components.

Any standard linear programming problem therefore is called a transportation problem if its technology matrix has the same structure as (19.8).

The simplex method therefore could directly be applied to find a solution of a transportation problem. However, due to the special structure of the coefficient matrix A, it has been possible to modify the simplex method which is more efficient computationally for such problems. It is also possible to employ methods which are not directly, related to the simplex method but in what follows we shall only discuss the method developed by Dantzig.

The Transportation Array

The special structure of the transportation model enables us to represent the system by a rectangular array having m rows, corresponding to the origins O_i (i = 1, 2, m) and n columns corresponding to the destinations D_j (j = 1, 2,..n) and is called the transportation array.

Transportation Array

	D_1	D_2		D_n	Row Total or Supply
O_1	x_{11} c_{11}	x_{12} c_{12}		x_{1n} c_{1n}	a_1
O_2	x_{21} c_{21}	x_{22} c_{22}		x_{2n} c_{2n}	a_2
	$\vdots$ $\vdots$	$\vdots$ $\vdots$		$\vdots$ $\vdots$	
O_m	x_{m1} c_{m1}	x_{m2} c_{m2}		x_{mn} c_{mn}	a_m
Column Total or Demand	b_1	b_2		b_n	(19.9)

Table 19.1

The (i,j)th cell of the array contains the values of c_{ij} in its lower right hand corner and of x_{ij} in its upper left hand corner. The values of the constants for the first m equations, i.e. the supply a_i are given in a marginal column and for the remaining n equations i.e. the demand b_j in a marginal row. Each row then corresponds to one of the first m equations and each column corresponds to the last n equations of the problem. At any stage of the algorithm, absence of an entry of x_{ij} implies that x_{ij} is nonbasic and hence of zero value. A zero-valued basic variable is indicated by a zero entry (degeneracy).

19.3. Fundamental Properties of Transportation Problems

Theorem 19.2. The matrix A is of rank (m + n – 1)

Proof: The system (19.6) consists of (m + n) equations in mn unknowns but the equations are not independent. Since the sum of the first m rows is equal to the sum of the n last rows, one of the equations is redundant, because it can be obtained from the others. If we add the first m equations and subtract from it the sum of the next (n – 1) equations, we get

$$\sum_{i=1}^{m} x_{in} = \sum_{i=1}^{m} a_i - \sum_{j=1}^{n-1} b_j = b_n, \tag{19.10}$$

which means that rank (A) $\leq$ m + n – 1.

Moreover, it is easy to extract from A, a square matrix of order m + n – 1 for which the determinant does not vanish. Hence the matrix A is of rank m + n = 1.

Corollary 19.1. The number of basic variables in a basis of (19.9) is m + n – 1.

Triangular Basis: A system of linear equations AX = b is said to be a triangular system if the matrix A is triangular, that is when all the elements below or above the main diagonal of A are zero. The elements in the main diagonal, however, must all be nonzero. The system has the property that there is at least one equation that contains only one unknown and when this unknown is eliminated from the remaining equations, the reduced system is again triangular.

A basis for the system AX = b, is said to be a triangular basis if a triangular system is obtained when the nonbasic variables are set equal to zero in the original system of equations.

Theorem 19.3. All bases for the transportation problem are triangular.

Proof: Suppose that we have a transportation array as in (19.9) with m rows and n columns and with marginal constants a_i and b_j all positive. Consider that a particular set of basic variables are placed in their appropriate cells. As each equation of (19.6) corresponds to a row or a column of the array, we need only to prove that there is at least one row or one column of the array that has only one basic variable and that when this row or column is deleted, the reduced array will also have the same property.

We note that it is impossible that any row or column might have no basic variable because this would mean that the corresponding constraint cannot be satisfied as a_i and b_j are positive. Hence all rows and columns must have one or more basic variables.

Let us assume that no row or column has exactly one basic variable. This will mean that all rows and columns must have two or more basic variables.

Let k be the total number of basic variables, in the array. Then, since there are at least two basic variables in each row and in each column, we must have,

$$k \geq 2m$$

and also $\qquad k \geq 2n$

so that we must have $k \geq m + n$ (19.11)

But this is impossible because the number of basic variables is m + n – 1.

It then follows that there must be at least one row or column which has exactly one basic variable. Deleting this row or column, we are left with a reduced array and the argument can be repeated to show that in the reduced array also, there is only one row or column which has exactly one basic variable. The procedure can be repeated again and again and the theorem is thus established.

Theorem 19.4: The values of the basic variables in a transportation problem are given by expressions of the form

$$x_{ij} = \pm \sum_{\text{Some p}} a_p \mp \sum_{\text{Some q}} b_q \tag{19.12}$$

where the upper signs apply to some basic variables and the lower signs apply to the remaining basic variables.

Proof: Consider a transportation problem arranged in an array as in (19.9) and suppose that the values of the basic variables are recorded in the appropriate cells. Since all the bases are triangular, there is at least one row or column which contains exactly one basic variable. Let the value of this basic variable be denoted by x_{pq}. Then

$$x_{pq} = a_p \text{ or } b_q. \tag{19.13}$$

depending on whether x_{pq} is the only basic variable in row p or column q. It should be noted that (19.13) is a special case of (19.12).

If x_{pq} is the only basic variable in row p, then $x_{pq} = a_p$ and we then obtain the reduced array by deleting row p and replacing b_q by $b_q - a_p$. This corresponds to discarding the equation in row p after evaluating x_{pq} and eliminating it from all the other equations. On the other hand, if x_{pq} is the only basic variable in column q, then we obtain the reduced array by deleting column q and replacing a_p by $a_p - b_q$.

As the basis of the reduced array is also triangular, we can repeat the foregoing procedure to find a new basic variable $x_{p'q'}$ such that

$$x_{p'q'} = a_{p'} \text{ or } b_{q'} \text{ or } a_p - b_q \text{ or } b_q - a_p. \tag{19.14}$$

which is again of the form (19.12)

Following the above procedure, we can further reduce the array and continue the process of selecting the basic variables. Clearly at any stage, the row totals for each reduced array are given by expressions of the form

$$\sum_{\text{some p}} a_p - \sum_{\text{some q}} b_q$$

while the column totals for each reduced array are of the form

$$-\sum_{\text{some p}} a_p + \sum_{\text{some q}} b_q$$

This shows that (19.12) is valid because the triangularity of the basis implies that the values of the basic variables are equal to a row total or a column total of some reduced array.

Corollary 19.2. If all a_i and b_j are integers, then the values of the basic variables for any basis are also integers.

Proof: The values of the basic variables can be expressed as the difference between partial sums of the original row and column totals (theorem 19.4). If now all a_i and b_j are integers, their partial sums are also integers.

19.4. Initial Basic Feasible Solution

Consider the transportation array as in (19.9). Select any variable x_{pq} as the first basic variable and make it as large as possible, consistent with the row and column totals, i.e. set

$$x_{pq} = \text{Min } (a_p, b_q) \tag{19.15}$$

If $a_p < b_q$, $x_{pq} = a_p$ and all the other variables in the pth row are nonbasic and given the value zero. The pth row is then removed from further consideration and b_q is replaced by $b_q - a_p$ and the process is repeated to evaluate a basic variable in the reduced array.

If $a_p > b_q$, $x_{pq} = b_q$ and all other variables in the qth column are nonbasic and are set equal to zero. The qth column is then removed and a_p is replaced by $a_p - b_q$.

In case, $a_p = b_q$, either the pth row or the qth column is deleted but not both. If there is only one row (column) but several columns (rows) remain in the reduced array, then the qth column (pth row) is dropped.

In this process, we select $(m + n - 1)$ variables for the initial basis because eventually there will be only one row and one column left and both are dropped after the last variable is evaluated. Thus the values of the variables are uniquely determined as linear combinations of any $m + n - 1$ of the $m + n$ marginals a_1, $a_2, \ldots a_m$; $b_1, b_2, \ldots b_n$.

Since the rank of the system is $m + n-1$, the set of these variables, with the remaining variables equal to zero, constitutes a basic feasible solution of the problem. It is clear that the basis selected by this procedure is triangular. Since all bases for the transportation problem are triangular, we can generate any of them by the above procedure.

Special cases

(a) North-West Corner Rule [68].

A particular case of the method discussed above, is the so called north–west corner rule.

In this method, we begin with the north–west corner cell of the array, i.e. cell (1,1) and allocate $x_{pq} = x_{11} = \min (a_1, b_1)$ and repeat the procedure with the reduced array.

Computationally, however, the above methods are not very practical as the costs are completely ignored here. Usually, the number of iterations required to achieve optimality can be greatly reduced if the basic set is selected with some reference to the values of the cost elements.

Several such methods of determining an initial basic feasible solution are available in the literature.

(b) Matrix Minima (Least Cost) Method. [235]

The method consists in choosing the cell with the smallest cost in the entire array and allocating the first basic variable in that cell.

Suppose that the smallest cost occurs in cell (p.q.). Then allocate x_{pq} = min $[a_p, b_q]$. If $x_{pq} = a_p$, cross out the pth row and replace b_q by $b_q - a_p$. If $x_{pq} = b_q$, cross out the qth column and replace a_p by $a_p - b_q$ and if $x_{pq} = a_p = b_q$, i.e. the row and the column requirements are satisfied simultaneously, cross out either the row or the column but not both. If the minimum cost is not unique, make an arbitrary choice among the minima.

Repeat the process with the reduced array, until all the row requirements and the column requirements are satisfied.

(c) Vogel's Approximation Method (VAM) [375]

Another method which usually provides a solution quite close to optimum is the Vogel's approximation method. The procedure is carried out as follows.

Step 1. Find the difference between the lowest and the next lowest cost entries in each row and each column of the array and record them in a column on the right and in a row at the bottom of the array.

Step 2. Select the row or the column with the largest difference. If the largest difference is not unique, then select the row or the column arbitrarily.

Step 3. Make the maximum possible allocation to the cell with the lowest cost in the selected row or column. If the cell having the lowest cost is not unique, then select from tied ones, the cell to which the maximum allocation can be made. The marginal totals are then adjusted accordingly.

Step 4. Cross out that row or column on which the supply has been exhausted or the demand has been fully satisfied. If both the row and the column totals are satisfied simultaneously, cross out either the row or the column but not both.

Step 5. Repeat the process with the reduced array.

We now illustrate the methods of finding an initial basic feasible solution to a transportation problem by a numerical example

Consider the transportation problem given by the table 19.2

Destination

		D_1	D_2	D_3	D_4	Supply
	O_1	11	13	17	14	250
Origin	O_2	16	18	14	10	300
	O_3	21	24	13	10	400
Demand		200	225	275	250	

Table 19.2

Note that it is a balanced transportation problem.

(a) North–West Corner Rule

An initial basic feasible solution obtained by the N.W.C.R. is given in the following table

	D_1	D_2	D_3	D_4	Supply
O_1	200	50			250
O_2		175	125		300
O_3			150	250	400
Demand	200	225	275	250	

(b) Matrix Minima (Least Cost) Method

Following the matrix minima (least cost) method, let us select the cell (2,4) with the least cost 10 to be allocated first. The method then yields a basic feasible solution as shown in the following table

	D_1	D_2	D_3	D_4	Supply
O_1	200	50			250
O_2		50		250	300
O_3		125	275		400
Demand	200	225	275	250	

Note that if we would have selected the first cell to be allocated to be the cell (3,4), we would have got the same solution as obtained by the north west corner rule,

(c) Vogel's Approximation Method.

	D_1	D_2	D_3	D_4	Supply	Difference
O_1	200 11	13	17	14	250	2
O_2	16	18	14	10	300	4
O_3	21	24	13	10	400	3
Demand	200	225	275	250		
Difference	5	5	1	0		

	D_2	D_3	D_4	Supply	Difference
O_1	50 13	17	14	50	1
O_2	18	14	10	300	4
O_3	24	13	10	400	3
Demand	225	275	250		
Difference	5	1	0		

	D_2	D_3	D_4	Supply	Difference
O_2	175 18	14	10	300	4
O_3	24	13	10	400	3
Demand	175	275	250		
Difference	6	1	0		

	D_3	D_4	Supply	Difference
O_2	125 14	10	125	4
O_3	13	10	400	3
Demand	275	250		
Difference	1	0		

	D_3	D_4	Supply
O_3	275 13	125 10	400
Demand	275	125	

The basic feasible solution thus obtained is shown in the following table,

	D_1	D_2	D_3	D_4	Supply
O_1	200 11	50 13	17	14	250
O_2	16	175 18	14	125 10	300
O_3	21	24	275 13	125 10	400
Demand	200	225	275	250	

19.5. Duality and Optimality Criterion

The dual problem to the standard transportation problem (19.2 – 19.5) can be easily obtained as

$$\text{Maximize} \quad w = \sum_{i=1}^{m} a_i u_i + \sum_{j=1}^{n} b_j v_j \tag{19.16}$$

$$\text{Subject to} \quad u_i + v_j \le c_{ij}; \qquad i = 1, 2, \ldots m$$
$$j = 1, 2, \ldots n$$
$$u_i \text{ and } v_j \text{ unrestricted}$$

From the complementary slackness theorem (Theorem 15.7) we know that the feasible solutions x_{ij} and u_i, v_j are optimal solutions to the transportation problem and its dual if and only if

$$c_{ij} - u_i - v_j \ge 0 \tag{19.17}$$

$$x_{ij}\,(c_{ij} - u_i - v_j) = 0 \tag{19.18}$$

the equation (19.18) implies that if $x_{ij} \ne 0$, $c_{ij} - u_i - v_j$ must be equal to zero. Hence for a basic feasible solution if $x_{ij} > 0$, we must have

$$u_i + v_j = c_{ij}, \quad \text{for } i, j \text{ in the basis} \tag{19.19}$$

a system of (m + n − 1) equations in (m + n) unknowns, (assumption). Thus if a set of values u_i and v_j satisfies (19.19), the basic feasible solution under consideration is optimal if and only if,

$$u_i + v_j \leq c_{ij}, \qquad \text{for all i j.} \qquad (19.20)$$

The dual variables u_i and v_j can be obtained uniquely from (19.19), if one of them is assigned an arbitrary value.

Optimality Criterion

From (19.20), we note that a basic feasible solution to the transportation problem is optimal if and only if it satisfies

$$u_i - v_j - c_{ij} \leq 0, \qquad \text{for all i and j} \qquad (19.21)$$

Following the simplex method for a standard linear program with basis B and the associated cost vector c_B, it can be seen that the optimality criterion for a transportation problem is the same as (19.21)

$$z_{ij} - c_{ij} = c_B^T B^{-1} a_{ij} - c_{ij} = u_i + v_j - c_{ij} \leq 0. \qquad i = 1, 2,...m \quad j = 1, 2,...n$$

since $c_B^T B^{-1}$ represents the dual variable (See section 15.6) and every column vector of A, $a_{ij} = e_i + e_{m+j}$.

The values of u_i and v_j are now so chosen that the coefficients of the basic variables in (19.18) vanish, i.e.

$$u_i + v_j = c_{ij}, \qquad \text{for all basic variables } x_{ij}. \qquad (19.22)$$

The system of equations (19.22) has a matrix which is the transpose of the matrix of the basis under consideration and is of rank m + n − 1, Since any equation in (19.22) may be considered redundant, one of the dual variables u_i, v_j can be assigned an arbitrary value and the remaining m + n − 1 variables can be obtained uniquely.

Now, since the basis is triangular and the transpose of a triangular matrix is also triangular, the matrix of (19.22) is triangular and the dual variables u_i, v_j can be easily obtained by back substitution. In practice, it is convenient to assign zero value to that variable which is associated with the row or the column of the transportation table, containing the greatest number of basic variables.

19.6. Improvement of a Basic Feasible Solution

If the current basic feasible solution is not optimal then there is at least one (i, j) for which $c_{ij} - u_i - v_j < 0$. Then an improved basic feasible solution can be obtained by increasing the value of the nonbasic variable that violates the optimality criterion and adjusting the basic variables so that the row totals and the column totals remain satisfied.

Suppose that for the nonbasic variable x_{pq}

$$c_{pq} - u_p - v_q < 0 \tag{19.23}$$

and the value of x_{pq} is increased to a value θ.

Since the values of the basic variables must satisfy the row and the column totals, some other variable in row p, say x_{pq_1} will have to be reduced by θ. This will necessitate an increase of θ somewhere in column q_1, say in $x_{p_1q_1}$. There should then be a decrease of θ somewhere in row p_1, say in $x_{p_1q_2}$ and so on until eventually balance all rows and columns by arriving back in column q thus forming a loop[1] and the algebraic sum of the θ adjustments is zero. Note that the value of θ is restricted by those x_{ij} from which it is subtracted.

To see that the new solution thus obtained is better than the preceding one, consider the four cells of the transportation table where x_{pq} is a nonbasic variable and the corners of cells of the loop formed with basic variables, have cost elements c_{pq}, c_{pq_1}, c_{p_1q}, $c_{p_1q_1}$ such that $c_{pq} + c_{p_1q_1} < c_{pq_1} + c_{p_1q}$

Let min $\left(x_{pq_1}, x_{p_1q}\right) = x_{pq_1} = \theta$.

In the new solution we have

$$x^1_{pq} = x_{pq} + \theta, \qquad x^1_{pq_1} = x_{pq_1} - \theta$$

$$x^1_{p_1q} = x_{p_1q} - \theta, \qquad x^1_{p_1q_1} = x_{p_1q_1} + \theta.$$

while other variables remain unchanged.

x^1_{pq} then becomes basic and $x^1_{pq_1}$, has value zero and is taken as nonbasic. Thus the new solution still contains m + n – 1 basic variables.

The value of z is then decreased by

$$\theta\left(c_{pq_1} + c_{p_1q} - c_{pq} - c_{p_1q_1}\right).$$

If x_{pq_1} and x_{p_1q} had the same value, only one of them is made nonbasic.

19.7. The Transportation Algorithm

We now summarize below the computational procedure for solving a transportation problem

(a) Find an initial basic feasible solution.

(b) Solve the system of equations

$$c_{ij} - u_i - v_j = 0 \text{ , for all basic variables } x_{ij},$$

[1]A loop in a transportation table is a path which begins and ends in the same cell and has no more than two cells in any row or column.

where the value of one of the variables u_i, v_j is fixed arbitrarily (say $u_1 = 0$). The system of equations thus obtained is triangular and may easily be solved by back substitution. Enter the values of u_i and v_j in the margins of the rows and columns of the transportation table.

(c) Calculate the expression

$$\bar{c}_{ij} = c_{ij} - u_i - v_j$$

for every nonbasic variable and enter them in the lower left corner of the corresponding cells in the transportation table.

(i) If $\bar{c}_{ij} \geq 0$ for all (i,j), the current basic feasible solution is optimal.

(ii) If $\bar{c}_{ij} < 0$, for at least one (i,j), a better solution can be obtained. Proceed to (d).

(d) Select the cell (i, j) for which $\bar{c}_{ij}$ is the most negative. Let this cell be (p,q). The nonbasic variable x_{pq} is then increased to a value θ (say).

(e) Determine a loop, that starts from the cell (p, q) and connects certain other cells occupied by basic variables and ends at the cell (p, q), i.e. the original cell. Assign the maximum possible value to θ without violating the structural and nonnegativity constraints. The variable x_{pq} then becomes basic and the basic variable whose value is now reduced to zero leaves the basis.

(f) Repeat the process from step (b), until an optimal solution is obtained.

19.8. Degeneracy

As in any linear programming problem, degeneracy may occur in a transportation problem also. One or more of the basic variables of a transportation problem may have the value zero and the problem becomes degenerate. We know that the values of the basic variables for a transportation problem are given by the positive or negative difference between a partial sum of the row totals a_i and a partial sum of the column totals b_j. Degeneracy can therefore occur in a transportation problem only if this difference is zero. Degeneracy may occur in the process of determining an initial solution or it may arise at some subsequent iteration.

Suppose that at some stage in the process of finding an initial basic feasible solution, we have a situation where $a_p = b_q$. We then delete either the row or the column and the reduced array thus obtained will have either a zero column total or a zero row total. We must then allocate a zero basic variable to a properly selected cell in this column or row so that the vectors associated with the resulting (m + n – 1) occupied cells must be linearly independent. We thus obtain an initial basic feasible solution but the basis is degenerate became one or more of the basic variables are zero.

Degeneracy may also arise at some subsequent iteration if two or more basic variables have the same value and there is a tie for the variable to leave the basis. In that case, one of these variables is dropped from the basis but the remaining ones are kept which become degenerate variables in the new basis.

In the next iteration, suppose that the θ adjustment indicates that we are to subtract θ from one or more of the degenerate variables. θ_{max} must then be equal to zero and we drop one of these degenerate variables that required a $-\theta$ and bring another variable into the basis with zero value. Thus from one degenerate solution we move to another degenerate solution with no improvement in the value of z and cycling may start.

To avoid these degenerate situations we have to make sure that no partial sum of the a_i can be equal to a partial sum of the b_j. This can be done by purturbing the constants a_i and b_j. One technique due to Orden [361] consists of replacing the values of the ai and bj by

$$\bar{a}_i = a_i + \in, \qquad i = 1, 2, \ldots m,$$

$$\bar{b}_j = b_j, \qquad j = 1, 2, \ldots n-1,$$

$$\bar{b}_n = b_{n+m\in}$$

for $\in = 0$.

We note that

$$\sum_{i=1}^{m} \bar{a}_i = \sum_{i=1}^{m} (a_i + \in) = \sum_{j=1}^{n-1} b_j + b_n + m \in = \sum_{j=1}^{n} \bar{b}_j$$

so that the perturbed problem also has a solution Orden has shown that there exists an $\in_0 > 0$ such that for all $\in$, $0 < \in < \in_0$, degeneracy will never occur. It is however, never necessary to determine θ_0 explicitly. The purpose of the $\in$-procedure is to break the ties and an arbitrarily small $\in$'s enable us to proceed with the computation without any degenerate solutions. When an optimal solution to the perturbed problem is found, we drop the $\in$'s to obtain an optimal solution to the original problem.

However, no transportation problem has ever been known to cycle.

19.9. Examples

Example 1. Let us consider the following problem (Hitchcock) and solve it starting with an initial solution obtained by the Vogel's approximation method.

Destination

Origin	1	2	3	4	a_i	(1)	(2)	(3)	(4)
1	10	5	6	25 / 7	25	1	1	3	X
2	8	20 / 2	7	5 / 6	25	4	1	2	2
3	15 / 9	3	30 / 4	5 / 8	50	1	4	1	1
b_j	15	20	30	35					
(1)	1	1	2	1					
(2)	1	X	2	1					
(3)	1	X	X	1					
(4)	1	X	X	2					

To find an improved solution, we compute $\bar{c}_{ij} = c_{ij} - u_i - v_j$. Let $u_1 = 0$. Since $\bar{c}_{ij} = 0$ for all basic variables, other values of u_i's and v_j's are easily obtained and then $\bar{c}_{ij}$ for all nonbasic variables are evaluated.

	1	2	3	4	a_i	u
1	2 10	2 5	3 6	25 / 7	25	0
2	1 8	20–θ / 2	5 7	5+θ / 6	25	–1
3	15 / 9	+θ / –1 3	30 / 4	5–θ / 8	50	1
b_j	15	20	30	35		
v	8	3	3	7		

	1	2	3	4	a_i	u
1	1 10	2 5	2 6	25 7	25	0
2	0 8	15 2	4 7	10 6	25	–1
3	15 9	5 3	30 4	1 8	50	0
b_j	15	20	30	35		
v	9	3	4	7		

Since $\bar{c}_{ij} \geq 0$, the solution obtained is optimal. The optimal solution is $x_{14} = 25$, $x_{22} = 15$, $x_{24} = 10$, $x_{31} = 15$, $x_{32} = 5$, $x_{33} = 30$.

Example 2. Consider the following transportation problem, where the initial solution is obtained by the matrix minima method. Note that the number of positive basic variables is 4, less than 5 (= 3 + 3 – 1) and hence the problem is degenerate. We therefore perturb the problem as shown below and apply the optimality test

	1	2	3	a_i	u
1	11 8	6 7	$60+\epsilon$ 3	$60 + \epsilon$	–2
2	50 3	1 8	$20+\epsilon$ 9	$70 + \epsilon$	4
3	11	80 3	ϵ 5	$80+\epsilon$	0
b_j	50	80	$80+3\epsilon$		
v	–1	3	5		

Since $\bar{c}_{ij} \geq 0$, for all i,j, the optimal solution obtained is, $x_{13} = 60$, $x_{21} = 50$, $x_{23} = 20$, $x_{32} = 80$ and $x_{33} = 0$.

19.10. Unbalanced Transportation Problem

So far we have assumed that the total supply at all the origins is equal to the total demand at all the destinations so that the supply at the origins is exhausted to meet the demand at the destinations exactly. This implies that

$$\sum_{i}^{m} a_i = \sum_{j}^{n} b_j$$

so that the system is in balance.

In many applications, however, it may be impossible or unprofitable to ship all that is available or to supply all that is required or the total production (supply) either exceeds or is less than the total demand. Such problems are called unbalanced and can be handled by the standard transportation algorithm as explained below.

(a) Supply Exceeds Demand. (overproduction)

Suppose that for a transportation problem, the total quantities available at the origins exceeds the total demand at the destinations and that the demand must be met exactly. This means that some origins will have some undispatched items.

The problem then has the form

$$\text{Minimize} \quad z = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \tag{19.24}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} x_{ij} \le a_i, \qquad i = 1, 2, \ldots m \tag{19.25}$$

$$\sum_{i=1}^{m} x_{ij} = b_j, \qquad j = 1, 2, \ldots n \tag{19.26}$$

$$x_{ij} \ge 0, \qquad i = 1, 2, \ldots m \tag{19.27}$$
$$j = 1, 2, \ldots n$$

$$\sum_{i=1}^{m} a_i > \sum_{j=1}^{n} b_j \tag{19.28}$$

Introducing slack variables x_{in+1} ($i = 1, 2, \ldots m$) in the equations (19.25), we have the problem

$$\text{Minimize} \quad z = \sum_{i=1}^{m}\sum^{n+1} c_{ij}x_{ij} \tag{19.29}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} x_{ij} + x_{in+1} = a_i, \qquad i = 1, 2, \ldots m \tag{19.30}$$

$$\sum_{i=1}^{m} x_{ij} = b_j, \qquad j = 1, 2, \ldots (n+1) \tag{19.31}$$

$$x_{ij} \geq 0, \qquad i = 1, 2, ...m \qquad (19.32)$$
$$j = 1, 2, ...(n+1)$$

where $$b_{n+1} = \sum_{i=1}^{m} a_i - \sum_{j=1}^{n} b_j \qquad (19.33)$$

The slack variable x_{in+1} represents the excess supply, i.e. the quantity of undispatched items at the ith origin which is stored there. Then c_{in+1} is interpreted as the storage cost per unit at the origin i. Suppose that $c_{in+1} = 0$, for all i. We can then imagine a fictitious destination where all the excess supply is sent at zero cost. We therefore add a column in the transportation array, which represents the fictitious destination whose requirement is b_{n+1}. We thus reduce the unbalanced problem to a balanced transportation problem and hence can be solved by the algorithm already discussed.

Even though we have assumed above that the storage costs of undispatched items are zero at all the origins, the assumption is not necessary. In fact, in many practical problems, positive storage costs are incurred at some or all the origins. No matter what values are assigned to c_{in+1}, the problem (19.29) to (19.33) is a balanced transportation problem and hence can be solved in the usual manner.

(b) Demand Exceeds Supply (Underproduction)

In a transportation problem if the total demand exceeds the total supply, the demand at some destinations cannot be completely satisfied, even when all the supplies are dispatched. The problem then is

Minimize $$z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} \qquad (19.34)$$

Subject to $$\sum_{j=1}^{n} x_{ij} = a_i, \qquad i = 1, 2, ...m \qquad (19.35)$$

$$\sum_{i=1}^{m} x_{ij} \leq b_j, \qquad j = 1, 2, ...n \qquad (19.36)$$

$$x_{ij} \geq 0, \qquad i = 1, 2, ...m \qquad (19.37)$$
$$j = 1, 2, ...n$$

where $$\sum_{i=1}^{m} a_i < \sum_{j=1}^{n} b_j \qquad (19.38)$$

Introducing slack variables $x_{m+1,j}$ (j = 1,2,..n) in (19.36), we have the balanced transportation problem,

Minimize $$z = \sum_{i=1}^{m+1} \sum_{j=1}^{n} c_{ij} x_{ij} \qquad (19.39)$$

$$\text{Subject to} \quad \sum_{j=1}^{n} x_{ij} = a_i, \qquad i = 1, 2, \ldots (m+1) \tag{19.40}$$

$$\sum_{i=1}^{m} x_{ij} + x_{m+1,j} = b_j, \quad j = 1, 2, \ldots n \tag{19.41}$$

$$x_{ij} \geq 0, \qquad i = 1, 2, \ldots m+1 \tag{19.42}$$
$$j = 1, 2, \ldots n.$$

$$\text{where} \quad c_{m+1,j} = 0 \qquad j = 1, 2, \ldots m$$

$$\text{and} \quad a_{m+1} = \sum_{j=1}^{n} b_j - \sum_{i=1}^{m} a_i \tag{19.43}$$

The slack variable $x_{m+1,j}$ represents the quantity of unsatisfied demand at the jth destination. We can imagine to create a fictitious origin with a_{m+1} quantities of the item in store which can be sent to the destinations to make up any deficiencies between supply and demand at zero cost. We therefore add a row in the transportation array which represents the fictitious origin where the row total is a_{m+1}, the quantity of the item available at this origin. The balanced transportation problem (19.39) to (19.43) can then be solved by the usual transportation algorithm. We have assumed here that the cost of unsatisfied demand ($c_{m+1,j}$) at all destinations is zero. If however, the cost of unsatisfied demand is positive, i.e. $c_{m+1,j} > 0$. (since failure to meet the demand results in a loss of revenue or goodwill, in many cases, shortages are associated with a penalty cost) the problem can still be solved without introducing any additional computational technique.

(c) Surplus Supply, Demand Unsatisfied.

Situations may also arise where it may not be possible to ship all the items available at the origins or to satisfy all the demands at the destinations.

The problem then is

$$\text{Minimize} \quad z = \sum_{i}^{m} \sum_{j}^{n} c_{ij} x_{ij} \tag{19.44}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} x_{ij} \leq a_i, \qquad i = 1, 2, \ldots m \tag{19.45}$$

$$\sum_{i=1}^{m} x_{ij} \leq b_j, \qquad j = 1, 2, \ldots n \tag{19.46}$$

$$x_{ij} \geq 0, \qquad \text{for all i and j} \tag{19.47}$$
$$j = 1, 2, \ldots n$$

$$\text{where} \quad \sum_{i} a_i \geq \sum_{j} b_j \tag{19.48}$$

Under this formulation, if all $c_{ij} \geq 0$, the problem is trivial, simply ship nothing, that is, $x_{ij} = 0$ for all i and j, gives the optimal solution with the minimum transportation cost as zero. A meaningful problem exists if there are penalty costs for unsatisfied demands and storage costs for surpluses at the origins.

Introducing the slack variables x_{in+1} and x_{m+1j} in (19.45) and (19.46) respectively, we get the problem as

$$\text{Minimize} \quad z = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} + \sum_{i=1}^{m} c_{in+1}x_{in+1} + \sum_{j=1}^{n} c_{m+1j}x_{m+1j} \tag{19.49}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} x_{ij} + x_{in+1} = a_i, \quad i = 1, 2,...m \tag{19.50}$$

$$\sum_{i=1}^{m} x_{ij} + x_{m+1j} = b_j, \quad j = 1, 2,...n \tag{19.51}$$

$$x_{ij} \geq 0, \quad i = 1, 2,...m + 1 \tag{19.52}$$

$$j = 1, 2,...n + 1 \tag{19.52}$$

where c_{in+1} is the storage cost per unit at the i th origin, $c_{m+1,j}$ is the penalty cost per unit of unsatisfied demand at the j th destination, x_{in+1} is the unutilized item at the ith origin and x_{m+1j} is the unsatisfied demand at the jth destination.

We also note that,

$$\sum_{i=1}^{m} a_i - \sum_{i=1}^{m} x_{in+1} = \sum_{j=1}^{n} b_j - \sum_{j=1}^{n} x_{m+1j} = \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij} \tag{19.53}$$

We can imagine a fictitious origin where the unutilized supply is stored and is shipped to the real destinations; and a fictitious destination where the demand is the total extra requirements and the above relations imply that they are numerically equal.

The problem is thus reduced to the balanced transportation problem,

$$\text{Minimize} \quad z = \sum_{i=1}^{m+1}\sum_{j=1}^{n+1} c_{ij}x_{ij} \tag{19.54}$$

$$\text{Subject to} \quad \sum_{j=1}^{n+1} x_{ij} = a_i, \quad i = 1, 2,...m+1 \tag{19.55}$$

$$\sum_{i=1}^{m+1} x_{ij} = b_j, \quad j = 1, 2,...n+1 \tag{19.56}$$

$$x_{ij} \geq 0, \quad \text{for all i and j} \tag{19.57}$$

where $c_{m+1n+1} = 0, \quad a_{m+1} = \sum_{j=1}^{n} b_i, b_{n+1} = \sum_{i=1}^{m} a_i$

There are many other variations of unbalanced transportation problem which can be reduced to balanced transportation problems and can be solved by the usual transportation technique.

19.11. The Transhipment Problem

In our discussion of the transportation problem we have so far assumed that shipments can only be done from origins to destinations directly. In many cases however, the amount that can be sent on the direct route may be limited or situations may arise where it may not be economical to ship directly from origins to destinations. It may therefore be required that the goods available at some origin reach their ultimate destinations via other origins and destinations which act as intermediary stations. It may be possible that some of these intermediary stations are neither origins nor destinations. They merely act as transhipment points and can be considered as origins that produce nothing or as destinations that have zero demand.

The generalized transportation problem in which transhipment through intermediary stations is permitted was first considered by Orden [343a] who has shown that such a problem can easily be formulated as a standard transportation problem.

Suppose that, we have m origins and n destinations. Since in a transhipment problem, any origin or destination can ship to any other origin or destination it would be convenient to number them successively so that the origins are numbered from 1 to m and the destinations from m + 1 to m + n.

Let a_i be the quantities available at the origins and b_j be the demands at the destinations and

$$\sum_{i=1}^{m} a_i = \sum_{j=m+1}^{m+n} b_j.$$

Let x_{ij} (i, j = 1,2,...m + n, j ≠ i) be the quantities shipped from station i to station j and c_{ij} be the unit cost of shipping from i to j (i, j = 1,2...m + n, j ≠ i) where c_{ij} need not be the same as c_{ji}.

The total amount shipped from an origin must be equal to the amount it produces plus what it tranships. Similarly, the total amount received at a destination must be equal to its demand plus what it tranships.

We then wish to find

$$x_{ij} \geq 0 \qquad i, j = 1, 2, \ldots m + n,\ i \neq j. \tag{19.58}$$

which Minimize $$z = \sum_{i=1}^{m+n} \sum_{\substack{j=1 \\ i \neq j}}^{m+n} c_{ij} x_{ij}$$

Subject to $\sum_{\substack{j=1\\j\neq i}}^{m+n} x_{ij} - \sum_{\substack{j=1\\j\neq i}}^{m+n} x_{ji} = a_i,$ $\quad i = 1, 2,...m \quad (19.60)$

$$\sum_{\substack{i=1\\i\neq j}}^{m+n} x_{ij} - \sum_{\substack{i=1\\i\neq j}}^{m+n} x_{ji} = b_j, \qquad j = m + 1,...m + n. \qquad (19.60)$$

The system (19.58) to (19.61) is a linear programming problem which is similar to a transportation problem but not exactly since the coefficients of Σx_{ji}'s are −1. The problem however, may easily be converted to a standard transportation problem.

Let $t_i = \sum_{\substack{j=1\\j\neq i}}^{m+n} x_{ij},$ $\quad i = 1, 2,...m \quad (19.62)$

and $t_j = \sum_{\substack{i=1\\i\neq j}}^{m+n} x_{ji},$ $\quad j = m + 1,...m + n \quad (19.63)$

where t_i represents the total amount of transhipment through the i th origin and t_j represents the total amount shipped out from the jth destination as transhipment.

Let T > 0 be a sufficiently large number so that

$$t_i \leq T, \text{ for all i and } t_j \leq T \text{ for all j} \qquad (19.64)$$

If we now write $t_i + x_{ii} = T$, then the nonnegative slack variable x_{ii} represents the difference between T and the actual amount of transhipment through the ith origin. Similarly if we let $t_j + x_{jj} = T$, then the nonnegative slack variable x_{jj} represents the difference between T and the actual amount of transhipment through the jth destination.

The transhipment problem then reduces to

Minimize $\sum_{i=1}^{m+n}\sum_{j=1}^{m+n} c_{ij}x_{ij}$ $\quad (19.65)$

Subject to $\sum_{j=1}^{m+n} x_{ij} = a_i + T,$ $\quad i = 1, 2,...m \quad (19.66)$

$$\sum_{j=1}^{m+n} x_{ij} = T, \qquad i = m + 1,...m + n. \qquad (19.67)$$

$$\sum_{i=1}^{m+n} x_{ij} = T, \qquad j = 1, 2,...m. \qquad (19.68)$$

$$\sum_{i=1}^{m+n} x_{ij} = b_j + T, \qquad j = m + 1,...(m + n). \qquad (19.69)$$

$$x_{ij} \geq 0, \qquad i = 1, 2, \ldots m + n \qquad (19.70)$$
$$j = 1, 2, \ldots m + n$$

where $\quad c_{ii} = 0, \qquad i = 1, 2, \ldots m + n.$

The system (19.65) to (19.70) represents a standard transportation problem with (m + n) origins and m + n destinations.

Note that T can be interpreted as a buffer stock at each origin and destination. The question now arises as to what value should be assigned to T. Since we are assuming that any amount of goods can be transhipped at any point, T should be large enough to take care of all transhipments. It is clear that the volume of goods transhipped at any point cannot exceed the amount produced (or received) and hence we take

$$T = \sum_{i=1}^{m} a_i \qquad (19.71)$$

The transhipment problem may also be presented in a tabular form as given below and the standard transportation algorithm may then be used to obtain an optimal solution. We note that the solution of the problem contains 2m + 2n – 1 basic variables. However, m + n of these variables appearing in the diagonal cells represent the remaining buffer stock and if they are omitted, we have (m + n – 1) basic variables of our interest

Transhipment Tableau

		Origin (O) 1 2 m	Destination (D) m+1 m+2 m+n	
origin (O)	1 2 ⋮ m	O to O	O to D	a_1+T Quantities a_2+T leaving ⋮ origins a_m+T
Destination (D)	m+1 m+2 ⋮ m+n	D to O	D to D	T Quantities T leaving ⋮ destinations T
		T T ...... T Quantities arriving at origins	b_1+T b_2+T ..... b_n+T Quantities arriving at destinations	

Table 19.3

Example. Consider the following transhipment problem involving two origins and two destinations. The availabilities at the origins, the requirements at the destinations and the costs of transportation are given in the table below.

	O_1	O_2	D_1	D_2	a_i
O_1	0	1	3	4	5
O_2	1	0	2	4	25
D_1	3	2	0	1	–
D_2	4	4	1	0	–
b_j	–	–	20	10	

Find the optimal shipping schedule.

Since $T = \Sigma_i a_i = 30$, we convert the problem into a balanced transportation problem by adding 30 units to each a_i and b_j.

	O_1	O_2	D_1	D_2	a_i
O_1	0	1	3	4	35
O_2	1	0	2	4	55
D_1	3	2	0	1	30
D_2	4	4	1	0	30
	30	30	50	40	

Using Vogel's approximation method for an initial basic feasible solution and checking its optimality, we have the following table

	O_1	O_2	D_1	D_2	
O_1	30 0	 0 1	 0 3	5 4	35
O_2	 2 1	30 0	25 2	 1 4	55
D_1	 6 3	 4 2	25 0	5 1	30
D_2	 8 4	 7 4	 2 1	30 0	30
	30	30	50	40	

Since all $\bar{c}_{ij}$ are nonnegative, the solution obtained is optimal. Ignoring the allocations in the diagonal cells, we obtain the optimal solution for the transhipment problem. The optimal schedule is

$O_1 \rightarrow D_2$, 5 units
$O_2 \rightarrow D_1$, 25 units
$D_1 \rightarrow D_2$, 5 units
and Min Z = 75

19.12. Exercises

1. Solve the transportation problem given by the following table.

Destination

		D_1	D_2	D_3	D_4	Supply
	O_1	5	9	4	2	20
Origin	O_2	1	4	1	2	30
	O_3	2	1	3	3	50
Demand		30	10	20	40	

2. Consider the following transportation problem

D_1	D_2	D_3	D_4	a_i	
O_1	63	71	99	31	93
O_2	27	48	14	56	71
O_3	81	21	82	35	47
b_j	60	45	71	35	

Find an initial basic feasible solution by the north–west corner rule and obtain an optimal solution.

3. Find an optimal solution of the problem discussed in section 19.4 starting with each of the initial basic feasible solution obtained.

4. Consider the following transportation problem.

D_1	D_2	D_3	D_4	a_i	
O_1	85	60	30	50	130
O_2	75	65	35	45	150
O_3	80	60	20	55	220
b_j	120	110	150	120	

Find an initial basic feasible solution by (i) the north–west corner rule and (ii) the matrix minima method and obtain optimal solutions.

5. Consider the following transportation problem

	D_1	D_2	D_3	D_4	a_i
O_1	5	8	3	6	30
O_2	4	5	7	4	50
O_3	6	2	4	5	40
b_j	30	20	40	30	

Find an initial basic feasible solution by the Vogel's approximation method and obtain an optimal solution.

6. Show that a balanced transportation problem with some or all $c_{ij} < 0$ can be converted to an equivalent transportation problem with all $c_{ij} > 0$.

7. Consider four bases of operations B_i and three targets T_j. The tons of bombs per aircraft from any base that can be delivered to any target are given in the following table:

$$\begin{array}{cc|c}
 & & \text{Target} \\
 & & \begin{array}{ccc} T_1 & T_2 & T_3 \end{array} \\
\text{Base} & \begin{array}{c} B_1 \\ B_2 \\ B_3 \\ B_4 \end{array} & \begin{bmatrix} 8 & 6 & 5 \\ 6 & 6 & 6 \\ 10 & 8 & 4 \\ 8 & 6 & 4 \end{bmatrix}
\end{array}$$

The daily sortie capability of each of the four bases is 150 sorties per day. The daily requirement in sorties over each individual target is 200. Find the allocation of sorties from each base to each target which maximizes the total tonnage over all the three targets.

8. Prove that for the balanced transportation problem, the optimal solution is not affected if for any row r, the unit cost c_{rj} is replaced by $c_{rj} + \alpha_r$ or for any column k, c_{ik} is replaced by $c_{ik} + \beta_k$ where α_r and β_k are constants.

9. Suppose that for a balanced transportation problem, an infeasible basic solution satisfying the optimality criterion (19.20) is available. Discuss how the dual simplex method can be applied to find an optimal solution of the problem.

10. Consider the following bounded variable transportation problem called the capacitated transportation problem:

$$\text{Minimize} \quad z = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} x_{ij} = a_i, \qquad i = 1, 2, \ldots m.$$

$$\sum_{i=1}^{m} x_{ij} = b_j, \qquad j = 1, 2, \ldots n.$$

$$0 < x_{ij} \leq d_{ij}, \qquad \text{for all } i, j.$$

$$\text{where} \quad \sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j, \qquad a_i > 0,\ b_j > 0,\ c_{ij} \geq 0$$

Modify the bounded variable method (chapter 18) to derive an algorithm that would solve the problem.

11. A company produces a product in three plants located at places L_1, L_2, L_3 and supplies its product to warehouses W_1, W_2, W_3, W_4 and W_5 at different places. The plant capacities, warehouse requirements and the unit transport cost is given in the following table

		Warehouse					
		W_1	W_2	W_3	W_4	W_5	Capacity
	L_1	6	6	4	8	4	900
Plant	L_2	7	6	5	4	7	500
	L_3	6	6	3	5	8	800
Requirement		500	400	800	400	400	

Find an optimum distribution plan for the company in order to minimize the total transportation cost.

12. Solve the following transportation problem

		Destination					
		D_1	D_2	D_3	D_4	D_5	a_i
	O_1	7	4	7	2	7	9
	O_2	2	3	4	7	8	4
Origin	O_3	7	2	4	8	4	2
	O_4	8	7	1	4	3	7
	b_j	2	7	8	3	2	

13. Consider the following transportation problem:

		Warehouse						
		W_1	W_2	W_3	W_4	W_5	W_6	Supply
	P_1	–	6	9	5	11	11	20
	P_2	9	9	9	12	2	10	50
Plant	P_3	2	6	11	8	6	10	90
	P_4	5	7	7	3	7	5	60
Demand			40	40	60	40	20	20

It is not possible to transport any quantity from plant P_1 to warehouse W_1.

Find an initial basic feasible solution by the Vogel's approximation method and obtain an optimal solution of the problem.

14. Generalized Transportation Problems:

Consider the problem

$$\text{Minimize} \quad Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} d_{ij} x_{ij} = a_i, \; a_i > 0, \qquad i = 1, 2, \ldots m.$$

$$\sum_{i=1}^{m} x_{ij} = b_j, \quad b_j > 0, \qquad j = 1, 2, \ldots n.$$

$$x_{ij} \geq 0$$

where $d_{ij} > 0$ for all i, j.

The problem differs from the transportation problem in that the coefficients of the x_{ij} in the constraints may not be unity in this case. However, the structure of these problems is similar to that of the transportation problem and are referred to as the generalized transportation problems.

Extend the transportation algorithm to solve such problems.

15. Solve the transhipment problem involving two origins and three destinations for which the origin availabilities, the destination requirements and the costs for shipping are given below.

	O_1	O_2	D_1	D_2	D_3	Available
O_1	0	2	6	4	2	50
O_2	2	0	1	3	5	50
D_1	6	1	0	1	2	–
D_2	4	3	1	0	3	–
D_3	2	5	2	3	0	–
Required	–	–	30	30	40	

Chapter 20

Assignment Problems

20.1. Introduction and Mathematical Formulation

Assignment problems are special type of allocation problems. In its simplest form the problem can be stated as follows. Suppose that we have n jobs to perform and n persons, each of whom can perform each of the jobs but with varying degree of efficiency. It is assumed that through performance tests (for example, the number of hours of ith person takes to perform the jth job), an estimate of the cost of assigning the ith person to the jth job can be determined. Let this be denoted by c_{ij}. The number c_{ij} is then a measure of effectiveness of the ith person to the jth job. We wish to assign only one job to each person so that the total cost of performance is minimum.

Suppose that the ith person is assigned to job p_i. The problem then is to find a permutation $(p_1, p_2, .. p_n)$ from the set of n! permutations such that the total assignment cost.

$$z = \sum_{i=1}^{n} c_{ip_i} \tag{20.1}$$

is minimum

Stated in this form, the problem is evidently combinatorial. For small n it may be possible to enumerate all the n! permutations and calculate the corresponding costs, the least value of which provides an optimal solution. But the number of possibilities grow rapidly with n and even for a moderately large n, this enumerative procedure is very much time consuming and not at all practical.

Alternatively, the problem can be formulated as follows.

$$\text{Let } x_{if} = \begin{cases} 1, & \text{if the ith person is assigned to the jth job} \\ 0, & \text{if not} \end{cases} \tag{20.2}$$

The problem can then be stated as

$$\text{Minimize} \quad z = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \tag{20.3}$$

$$\text{Subject to} \quad \sum_{i=1}^{n} x_{ij} = 1, \quad j = 1, 2, \ldots n \quad \text{(each job is assigned to only one person)} \tag{20.4}$$

$$\sum_{j=1}^{n} x_{ij} = 1, \quad i = 1, 2, \ldots n \quad \text{(each person is assigned only one job)} \tag{20.5}$$

$$x_{ij} = 0 \text{ or } 1 \text{ for all } i \text{ and } j \tag{20.6}$$

This is an integer linear programming problem (i.e. a linear programming problem where the variables are restricted to be integers) which may be very difficult to solve. If however, the last constraint (20.6) is replaced by the condition $x_{ij} \geq 0$, it reduces to a transportation problem with each $a_i = b_j = 1$. We know that (see corollary 19.2) the solution of a transportation problem will be nonnegative integers when a_i, b_j are integers and therefore in this case, the solution will automatically satisfy the constraint (20.6). The assignment problem is therefore, a special case of the transportation problem. The structure of the problem however is such that every basic feasible solution is degenerate since exactly n basic variables must have the value unity and the (n – 1) basic variables must therefore, all be equal to zero. The problem therefore is highly degenerate and solving an assignment problem by transportation technique can be very frustrating. A very convenient procedure however, is available for assignment problems and is known as Hungarian method.

20.2. The Hungarian Method

In 1931, the Hungarian mathematician König [280] published a theorem on linear graphs, which was generalized in the same year by another Hungarian mathematician Egerváry [144]. Based on König–Egerváry theorem, Kuhn [290] designed a technique to solve the so-called assignment problem and called his algorithm the Hungarian method. As we shall see, the Hungarian method is much more suitable as it is not affected by degeneracy.

The method relies on the following theorems.

Theorem 20.1. If all the elements in any row or column of the cost matrix in an assignment problem are increased or decreased by the same amount, an equivalent assignment problem is obtained.

Proof: We are to show that if $X = (x_{ij})$ is an optimal solution to the assignment problem with the cost matrix (c_{ij}), then it is also optimal for the problem with the cost matrix (c'_{ij}) where

$$c'_{ij} = c_{ij} \pm p_i \pm q_j, \qquad p_i, q_j \text{ being arbitrary real numbers.}$$

$$\text{Now,} \quad z' = \sum_{i=1}^{n}\sum_{j=1}^{n} c'_{ij} x_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} \left(c_{ij} \pm p_i \pm q_j\right) x_{ij}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij} \pm \sum_{i=1}^{n}\sum_{j=1}^{n} p_i x_{ij} \pm \sum_{i=1}^{n}\sum_{j=1}^{n} q_j x_{ij}$$

$$= z \pm \sum_{i=1}^{n} p_i \pm \sum_{j=1}^{n} q_j$$

since $\sum_{i=1}^{n} x_{ij} = 1, \ \sum_{j=1}^{n} x_{ij} = 1.$

which shows that the set of x_{ij} will optimize both the problems.

Theorem 20.2: If in an assignment problem, some of the cost elements c_{ij} are zero and the others are positive, then a set of x's that are all zero except perhaps where $c_{ij} = 0$, must be optimal to the problem.

Proof: Since all $c_{ij} \geq 0$, for any assignment schedule, $z \geq 0$. Now, the value of z corresponding to the given set of x's, is zero and hence it must be optimal to the problem.

By applying theorem 20.1, the cost matrix (c_{ij}) is transformed to (c'_{ij}), where each $c'_{ij} \geq 0$ with at least one zero element in each row and each column. We then need to find the maximum number of independent zero elements so that an attempt can be made to find an assignment among the zeros. For this we state the König–Egerváry theorem, where we make use of the following definitions.

Definitions: The elements of a matrix are also called points. A row or a column of the matrix is known as a line. A set of points of a matrix is said to be independent if none of the lines contains more than one point of the set. A single point is regarded as independent.

Theorem 20.3. (König–Egerváry)

The maximum number of independent zero elements in a square matrix is equal to the minimum number of lines required to cover all the zeros in the matrix.

We therefore, repeatedly use theorem 20.1 to create zeros in the cost matrix of the problem and then make use of theorem 20.3 to find a set of independent zeros which provides an optimal solution to the problem.

20.3. The Assignment Algorithm

Various steps of the Hungarian method for obtaining an optimal solution to an assignment problem may be summarized as follows.

Step 1. Subtract the smallest element in each row of the cost matrix from every element in that row.

Step 2. Subtract the smallest element in each column of the reduced matrix from every element in that column.

The reduced matrix will then have nonnegative elements with at least one zero in each row and in each column. If it is possible to find a set of n independent zeros, then an assignment among the independent zeros will provide an optimal solution to the problem. According to theorem 20.3 we then proceed as follows.

Step 3. Draw the minimum number of lines say n_1, through rows and columns which will cover all the zeros. This can be done by drawing the lines in such a way that each line covers as many zeros as possible If $n_1 = n$, an optimal solution has been reached. This optimal assignment corresponds to n independent zeros, i.e. where there is one zero in each row and in each column.

If however, $n_1 < n$, we further create zeros in the cost matrix and once again verify whether it is possible to get a set of n independent zeros.

Step 4. Find the smallest element not covered by the n_1 lines drawn in step 3. Let this element be θ. Subtract θ from each of the elements not covered by the n_1 lines and add it to the elements at the intersections of these lines. This procedure is equivalent to

(i) subtracting θ from all the elements of the cost matrix.

(ii) adding θ to all the elements of the covered rows.

(iii) adding θ to all the elements of the covered columns.

and hence is carried out according to theorem 20.1 It is therefore clear that this procedure does not affect the optimal assignments.

Step 5. Draw the minimum number of lines say n_2 through rows and columns that cover all the zeros in the new reduced matrix.

If $n_2 = n$, an optimal solution is reached.

If $n_2 < n$, repeat step 4, as often as necessary until a reduced matrix is obtained for which $n_r = n$ and then the independent zeros provide an optimal solution to the problem. (follows from theorem 20.3)

The optimum is always reached after a finite number of steps.

To find the independent zero elements, the following procedure may be followed.

A row which contains only one zero is first selected. This zero becomes an independent zero element and is enclosed by a square. All other zeros in the column of this enclosed zero are then crossed. Proceed in this manner until all the rows have been examined. The process is then repeated with the columns of the matrix thus reduced.

Two things may happen,

(a) All the zeros in the matrix are either enclosed or crossed, or

(b) there is no row or column in the reduced matrix containing a single zero.

In case (a), the number of independent points determined, is maximal and the enclosed zeros provide an optimal solution.

In case (b), a row or a column having the minimum number of zeros is chosen arbitrarily and one of the zeros is enclosed as an independent point.

The remaining zeros in the selected row or column are then crossed. The process is continued and we eventually obtain a matrix with all the zeros either enclosed or crossed. The enclosed zeros are then independent points and yield an optimal solution to the problem.

Remarks

1. There may be situations where a particular assignment is not permissible. In such cases, the cost of such an assignment is taken to be very high in order to prevent this assignment in the optimal solution.
2. If the number of jobs is less than the number of workers, fictitious jobs are added to the problem and if the number of workers is less than the number of jobs, fictitious workers are introduced. The corresponding costs are taken to be zero.

20.4. Variations of the Assignment Model

(a) In the assignment problem discussed above, we considered that each person is assigned one job and each job is assigned to one person. There may however, be situations where a person can spend a fraction of his time in one job and a fraction of his time in another. The problem can then be expressed as

$$\text{Minimize} \quad z = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij}$$

$$\text{Subject to} \quad \sum_{j=1}^{n} x_{ij} = 1, \qquad i = 1, 2, \ldots n$$

$$\sum_{i=1}^{n} x_{ij} = 1, \qquad j = 1, 2, \ldots n$$

$$x_{ij} \geq 0, \qquad \text{for all } i, j.$$

where x_{ij} is the fractions of time that the ith person spends in the jth job. By corollary 19.2 it is clear that in the optimal solution x_{ij} can take only the values zero or one. This implies that any optimal solution to the problem will assign one person full time to one job.

(b) Frequently, we have situations, where there are many identical jobs which require the same basic qualifications. Such jobs can be grouped into a job category and similarly if there are individuals having approximately the same measure of efficiency, they can be grouped into a personal category.

Let us assume that there are n job categories with b_j number of jobs in the jth category and there are m personal categories with a_i individuals in the ith category.

Let c_{ij} be the measure of efficiency of the ith individual to the jth job.

The general assignment problem can then be formulated as,

Maximize $z = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij}$

Subject to $\sum_{j=1}^{n} x_{ij} = a_i, \quad i = 1, 2, \ldots m.$

$\sum_{i=1}^{n} x_{ij} = b_j, \quad j = 1, 2, \ldots n.$

$x_{ij} \geq 0, \quad i = 1, 2, \ldots m; j = 1, 2, \ldots n.$

where $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$ or $\sum_{i=1}^{m} a_i < \sum_{j=1}^{n} b_j$ or $\sum_{i=1}^{m} a_i > \sum_{j=1}^{n} b_j$

The problem is a transportation problem and since a_i, b_j are integers x_{ij} take on the values zero or positive integer indicating the number of individuals in the ith personal category assigned to the jobs in the jth job category.

If all $a_i = 1$ and all $b_j = 1$, the problem reduces to a simple assignment problem already discussed.

20.5. Some Applications of the Assignment Model

(a) The Marriage Problem: Throughout the ages, our society has been debating the question that of all the possible forms of marriage (monogamy, bigamy, polygamy etc.) which one is the best. It is rather interesting that with the help of linear programming, the issue can now be settled.

Consider n men and n women. Let each man rate the women according to his preference and vice-versa. Let these ratings be denoted by c_{ij}^{M} and c_{ji}^{W} respectively which we assume to be valid measures of happiness.

Define $c_{ij} = \frac{1}{2}\left(c_{ij}^{M} + c_{ji}^{W}\right),$

as the average rating of the pair of the ith man and the jth woman.

Let x_{ij} be the fraction of time that the (i,j) couple spends together.

Since we wish to maximize overall happiness, the problem is

Maximize $z = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij}$

Subject $\sum_{j=1}^{n} x_{ij} = 1, \quad i = 1, 2, \ldots n$

$$\sum_{i=1}^{n} x_{ij} = 1, \qquad j = 1, 2, \ldots n$$

$$x_{ij} \geq 0, \qquad \text{for all } i, j.$$

where $c_{ii} = 0$.

It is an assignment problem and is clear that the optimal values of x_{ij} will either be zero or one which implies that monogamy is an optimal social structure.

It should be noted that if in the above problem c_{ij} is considered to be the measure of average unhappiness of the pair (i, j), it would correspond to minimizing the overall happiness. An optimal solution of the problem will then show that monogamy is a worst form of marriage. Hence monogamy is both a best and a worst social structure. The resulting couples however, are different. This shows that the determination of values of c_{ij} depends on many other factors besides simple preference ratings by couples

(b) Machine Set-up Problem: Suppose that a production company has n machines on which n jobs are to be performed. Each machine can do each of n jobs but it needs an adjustment so as to adopt it to the particular job assigned. The set up time of a machine is different for different job. It depends on what the machine was doing previously; if the previous job is of the same kind, it will not be necessary to reset the machine. The company wants to find an assignment which minimizes the total set up time.

Let c_{ij} be the time needed to set up the ith machine for the jth job. The problem then is,

$$\text{Minimize} \qquad z = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij}$$

$$\text{Subject to} \qquad \sum_{j=1}^{n} x_{ij} = 1, \qquad i = 1, 2, \ldots n$$

$$\sum_{i=1}^{n} x_{ij} = 1, \qquad j = 1, 2, \ldots n$$

$$x_{ij} \geq 0, \qquad \text{for all } i \text{ and } j.$$

Example 1

A company has 4 machines on which 4 jobs are to be performed. Each job can be assigned to one and only one machine. The set up time taken by the machines to do the jobs is given in the following table.

Find the optimal assignment.

		Machine			
		1	2	3	4
	1	2	5	5	4
Job	2	8	6	7	9
	3	4	5	8	7
	4	6	7	6	5

Solution:

Step 1. Subtracting the smallest element in each row from every element in that row, we obtain the reduced matrix

	1	2	3	4
1	0	3	3	2
2	2	0	1	3
3	0	1	4	3
4	1	2	1	0

Step 2. Subtracting the smallest element in each column of the reduced matrix from every element in that column, we get

	1	2	3	4
1	0	3	2	2
2	2	0	0	3
3	0	1	3	3
4	1	2	0	0

The minimum number of lines that cover all zeros is 3 which is less then the order of the cost matrix This implies that an optimal solution has not yet been reached and we proceed to step 3.

Step 3. The smallest element not covered by the lines drawn in step 2 is 1. Subtracting this element from all elements not covered by the lines and adding it to the elements at the intersection of the lines, we obtain the new reduced matrix

$$\begin{array}{c|cccc|}
 & 1 & 2 & 3 & 4 \\
\hline
1 & \boxed{0} & 2 & 1 & 1 \\
2 & 3 & 0 & \boxed{0} & 3 \\
3 & 0 & \boxed{0} & 2 & 2 \\
4 & 2 & 2 & 0 & \boxed{0} \\
\hline
\end{array}$$

The smallest number of lines that cover all the zeros of the reduced matrix is now 4 and 4 independent zeros can be easily determined which provide an optimal solution.

Thus the optimal solution is (1,1), (2,3), (3,2), (4,4) and the minimum set up time is 2 + 7 + 5 + 5 = 19.

(c) Machine Installation problem: A job shop has purchased m new machines and there are n locations available in the shop where a machine could be installed. Some of these locations are more desirable than others for particular machines from the stand point of materials handling. The problem therefore, is to assign the new machines to the available locations so that the total cost of materials handling is minimum.

Example 2

A job shop has purchased 3 machines and there are 4 locations available in the shop for their installation. Selection of locations for installation of the machines are important from the standpoint of materials handling. The estimated cost per unit time of materials handling is given in the following table.

Find an optimal assignments of the machines to the locations

		Location			
		1	2	3	4
	A	18	14	16	15
Machine	B	19	20	15	17
	C	10	X	12	11

Location 2 is not suitable for machine c.

Solution: Since the number of machines is less than the number of locations, a fictitious machine is introduced in order to formulate the problem as an assignment problem and the cost of materials handling for this machine at any of the four location is considered to be zero. Also, an extremely large cost M is attached to the assignment of machine C to location 2, so that this assignment does not appear in the optimal solution. The resulting cost matrix then is

		Location 1	2	3	4
	A	18	14	16	15
	B	19	20	15	17
Machine	C	10	M	12	11
	D	0	0	0	0

Now, subtracting the smallest element in each row from all the elements in that row, we get the reduced matrix

	Location 1	2	3	4
A	4	$\boxed{0}$	2	1
B	4	5	$\boxed{0}$	2
C	$\boxed{0}$	M	2	1
D	0	0	0	$\boxed{0}$

As indicated in the table above, the minimum number of lines covering all the zeros is 4 and therefore yields an optimal solution. The independent zero elements are enclosed and the optimal assignment is given by

$$A \to 2, B \to 3, C \to 1 \text{ and } D \to 4.$$

The total cost of materials handling is $14 + 15 + 10 + 0 = 39$

Example 3. Marketing Problem: A marketing manager has four territories open and four salesmen available for assignment. The territories differ in their sales potential and the salesmen differ in their ability. The estimates of normal sales by the salesmen in the four territories are given below.

Territory

		1	2	3	4
	A	40	33	30	23
	B	25	27	22	17
Salesman	C	28	25	22	16
	D	24	30	18	20

Find the assignment of salesmen to the territories in order to have maximum sales.

Solution: The problem of maximization is converted to a minimization problem by multiplying each element of the sales matrix by –1 and then subtracting the smallest element in each row from all the elements in that row. The efficiency matrix for the minimization problem then becomes

	1	2	3	4
A	0	7	10	17
B	2	0	5	10
C	0	3	6	12
D	6	0	12	10

Now, subtracting the smallest element in each column of the reduced matrix from all the elements of that column, we obtain

	1	2	3	4
A	0	7	5	7
B	2	0	0	0
C	0	3	1	2
D	6	0	7	0

The minimum number of lines that cover all the zeros is 3, which is less than the order of the matrix. We therefore, find the smallest element in the matrix, not covered by the lines (which is 1). Subtracting this element from all the elements not covered by the lines and adding it to the elements at the intersection of the lines, we obtain the new reduced matrix

$$\begin{array}{c|cccc|}
 & 1 & 2 & 3 & 4 \\
\hline
A & \boxed{0} & 6 & 4 & 6 \\
B & 3 & \boxed{0} & 0 & 0 \\
C & 0 & 2 & \boxed{0} & 1 \\
D & 7 & 0 & 7 & \boxed{0} \\
\hline
\end{array}$$

The minimum number of lines covering all the zeros now is 4 which implies that an optimal solution has been reached. The independent points are enclosed which provide an optimal solution

$$(A \to 1), (B \to 2), (C \to 3) \text{ and } (D \to 4).$$

An alternative optimal solution exists for this problem

$$(A \to 1), (B \to 4), (C \to 3) \text{ and } (D \to 2).$$

The optimal sale is 109.

Example 4

A firm has five garages in each of which is stationed 5 trucks which are of the same type. The trucks have to go on hire to five different places one to each. The following table gives the distances between the garages and the destinations. How should the trucks be dispatched so as to minimize the total distance travelled?

	Place P_1	P_2	P_3	P_4	P_5
A	20	20	16	18	20
B	18	15	14	10	16
Garage C	14	18	17	12	14
D	19	22	15	24	22
E	14	10	16	11	13

Solution: Subtracting the smallest element in each row of the distance matrix from all the elements in that row and then subtracting the smallest element in each column of the reduced matrix from all the elements in that column, we get the matrix reduced to

		P_1	P_2	P_3	P_4	P_5
	A	2	4	0	2	2
	B	6	5	4	0	4
Garage	C	0	6	5	0	0
	D	2	7	0	9	5
	E	2	0	6	1	1

As indicated in the table above, the minimum number of lines covering all the zeros is 4, and hence this reduced matrix does not contain 5 independent zero elements. We then find the smallest element not covered by the lines. This element is 1 in our case. Subtracting this element from all the elements not covered by the lines and adding it to the elements at the intersections of the lines, we get the new reduced matrix

	P_1	P_2	P_3	P_4	P_5
A	1	4	0	2	1
B	5	5	4	0	3
C	0	7	6	1	0
D	1	7	0	9	4
E	1	0	6	1	0

Again we find that the minimum number of lines covering all the zeros is 4. The above matrix therefore has to be subjected to further transformation. We find the smallest element not covered by the lines which is 1. We subtract this element to all the elements not covered by the lines and adding to the elements at the intersections, we get the matrix:

	P_1	P_2	P_3	P_4	P_5
A	$\boxed{0}$	3	0	1	0
B	5	5	5	$\boxed{0}$	3
C	0	7	7	1	$\boxed{0}$
D	0	6	$\boxed{0}$	8	3
E	1	$\boxed{0}$	7	1	0

The minimum number of lines covering all the zeros is now 5 and the optimal solution has been reached. The optimal assignment is

$$(A \to P_1), (B \to P_4), (C \to P_5), (D \to P_3), (E \to P_2).$$

and the minimum distance travelled is 69.

Note that alternative optimal solution is also available.

20.6. Exercises

1. Solve the following assignment problems.

(a)

$$\text{Man} \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \overset{\displaystyle\text{Job}}{\begin{array}{cccc} \text{I} & \text{II} & \text{III} & \text{IV} \\ \end{array}} $$

$$\begin{array}{cc} & \begin{array}{cccc} \text{I} & \text{II} & \text{III} & \text{IV} \end{array} \\ \text{Man} \begin{array}{c} A \\ B \\ C \\ D \end{array} & \begin{bmatrix} 19 & 12 & 11 & 10 \\ 7 & 10 & 8 & 5 \\ 13 & 14 & 11 & 12 \\ 11 & 15 & 9 & 8 \end{bmatrix} \end{array}$$

(a)

$$\begin{array}{cc} & \begin{array}{cccc} \text{I} & \text{II} & \text{III} & \text{IV} \end{array} \\ \text{Man} \begin{array}{c} A \\ B \\ C \\ D \end{array} & \begin{bmatrix} 3 & 6 & 4 & 1 \\ 9 & 10 & 7 & 9 \\ 7 & 11 & 5 & 4 \\ 5 & 8 & 7 & 8 \end{bmatrix} \end{array}$$

2. There are four engineers available for designing four projects. Engineer E_3 is not competent to design project P_2. Given the time estimate required by each engineer to design a given project in the table below, find an assignment which minimizes the total time.

		Project			
		P_1	P_2	P_3	P_4
	E_1	10	2	3	7
	E_2	9	1	8	6
Engineer	E_3	8	X	10	3
	E_4	5	1	8	6

3. A marketing manager has five territories open and five salesmen available for assignment. Considering the capabilities of the salesmen, the marketing manager estimates, sales by the salesmen in each territory as given in the table below:

		Territory				
		T_1	T_2	T_3	T_4	T_5
	A	35	39	33	40	29
	B	36	36	38	41	22
	C	28	40	38	40	32
Salesman	D	21	36	24	28	40
	E	30	37	27	33	41

Find an assignment which maximizes the total sales.

4. A manufacturing company has four jobs to be done and three workers are available. The time taken by each worker to complete the job is given below.

		Job			
		J_1	J_2	J_3	J_4
	W_1	1	5	3	3
Worker	W_2	5	3	4	2
	W_3	4	2	1	3

Which worker should be assigned to which job? Which job will remain unfinished?

5. The captain of a cricket team has to allot five batting positions to six batsmen. The average runs scored by each batsman at these positions are as follows:

		Batting position				
		P_1	P_2	P_3	P_4	P_5
	B_1	50	48	40	60	50
	B_2	45	52	38	50	49
Batsman	B_3	58	60	59	55	53
	B_4	20	19	20	18	25
	B_5	42	30	16	25	27
	B_6	40	40	35	25	50

Find an assignment of batsmen to batting position which would give the maximum number of runs. Which of the batsmen is to be dropped?

CHAPTER 21

The Decomposition Principle for Linear Programs

21.1. Introduction

The solution of linear programming problems with large number of constraints by means of the simplex method usually requires considerable computational time. In many cases however, the matrix of the coefficients has a special structure, which can be utilized to improve the computational efficiency of the problem. For example, such a situation may arise in a big business enterprise with many branches. Suppose that the branches are almost independent in the sense that the activities of one branch are not affected by the activities of the other branches but there are a few constraints and a common objective that tie them together.

Several methods for handling such linear programs efficiently have been suggested by various authors [119, 47, 40, 385, 296 and others].

In this chapter we shall describe only the decomposition algorithm of Dantzig and Wolfe [112].

If the coefficient matrix of the problem is black angular, the set of constraints may be partitioned into (k + 1) subsets such that k subsets are mutually independent systems, each one containing different unknowns and the constraints of the $(k + 1)^{th}$ subset are the linking constraints involving all the variables.

The linear programs over the sets of independent constraints are called the subprograms and the solutions of the subprograms enable us to obtain a solution of the original problem efficiently.

21.2. The Original Problem and its Equivalent

Consider the linear programming problem,

$$\begin{aligned} &\text{Minimize} && z = c^T X \\ &\text{Subject to} && AX = b \\ & && X \geq 0. \end{aligned} \tag{21.1}$$

where the $m \times n$ coefficient matrix A is of the form

$$A = \begin{bmatrix} P_1 & P_2 \dots & P_k \\ A_1 & 0 & 0 \\ 0 & A_2 \dots & 0 \\ & & \\ 0 & 0 & A_k \end{bmatrix} \tag{21.2}$$

where P_j is $m_0 \times n_j$ matrix, A_j, an $m_j \times n_j$ matrix,

$$\sum_{j=0}^{k} m_j = m, \quad \text{the number of constaints and}$$

$$\sum_{j=1}^{k} n_j = n, \quad \text{the number of variables.}$$

The matrix of the form (21.2) is called a block-angular matrix.

The problem (21.1) can then be expressed in the form,

$$\text{Minimize} \quad z = \sum_{j=1}^{k} c_j^T X_j \tag{21.3}$$

$$\text{Subject to} \quad \sum_{j=1}^{k} P_j X_j = b_0 \tag{21.4}$$

$$A_j X_j = b_j, \qquad j = 1, 2, \dots k. \tag{21.5}$$

$$X_j \geq 0, \qquad j = 1, 2, \dots k. \tag{21.6}$$

where the vectors c, b and X are also partitioned in the same manner as done in A. Thus,

$$c^T = \left(c_1^T, c_2^T, \dots c_k^T\right),$$

$$b^T = \left(b_0^T, b_1^T, \dots b_k^T\right), \text{ and}$$

$$X^T = \left(X_1^T, X_2^T, \dots X_k^T\right).$$

Let us now suppose that the polytopes S_j defined by the constraint sets

$$A_j X_j = b_j, \quad X_j \geq 0 \qquad j = 1, 2 \dots k. \tag{21.7}$$

are bounded and hence are convex polyhedra.

Each S_j, therefore has a finite number of extreme points say X_{ji} ($i = 1,2..t_j$).

Every point X_j of S_j can therefore be written as

$$X_j = \sum_{i=1}^{t_j} \lambda_{ji} X_{ji}, \quad \sum_{i=1}^{t_j} \lambda_{ji} = 1, \quad \lambda_{ji} \geq 0, \qquad i = 1, 2 \dots t_j \tag{21.8}$$

Suppose that all the extreme points X_{ji} are known and let

$$p_{ji} = P_j X_{ji} \qquad i = 1,2 \ldots t_j \tag{21.9}$$

$$c_{ji} = c_j^T X_{ji} \tag{21.10}$$

Now, when X_j in (21.3) and (21.4) are replaced by (21.8), the problem is reduced to

$$\text{Minimize} \qquad z = \sum_{j=1}^{k} \sum_{i=1}^{t_j} c_{ji} \lambda_{ji}$$

$$\text{Subject to} \qquad \sum_{j=1}^{k} \sum_{i=1}^{t_j} p_{ji} \lambda_{ji} = b_0.$$

$$\sum_{i=1}^{t_j} \lambda_{ji} = 1, \qquad j = 1,2 \ldots k \tag{21.11}$$

$$\lambda_{ji} \geq 0, \qquad i = 1,2 \ldots t_j;\ j = 1,2 \ldots k.$$

The linear program (21.11) in the unknowns λ_{ji} equivalent to the original problem and is called the master program. Indeed, corresponding to every feasible solution to (21.1), there is a feasible solution to (21.11) and conversely.

The decomposition principle is thus based on representing the original problem by a linear program in terms of the extreme points of the sub programs. Once an optimal solution of the master program is obtained, we can find an optimal solution of the original problem from (21.8). It should be noted that the procedure reduces the number of constraints from $\sum_{j=0}^{k} m_j$. to $(m_0 + k)$ but there is an increase in the number of variables from $\sum_{j=0}^{k} n_j$ to $\sum_{j=0}^{k} t_j$. Since the computational effort in any linear program mainly depends on the number of constraints rather than on the number of variables, it may be considered advantageous to apply the decomposition method.

It may appear that the solution of the master program requires prior determination of all the extreme points X_{ji} which may be numerous and hence to enumerate all the extreme points explicitly is a very difficult task. However, this is not the case. In fact, to determine an initial basic feasible solution for the master program, only one extreme point of each of the subprograms is needed and the other extreme points required to find an optimal solution of the problem are generated at the succeeding iterations of the computing process.

21.3 The Decomposition Algorithm

Let us now describe the algorithm for the solution of the master program (21.11) The algorithm is based on the revised simplex method discussed in Chapter 14.

Let B be the basis matrix associated with the current basic feasible solution of (21.11) and c_B be the corresponding cost vector. We use the revised simplex method, so that the dual vector $c_B^T B^{-1}$ of $(m_o + k)$ components is known.

Let $c_B^T B^{-1} = V = (V_1, V_2)$, where V_1 is formed from the first m_0 components and V_2 is formed from the last k components of V.

Then
$$z_{ji} - c_{ji} = c_B^T B^{-1} \begin{pmatrix} p_{ji} \\ e_j \end{pmatrix} - c_{ji}$$

$$= V_1 p_{ji} + v_{2j} - c_{ji} \tag{21.12}$$

where e_j is the jth unit rector,

By (21.9) and (21.10), we have

$$z_{ji} - c_{ji} = (V_1 P_j - c_B^T) X_{ji} + v_{2j}. \tag{21.13}$$

If $z_{jj} - c_{ji} \leq 0$, for all (j, i), the current solution is optimal.

If $z_{ji} - c_{ji} > 0$, for at least one (j, i), the solution is not optimal and the basis is changed by replacing a vector of B by one of the vectors

$$\begin{pmatrix} p_{ji} \\ e_j \end{pmatrix}$$

for which $z_{ji} - c_{ji} > 0$.

Thus to determine whether the current solution is optimal or not, we are to compute

$$\underset{j,i}{\text{Max}}(z_{ji} - c_{ji}) = \underset{j}{\text{Max}}\left[\underset{i}{\text{Max}}(z_{ji} - c_{ji})\right] \tag{21.14}$$

Now, this computation depends on the values of the extreme points of the set $A_j X_j = b_j$, $X_j \geq 0$. Since by assumption, this set is bounded, it follows that, for a given j, $\underset{i}{\text{Max}}\,(Z_{ji} - C_{ji})$ is also bounded and must occur at an extreme point of the set. We therefore solve k linear programming problems,

$$\begin{aligned} &\text{Maximize} && z_j = (V_1 P_j - c_j^T) X_j && (21.15) \\ &\text{Subject to} && A_j X_j = b_j && \\ & && X_j \geq 0 \qquad \text{for } j = 1, 2, \ldots k. && \end{aligned}$$

Let X_{jr_j} be an optimal solution for the jth such problem and the corresponding optimal value of z_j, is

$$z_j^* = \left(V_1 P_j - c_j^T\right) X_{jr_j}$$

We then have,

$$\max_{j,i}\left(z_{ji} - c_{ji}\right) = \max_{j}\left(z_j^* + v_{2j}\right). \tag{21.16}$$

Let this maximum be achieved for j = s and the corresponding optimal solution of (21.15) be X_{sr_s} or X_{sr} (where r_s = r say).

Thus $\quad \max_{j,i} (z_{ji} - c_{ji}) = z_j^* + v_{2s} = z_{sr} - c_{sr}$ (21.17)

If $z_{sr} - c_{sr} \leq 0$, the current solution is optimal for the master program and an optimal solution of the original problem is then obtained using the relation (21.8)

If $z_{sr} - c_{sr} > 0$, the vector

$$\begin{pmatrix} p_{sr} \\ e_s \end{pmatrix}$$

is introduced in the basis and the vector to leave the basis is determined by the usual simplex procedure. The process is then repeated.

21.4. Initial Basic Feasible Solution

In the foregoing discussions, we considered the computational procedure involving all the extreme points of each of the subprograms. However, to obtain an initial basic feasible solutions to the master program, we need to know only one basic feasible solution for each of the k subprograms. These solutions (if exist) are then applied to the master program and the problem is solved by the two-phase method after adding artificial variables. The solution in Phase I provides a basic feasible solution to the master program or indicates that the problem is infeasible.

Suppose that a basic feasible solution X_{j1} to the subprogram j = 1,2,...k. is known (The use of artificial variables may be necessary to obtain such a solution)

Let $\quad p'_{ji} = p_J x_{j1}, \qquad j = 1,2\ldots k.$

The Phase I problem for (21.11), then becomes

Minimize $F(W) = w_1 + w_2 + \ldots + w_{m_0}$

Subject to $\quad p'_{11}\lambda_{11} + p'_{21}\lambda_{21} + \ldots + p'_{k1}\lambda_{k1} \pm e_1 w_1 \pm e_2 w_2 \pm \ldots \pm e_{m_0} w_{m_0} = b_0$

$\lambda_{j1} = 1, j = 1,2,\ldots.k$

$\lambda_{j1} \geq 0, j = 1,2,\ldots.k$

$w_i \geq 0, I = 1,2,\ldots.m_0$

where w_i is the ith artificial variable and e_i is an m_0-component unit vector whose ith component is one. The sign preceeding $e_i w_i$ is so determined that w_i would be nonnegative.

At the end of phase I, a basic feasible solution to the master program is obtained if Min F(W) = 0. We then apply Phase II, to obtain an optimal solution of the given problem. If any of the k subprograms is not feasible or Min F(w) ≠ 0, the given problem has no feasible solution.

21.5. The Case of Unbounded S_j

If the set S_j is unbounded, the decomposition algorithm has to be slightly modified. In this case, $X_j \in S_j$, if and only if

$$X_j = \sum_{i=1}^{t_j} \lambda_{ji} X_{ji} + \sum_{i=1}^{l_j} \mu_{ji} d_{ji} \tag{21.18}$$

$$\sum_{i=1}^{t_j} \lambda_{ji} = 1$$

$$\lambda_{ji} \geq 0, \qquad i = 1, 2 \ldots t_j.$$

$$\mu_{ji} \geq 0, \qquad i = 1, 2 \ldots l_j.$$

where X_{ji} and d_{ji} are the extreme points and the extreme directions of S_j (See Chapter 8).

Replacing each X_j in (21.3) – (21.6) by the above representation, the original problem can be written as follows.

$$\text{Minimize} \quad z = \sum_{j=1}^{k} \sum_{i=1}^{t_j} \left(c_j^T X_{ji}\right) \lambda_{ji} + \sum_{j=1}^{k} \sum_{i=1}^{l_j} \left(c_j^T d_{ji}\right) \mu_{ji}$$

$$\text{Subject to} \quad \sum_{j=1}^{k} \sum_{i=1}^{t_j} \left(P_j X_{ji}\right) \lambda_{ji} + \sum_{j=1}^{k} \sum_{i=1}^{l_j} \left(P_j d_{ji}\right) \mu_{ji} = b_0 \tag{21.19}$$

$$\sum_{i=1}^{t_j} \lambda_{ji} = 1, \qquad j = 1, 2, \ldots k \tag{21.20}$$

$$\lambda_{ji} \geq 0, \qquad i = 1, 2, \ldots t_j; \; j = 1, 2, \ldots k$$

$$\mu_{ji} \geq 0, \qquad i = 1, 2, \ldots l_j; \; j = 1, 2, \ldots k \tag{21.21}$$

Suppose that a basic feasible solution of the above problem is known with an $(m_0 + k) \times (m_0 + k)$ basis matrix B. It should be noted that each basis must contain at least one variable λ_{ji}; from each block j.

Let c_B be the cost vector of the basic variables with $c_{ji} = c_j^T X_{ji}$ for λ_{ji} and $c_{ji} = c_j^T d_{ji}$ for μ_{ji}.

Further, let π and α be the vectors of dual variables corresponding to the constraints (21.19) and (21.20) respectively. So that $(\pi^T \alpha^T) = c_B^T B^{-1}$

The current basic feasible solution is optimal if

$$z_{ji} - c_{ji} = \pi^T (P_j X_{ji}) + \alpha_j - c_j^T X_{ji} \leq 0 \text{ for } \lambda_{ji} \text{ nonbasic} \tag{21.22}$$

$$z_{ji} - c_{ji} = \pi^T (P_j d_{ji}) - c_j^T d_{ji} \leq 0 \text{ for } \mu_{ji} \text{ nonbasic.} \tag{21.23}$$

To verify whether the optimality conditions hold or not, we now solve the k subprograms

$$\text{Maximize } z_j = (\pi^T P_j - c_j^T) X_j \tag{21.24}$$

$$\text{Subject to } X_j \in S_j \qquad j = 1, 2, \ldots k.$$

If the solution of the jth subprogram is unbounded, then an extreme direction d_{jr} is found such that $(\pi^T P_j - c_j^T) d_{jr} > 0$, which means that the condition (21.23) is violated and μ_{jr} is eligible to enter the basis. Since $z_{jr} - c_{jr} > 0$, the insertion of μ_{jr} in the basis improves the value of the objective function.

If a finite optimal solution to the jth subproblem is found, then obviously the condition (21.23) holds for the jth subproblem. Let X_{jr} be an optimal extreme point of the jth subproblem. Now, if the optimal value of the objective function $z_j^0 = (\pi^T P_j - c_j^T) X_{jr}$ is such that $z_j^0 + \alpha_j \leq 0$, then the conditions (21.22) holds for the subproblem J. Otherwise λ_{jr} can be introduced in the basis. When for each subprogram in (21.24) $z_j^0 + \alpha_j \leq 0$, then the optimal solution to the original problem is obtained.

Thus, according to the above procedure, each subprogram j (j = 1,2..k) is solved in turn. If the subprogram j yields an unbounded solution, then an extreme direction d_{jr} is found and the vector

$$\begin{pmatrix} P_j d_{jr} \\ 0 \end{pmatrix}$$

is introduced into the basis of the problem (21.19) – (21.21). If the subprogram j yields a bounded optimal solution X_{jr} and $z_j^0 + \alpha_j = (\pi^T P_j - c_j^T) X_{jr} + \alpha_j > 0$, then the vector

$$\begin{pmatrix} P_j X_{jr} \\ e_j \end{pmatrix}$$

is introduced in the basis matrix. If neither of these conditions hold, then there is no vector to enter the basis from the subprogram j. If none of the k subprograms yields a vector to enter the basis, then an optimal solution of the original problem is obtained. Otherwise, we select one from the eligible vectors to enter the basis. We then update the entering column by premultiplying it by the current B^{-1}, pivot on the master array and repeat the process.

21.6. Remarks on Methods of Decomposition

It should be noted that in solving a structured linear programming problem by means of the decomposition algorithm, the problem can be decomposed in various ways. Instead of considering each of the subprograms $A_j X_j = b_j$ in (21.5) separately, they can be grouped into a single problem or grouped in parts giving several linear programs. However, there are advantages and disadvantages of such various types of decomposition in solving the original linear program.

The decomposition principle can also be applied to problems where the constraints form a staircase structure (multistage) such as

$$\begin{array}{lllll}
\text{Maximize} & x_0 & & & \\
\text{Subject to} & A_1X_1 & & & = b_1 \\
& \bar{A}_1 X_1 + A_2X_2 & & & = b_2 \\
& & \bar{A}_2 X_2 + A_3X_3 & & = b_3 \\
& & & \bar{A}_3 X_3 + A_4X_4 + P_0x_0 & = b_4 \\
& & & X_t \geq 0,\ t = 1, 2, 3, 4, &
\end{array}$$

where $\bar{A}_i$ $(i = 1, 2, 3)$, A_t are matrices, X_t, b_t $(t = 1, 2, 3, 4)$ and P_0 are vectors.

Such problems often arise in the study of processes through time in which the activities of one period are directly connected with those of the preceeding and following periods but with no others. (See section 23-4 in [109])

These observations show that the decomposition method is extremely flexible. In fact, application of the decomposition principle is not limited to problems of special structure. Actually, any linear programming problem can be decomposed by splitting up the set of constraints into two subproblems. An application of this variant of decomposition to a multi-commodity transportation problem furnishes a particularly efficient algorithm.

21.7. Example

$$\begin{array}{lll}
\text{Minimize} & z = -5x_1 - 3x_2 - 5x_3 - 2.5x_4 & \\
\text{Subject to} & 2x_1 + 3x_2 + 3x_3 + 2x_4 & \leq 15. \\
& x_1 + x_2 & \leq 4 \\
& 2x_1 + x_2 & \leq 6 \\
& \qquad 2x_3 + 5x_4 & \leq 10 \\
& \qquad x_3 & \leq 4, \\
& x_1, x_2, x_3, x_4 \geq 0 &
\end{array}$$

In terms of notations in section 21.2

$$P_1 = (2, 3), \quad P_2 = (3,2), \quad A_1 = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & 5 \\ 1 & 0 \end{pmatrix}$$

$$c_1^T = (-5, -3), \quad c_2^T = (-5, -2.5), \quad b_0 = 15,$$

$$b_1^T = (4, 6), \quad b_2^T = (10, 4) \text{ and } X_1^T = (x_1, x_2), \quad X_2^T = (x_3, x_4)$$

The convex polyhedron S_1 is given by the constraints

$$x_1 + x_2 \le 4$$
$$2x_1 + x_2 \le 6$$
$$x_1, x_2 \ge 0$$

and the convex polyhedron S_2 by the constraints

$$2x_3 + 5x_4 \le 10$$
$$x_3 \quad \le 4$$
$$x_3, x_4 \ge 0$$

It can be noted that S_1 and S_2 each has four extreme points X_{ji}, j = 1, 2; i = 1, 2, 3, 4, which can be easily obtained from their graphs.

The extreme points of S_1 are

$$X_{11} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad X_{12} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \quad X_{13} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad X_{14} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$$

and the extreme points of S_2 are

$$X_{21} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad X_{22} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \quad X_{23} = \begin{pmatrix} 4 \\ 2/5 \end{pmatrix}, \quad X_{24} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

To construct the master program, we note that

$$p_{11} = P_1X_{11} = (2,3) \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0, \qquad p_{21} = P_2X_{21} = (3,2) \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0$$

$$p_{12} = P_1X_{12} = (2,3) \begin{pmatrix} 3 \\ 0 \end{pmatrix} = 6, \qquad p_{22} = P_2X_{22} = (3,2) \begin{pmatrix} 4 \\ 0 \end{pmatrix} = 12$$

$$p_{13} = P_1X_{13} = (2,3) \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 10, \qquad p_{23} = P_2X_{23} = (3,2) \begin{pmatrix} 4 \\ 2/5 \end{pmatrix} = \frac{64}{5}$$

$$p_{14} = P_1X_{14} = (2,3) \begin{pmatrix} 0 \\ 4 \end{pmatrix} = 12, \qquad p_{24} = P_2X_{24} = (3,2) \begin{pmatrix} 0 \\ 2 \end{pmatrix} = 4$$

and from $C_{ji} = X_{ji}$ we have

$$c_{11} \, (-5, -3) \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0, \qquad c_{21} \, (-5, -2.5) \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0,$$

$$c_{12}(-5,-3)\begin{pmatrix}3\\0\end{pmatrix}=-15, \qquad c_{22}(-5,-2.5)\begin{pmatrix}4\\0\end{pmatrix}=-20,$$

$$c_{13}(-5,-3)\begin{pmatrix}2\\2\end{pmatrix}=-16, \qquad c_{23}(-5,-2.5)\begin{pmatrix}4\\2/5\end{pmatrix}=-21,$$

$$c_{14}(-5,-3)\begin{pmatrix}0\\4\end{pmatrix}=-12, \qquad c_{24}(-5,-2.5)\begin{pmatrix}0\\2\end{pmatrix}=-5.$$

Now, $\quad X_j=\sum_{i=1}^{4}\lambda_{ji}X_{ji}, \qquad j=1,2$

$$\sum_{i=1}^{4}\lambda_{ji}=1,, \qquad j=1,2$$

$$\lambda_{ji}\geq 0, \qquad i=1,2,3,4;\ j=1,2.$$

and the master program is given by

Minimize $z= -15\lambda_{12}-16\lambda_{13}-12\lambda_{14}-20\lambda_{22}-21\lambda_{23}-5\lambda_{24}.$

Subject to $\quad 6\lambda_{12}+10\lambda_{13}+12\lambda_{14}+12\lambda_{22}+64/5\lambda_{23}+4\lambda_{24}\leq 15$

$\lambda_{11}+\lambda_{12}+\lambda_{13}+\lambda_{14}=1$

$\lambda_{21}+\lambda_{22}+\lambda_{23}+\lambda_{24}=1$

$\lambda_{j,i}\geq 0$ for all i and j.

Solving the above linear programming problem, we have the optimal solution

$$\lambda_{12}=1,\ \lambda_{21}=1/4,\ \lambda_{22}=3/4,\ \text{all other } \lambda_{ji}=0.$$

The optimal solution to the original problem is then given by

$$X_1=\sum_{i=1}^{4}\lambda_{1i}X_{1i}=\lambda_{11}X_{11}+\lambda_{12}X_{12}+\lambda_{13}X_{13}+\lambda_{14}X_{14}$$

$$=\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}3\\0\end{pmatrix}$$

$$X_2=\sum_{i=1}^{4}\lambda_{2i}X_{2i}=\lambda_{21}X_{21}+\lambda_{22}X_{22}+\lambda_{23}X_{23}+\lambda_{24}X_{24}$$

$$=\begin{pmatrix}x_3\\x_4\end{pmatrix}=\begin{pmatrix}3\\0\end{pmatrix}$$

Hence, the optimal solution of the original problem is

$$x_1=3,\ x_2=0,\ x_3=3,\ x_4=0$$

and Min $z=-30$.

The above problem is solved by using all the extreme points of S_1 and S_2. In general, it is difficult to determine all the extreme points of each S_j. The problem can however be solved when only one extreme point of each S_j is known (See section 21.4)

It is left to the reader to solve the above given problem with only one extreme point of each S_1 and S_2.

21.8. Exercises

1. Solve the following linear programming problem by using the decomposition technique.

(i) Maximize $z = 8x_1 + 7x_2 + 6x_3 + 5x_4$

Subject to

$$\begin{aligned}
&2x_1 + x_2 + x_3 + 3x_4 \le 13\\
&x_1 + x_2 + 2x_3 + x_4 \le 10\\
&2x_1 + x_2 \le 11\\
&x_1 + x_2 \le 8\\
&x_3 + 2x_4 \le 9\\
&x_3 + x_4 \le 6\\
&x_1\ x_2, x_3, x_4 \ge 0.
\end{aligned}$$

(ii) Maximize $z = 6x_1 + 8x_2 + 3x_3 + 8x_4$

Subject to

$$\begin{aligned}
&3x_2 - x_3 + 4x_4 \le 12\\
&3x_1 + x_2 + 3x_3 + 4x_4 \le 16\\
&x_1 + 4x_2 \le 12\\
&3x_1 + 2x_2 \le 9\\
&x_3 + 3x_4 \le 10\\
&2x_3 + x_4 \le 8\\
&x_1, x_2, x_3, x_4 \ge 0.
\end{aligned}$$

(iii) Minimize $z = x_1 - x_2 - x_3 - 2x_4$

Subject to

$$\begin{aligned}
&2x_1 + x_3 + x_4 \le 3\\
&x_2 + x_4 \le 2\\
&x_1 + 2x_2 \le 6\\
&x_1 - x_2 \le 2\\
&2x_3 + x_4 \le 5\\
&x_4 \le 2\\
&x_1, x_2, x_3, x_4 \ge 0
\end{aligned}$$

2. Show how the decomposition technique can be applied to solve a problem of the following structure

$$\begin{array}{lll} \text{Minimize} & C^T_0 X_0 + C^T_1 X_1 + .. + C^T_k X_k & \\ \text{Subject to} & B_0 X_0 + P_1 X_1 + ... + P_k X_k & = b_0 \\ & B_1 X_0 + A_1 X_1 & = b_1 \\ & B_2 X_0 + A_2 X_2 & = b_2 \\ & \vdots & \\ & B_k X_0 + B_k X_k & = b_k \\ & \quad X_j \geq 0, \quad j = 0, 1,..k. & \end{array}$$

[Hint: Take the problem consisting of the last k constraints as the subproblem. Solve its dual by the decomposition algorithm.]

3. A linear programming problem can be decomposed in various ways. Discuss their relative advantages and disadvantages, when the decomposition algorithm is applied to solve the problem.

4. Apply the decomposition algorithm to solve the following problem.

$$\begin{array}{ll} \text{Minimize} & z = 10x_1 + 2x_2 + 4x_3 + 8x_4 + x_5 \\ \text{Subject to} & 2x_1 + x_2 + x_3 \geq 2 \\ & x_1 + 4x_2 - x_3 \geq 8 \\ & x_1 + 2x_4 - x_5 \geq 10 \\ & 3x_1 + x_4 + x_5 \geq 4 \\ & x_1, x_2, x_3, x_4, x_5 \geq 0. \end{array}$$

CHAPTER 22

Polynomial Time Algorithms for Linear Programming

22.1. Introduction

The development of the simplex method by George B Dantzig in the mid-1940s for solving linear programming problems is a great achievement in the theory of optimization. An optimal solution of a linear programming problem always lies at a vertex of the feasible region which is a polyhedron. The simplex method moves along the edges of the polyhedron from one vertex to the next in an orderly fashion until it arrives at an optimal vertex. Since the number of vertices of the polyhedron associated with the m x n coefficient matrix of the problem, is quite large for large values of m and n, there was apprehension that the simplex method would not prove to be efficient. However, in practice, it was found to perform exceedingly well. It was observed that for most of the practical problems, the number of iterations the method needs, is only between m and 3m. However, the fact remains that for certain problems, the simplex method may require to examine all the vertices to arrive at an optimal solution and therefore the number of iterations can grow exponentially An example given by Klee and Minty [275] indeed shows that in the worst case, the simplex method needs 2^n iterations to find an optimal solution.

The search for methods with better complexity (algorithm efficiency) than the simplex method led to the development of polynomial time algorithms, that is, where the computation time is bounded above by a polynomial in the size or the total data length of the problem.

22.2. Computational Complexity of Linear Programs

We now discuss the computational complexity (algorithm efficiency) of linear programming problems, that is, we determine the growth in computational effort of an algorithm as a function of the size of the problem in the worst case.

Consider the linear programming problem

$$\begin{aligned} &\text{Minimize} && c^T x \\ &\text{Subject to} && AX = b \\ &&& X \geq 0 \end{aligned} \tag{22.1}$$

where A is $m \times n$ and $m \leq n$, $n \geq 2$ and all data are integers. If the data are rational, they can be easily converted to integers by multiplying them by the least common denominator.

By an instance of the problem we mean a linear programming problem with specific values of m, n, c, A and b and the size of an instance of the problem is represented by (m, n, L), where L is the number of bits required to store the problem data in a computer and known as the input length of an instance of the problem.

Let f(m, n, L) be a function in terms of the size of the problem such that the total number of computational operations required by the algorithm to solve the problem is bounded by pf(m, n, L), where $p > 0$. We then say that the order of complexity of the algorithm is 0 (f(m, n, L)). If f(m, n, L) is a polynomial in m, n and L, the algorithm is said to be a polynomial algorithm and the problem is polynormally solvable.

Thus if the bound on the number of operations in an algorithm is αL^{β} for some $\alpha > 0$, $\beta > 0$, it is a polynomial algorithm and if the bound is $\gamma 2^{L}$, for $\gamma > 0$, we say that it is an exponential algorithm. Hence the simplex method is not a polynomial algorithm. Since for large L, $\gamma 2^{L} > \alpha L^{\beta}$, in theory, a polynomial algorithm is more efficient than an exponential algorithm.

On the other hand, if f(m, n, L) is independent of L and polynomial in m and n, then we say that the algorithm is strongly polynomial. No strongly polynomial algorithm is known to exist for general linear programming problems but it does exist for combinatorial linear programs in network flow problems. (see [460]).

22.3. Khachiyan's Ellipsoid Method

The first polynomial time algorithm for linear programming was given by Khachiyan in 1979 [274]. Khachiyan showed how the ellipsoid method developed during the 1970s by Russian mathematicians could be adapted to give a polynomial time algorithm for linear programming. There are several variants of Khachiyan's algorithm, we follow the interpretation given by Gacs and Lövask [169].

Consider the problem of determining a solution to the set of linear inequalities

$$S = \{X \mid AX \leq b\} \tag{22.2}$$

where A is $m \times n$, b is an m-vector; m, $n \geq 2$ and the data are all integers.

Assuming that S is nonempty, algorithm starts by constructing an appropriate ball E_0 centered at the origin to contain a large portion of S. If the center of the ball lies in S, then the algorithm terminates. Otherwise, a sequence of ellipsoids $E_1, E_2, E_{k\ldots\ldots}$ in decreasing volume are constructed each of which containing the region of S covered by E_0. If at the kth iteration, (k = 1, 2...), the center X_k of E_k lies in S, the algorithm terminates. If not, some constraints are violated. Let the most violated constraint be

$$A_v X \leq b_v \tag{22.3}$$

Hence

$$A_v X_k > b_v \tag{22.4}$$

A new ellipsoid E_{k+1} of smaller volume is then constructed which contains that region of S which was covered by E_k. If the center of E_{k+1} lies in S, the algorithm terminates. Otherwise the above steps are repeated. Khachiyan has shown that if S is nonempty, one can determine a feasible point within a number of iterations bounded above by a polynomial in the size of the problem. However, if this bound on the number of iterations is exceeded, the set is indeed empty.

The ellipsoid E_k can be represented by

$$E_k = \{X \mid (X - X_k)^T B^{-1}{}_k(X - X_k) \leq 1\} \tag{22.5}$$

where X_k is the center and B_k is a n x n symmetric positive definite matrix.

To construct E_{k+1} (see [34]), the hyperplane $A_v X = b_v$ is moved parallel to itself in the feasible direction of A_v^T until it becomes tangential to the ellipsoid E_k at the point Y_k (See Figure 22.1) and thus

$$Y_k = X_k + d_k,$$

$$\text{where} \quad d_k = -\frac{B_k A_V^T}{(A_v B_k A_V^T)^{1/2}}$$

Now, the point $\overline{Y}_k$ on the hyperplane $A_v X = b_v$ is obtained by moving from X_k in the direction d_k and hence

$$\overline{Y}_k = X_k + \lambda_k d_k,$$

$$\text{where the step length } \lambda_k = \frac{A_v X_k - b_v}{(A_v B_k A_V^T)^{1/2}} > 0$$

Let X_{k+1} be the point on the line segment $[\overline{Y}_k Y_k]$ that divides this in the ratio 1: n

Hence

$$X_{k+1} = \left(\frac{n}{n+1}\right)\overline{Y} + \frac{1}{n+1} Y_k$$

$$= X_k - \frac{1 + n\lambda_k}{n+1} \frac{B_k A_V^T}{\left(A_V B_k A_V^T\right)^{1/2}}$$

$$= X_k - \alpha_k \frac{B_k A_V^T}{\left(A_V B_k A_V^T\right)1/2} \tag{22.6}$$

where we assume that $\lambda_k < 1$.

The ellipsoid E_{k+1} with center X_{k+1} can then be given by

$$E_{k+1} = \{X \mid (X - X_{k+1})^T B^{-1}_{k+1} (X - X_{k+1}) \le 1\} \tag{22.7}$$

where $B_{k+1} = \delta_k \left[B_k - \sigma_k \dfrac{[B_k A_v^T][A_v B_k]}{(A_v B_k A_v^T)^{1/2}} \right]$, is a symmetric positive definite matrix.

$$\alpha_k = \frac{1 + n\lambda_k}{n+1}$$
$$\lambda_k = \frac{A_v X_k - b_v}{(A_v B_v A_v^T)^{1/2}}$$
$$\sigma_k = \frac{2(1 + n\lambda_k)}{(n+1)(1+\lambda_k)}$$
$$\delta_k = \frac{n^2}{n^2 - 1}(1 - \lambda_k^2).$$

Note that if $\lambda_k > 1$, (22.2) is infeasible and if $\lambda_k = 1$, E_{k+1} degenerates to a point.

It can be shown that the ellipsoid E_{k+1} defined as above is the minimum volume ellipsoid containing the appropriate part of E_k.

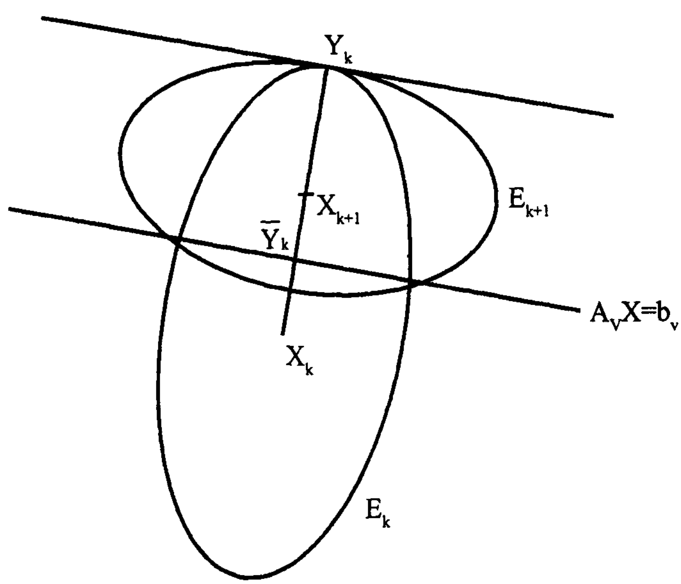


Figure 22.1

22.4. Solving Linear Programming Problems by the Ellipsoid Method

Let us now see how Khachiyan's algorithm can be used to solve a linear programming problem.

Consider the linear programming problem

$$\begin{aligned} &\text{Maximize} && c^T X \\ &\text{Subject to} && AX \leq b \\ & && X \leq 0 \end{aligned} \tag{22.8}$$

where A is m × n, m, n, ≥ 2 and the data are all integers.

There are several ways to apply the ellipsoid method to solve the above problem. We here discuss the approach where we apply the ellipsoid method to a certain system of linear inequalities that yields an optimal solution of the linear programming problem.

The dual problem to (22.8) is given by

$$\begin{aligned} &\text{Minimize} && b^T Y \\ &\text{Subject to} && A^T Y \leq c \\ & && Y \geq 0. \end{aligned} \tag{22.9}$$

and by duality theory in linear programming, X_0 and Y_0 are optimal solutions to (22.8) and (22.9) respectively if and only if X_0, Y_0 solve the system of linear inequalities

$$\begin{aligned} &AX \leq b \\ &-A^T Y \leq -c \\ &-c^T X + b^T Y \leq 0 \\ &-X \leq 0 \\ &-Y \leq 0 \end{aligned} \tag{22.10}$$

Now, (22.10) is in the form of (22.2) and we may apply the ellipsoid method to (22.10) and it produces optimal solutions to the primal and the dual problems simultaneously.

However, in this approach, the ellipsoid method is applied to a system of linear inequalities in R^{m+n} and the high dimensionality slows convergence.

To reduce the dimensionality of the problem, Jones and Marwil [261] suggested a variant of this approach of simultaneously solving the primal and dual using the complementary slackness conditions for the problems (22.8) and (22.9).

The overall complexity of Khachiyan's algorithm for linear programming is $0(n^6L^2)$ where L is the number of bits required to store the problem data in a computer. Unfortunately, this worst case bound is actually achieved in most problems and Khachiyan's ellipsoid method, though expected to be faster than the simplex method is no better. It was therefore, a great disappointment and the simplex method remained the best possible method for solving linear programs. This disappointment however gave great impetus to researchers for making exciting new developments.

22.5. Karmarkar's Polynomial-Time Algorithm

In 1984 Karmarkar [269] proposed a new polynomial-time algorithm for linear programming which performs much better than Khachiyan's ellipsoid method.

Karmarkar considers the linear programming problem in the form

$$\begin{aligned} \text{Minimize} \quad & c^T X \\ \text{Subject to} \quad & AX = 0 \\ & e^T X = 1 \\ & X \leq 0 \end{aligned} \tag{22.11}$$

where $X = (x_1, x_2, \dots x_n)^T \in R^n$

$e = (1,1,..1)^T$, an n-vector with all elements equal to unity and

A is an $m \times n$ matrix with rank m, $n \geq 2$ and the data are all integers.

It is assumed that the following two conditions hold:

A1. The point $X_0 = (l/n, l/n,..l/n)$ is feasible to the problem (22.11)

A2. The optimal value of the objective function in (22.11) is zero.

At first glance, it may appear that the above form of the linear programming problem along with the assumptions Al and A2 is very much restrictive but it can be shown that any general linear programming problem can be easily converted to this form. (See section 22.7)

We now describe Karmarkar's algorithm for solving the problem (22.11). Nice descriptions of the algorithm can also be found in Fletcher [165] and in Bazaraa, Jarvis and Sherali. [34]

Starting with the feasible point $X_0 = (l/n, l/n,.. l/n)$ representing the center of the (n–1) dimensional simplex $\Lambda_X = \{X|e^T X = 1, X \geq 0\}$, the algorithm generates a sequence of feasible points which converges to the optimal solution in polynomial time. Let at step k, $X_k > 0$ be a feasible solution where $c^T X_k \neq 0$. The current feasible point X_k will no longer be at the center of the simplex. For the procedure to be iterative, we use a projective transformation to bring this feasible point to the center of the simplex in the transformed space. Then we optimize the transformed problem under a restriction so that the new feasible solution is an interior point.

Karmarkar uses the projective transformation

$$Y = \frac{D_k^{-1} X}{e^T D_k^{-1} X}. \tag{22.12}$$

where D_k is the diagonal matrix, diag $(x_{k1}, x_{k2}, \dots x_{kn})$, x_{ki} being the ith element of the solution point X_k

Equivalently, the transformation is

$$y_i = \frac{x_i / x_{ki}}{\sum_{j=i}^{n} x_j / x_{kj}}, \qquad i = 1,2,..n \tag{22.13}$$

The inverse transformation is given by

$$X = \frac{D_k Y}{e^T D_k Y} \tag{22.14}$$

Therefore, the transformation (22.12) is one-to-one and maps the X-space into the Y-space uniquely. Thus under the transformation (22.12), any point in the simplex Λ_X is transformed into a point in the simplex $\Lambda_Y = \{Y|e^TY = 1, Y \geq 0\}$ in the Y-space. In particular, the point X_k is transformed into the point $Y_0 = (1/n, 1/n, \ldots 1/n)$, the center of the simplex Λ_Y. Each facet of Λ_X is mapped onto the corresponding facet of Λ_Y and the feasible region in the X-space now becomes the feasible region in the Y-space.

Thus under the transformation (22.12), the problem (22.11) is transformed into the problem

$$\begin{aligned} &\text{Minimize} && \frac{c^T D_k Y}{e^T D_k Y} \\ &\text{Subject to} && AD_k Y = 0 \\ & && e^T Y = 1 \\ & && Y \geq 0 \end{aligned} \tag{22.15}$$

Note that by assumption A2, the optimal value of the objective function in (22.15) is zero and since $e^T D_k Y$ is positive, the problem (22.15) is reduced to the problem

$$\begin{aligned} &\text{Minimize} && c^T D_k Y \\ &\text{Subject to} && AD_k Y = 0 \\ & && e^T Y = 1 \\ & && Y \geq 0 \end{aligned} \tag{22.16}$$

However, to ensure that the new feasible solution is an interior point, we proceed as follows.

Let B (Y_0, r) be the n-dimensional sphere or ball inscribed in the simplex Λ_Y, with the center Y_0, the same as the center of the simplex and radius r. The intersection of this ball with the simplex is then an (n–1) dimensional ball with the same center and radius. This radius r is then the distance from the center (1/n, 1/n,... 1/n) of the simplex to the center of one of its facets. Since the facets of Λ_Y are one lower dimensional simplices, r can be obtained as the distance from (1/n, 1/n,...1/n) to say (0, 1/n–1, 1/n–1,....1/n–1).

$$\text{Hence } r = \frac{1}{\sqrt{n(n-1}}. \tag{22.17}$$

The problem (22.16) is then optimized over a smaller ball B $(Y_0, \alpha r)$, $0 < \alpha < 1$, where α is a parameter which can be chosen suitably between 0 and 1. A smaller ball is considered as it ensures that the feasible solutions of the problem will be interior points with all the coordinates being strictly positive.

We are therefore concerned with the problem

$$\begin{aligned} \text{Minimize} \quad & c^T D_k Y \\ \text{Subject to} \quad & AD_k Y = 0 \\ & e^T y = 1 \\ & Y \in B\,(Y_0, \alpha r) \end{aligned} \tag{22.18}$$

The condition $Y \geq 0$ is disregarded since it is implied by the intersection of the constraint $e^T Y = 1$ and the ball.

For convenience, the problem (22.18) is rewritten as

$$\begin{aligned} \text{Minimize} \quad & \bar{c}^T Y. \\ \text{Subject to} \quad & PY = P_0 \\ & (Y-Y_0)^T (Y-Y_0) \leq \alpha^2 r^2\,, \end{aligned} \tag{22.19}$$

where $\bar{c}^T = c^T D_k$,

$$P = \begin{bmatrix} AD_k \\ e^T \end{bmatrix}, \quad P_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ and}$$

$$B\,(Y_0 \alpha r) = \{Y \mid (Y - Y_0)^T (Y - Y_0) \leq \alpha^2 r^2\}$$

Since $PY = P_0$ is an (n-m-1) dimensional affine subspace that passes through the center of the ball $B\,(Y_0, \alpha r)$, the feasible region of the problem (22.19) is an (n–m–1) dimensional ball centered at Y_0. It is thus evident that the optimal solution to the problem (22.19) is obtained by projecting the negative gradient of the objective function at Y_0 onto the null space of P and moving along this projected direction from the centre Y_0 to the boundary of the ball $B\,(Y_0, \alpha r)$.

Let c_p be the projection of the gradient vector $\bar{c}$ onto the null space of P. c_p can then be obtained by a projection theorem in linear algebra as

$$c_p = [1 - P^T (PP^T)^{-1} P]\ \bar{c} \tag{22.20}$$

Taking a step of length αr in the direction of normalized projection $-c_p/\| c_p \|$, we obtain the optimal solution Y_{new} of the problem (22.19) as

$$Y_{new} = Y_0 - \alpha r \frac{c_p}{\| c_p \|} \tag{22.21}$$

Since from (22.20), we note that $Pc_p = 0$, equation (22.21) implies that

$$\bar{c}^T Y_{new} < \bar{c}^T Y_0 \tag{22.22}$$

The next point X_{k+1} in the sequence in the X-space is then obtained by the inverse transformation (22.14) as

$$X_{k+1} = \frac{D_k Y_{new}}{e^T D_k Y_{new}} \tag{22.23}$$

and since Y_{new} lies in the interior of the ball inscribed in the simplex Λ_Y, it is positive. Hence $X_{k+1} > 0$.

The steps are then repeated and it can be shown that in each step there is a reduction in the value of the objective function of (22.19) and the process approaches the optimal value zero.

The convergence and the complexity of the algorithm is discussed in the next section.

22.6. Convergence and Complexity of Karmarkar's Algorithm

Consider the n-dimensional ball B (Y_0, R) in the Y-space with center Y_0, the same as the center of the simplex Λ_Y and radius R that circumscribes the simplex Λ_Y. R is then the distance from the center Y_0 to any vertex of Λ_Y, that is the distance from the point (l/n, l/n,..l/n) to say (1,0,0..0).

Hence $$R = \sqrt{\frac{n-1}{n}} \tag{22.24}$$

Consider now the problem (22.19), where the constraint $Y \in B(Y_o, \alpha r)$ is replaced by $Y \in B\,(Y_0, R)$, that is the problem

$$\begin{aligned} &\text{Minimize} && \bar{c}^T Y \\ &\text{Subject to} && PY = P_0 \\ & && (Y-Y_0)^T(Y-Y_0) \leq R^2 \end{aligned} \tag{22.25}$$

As in (22.21), the optimal solution $\overline{Y}_{new}$ to the problem (22.25) is given by,

$$\overline{Y}_{new} = Y_0 \frac{Rc_P}{\| c_P \|} \tag{22.26}$$

We note that

$$B\,(Y_0, r) \subseteq \Lambda_y \subseteq B\,(Y_0, R)$$

$$B\,(Y_0,r) \cap \Omega' \subseteq \Lambda_Y \cap \Omega' \subseteq B\,(Y_0 R) \cap \Omega',$$

where $\Omega' \{Y \mid AD_k Y = 0\}$ (22.27)

Denoting Y^*, as an optimal solution of the problem (22.16), we then have

$$\bar{c}^T \overline{Y}_{new} \leq \bar{c}^T Y^* \leq \bar{c}^T Y_{new} < \bar{c}^T Y_0 \tag{22.28}$$

where the last inequality follows from (22.22)

From (22.28) we get

$$0 < \bar{c}^T(Y - Y_{new}) \le \bar{c}^T(Y_0 - Y^*) \le \bar{c}^T(Y_0 - \bar{Y}_{new}) = \frac{R\bar{c}^T c_P}{\| c_P \|} \qquad \text{(by 22.26)}$$

$$= \frac{R}{\alpha r}\bar{c}^T(Y_0 - Y_{new}) \qquad \text{(by 22.21)}$$

Thus, $\bar{c}^T(Y_0 - Y^*) \le \frac{R}{\alpha r}\bar{c}^T(Y_0 - Y_{new})$

$$= \frac{R}{\alpha r}\left[\bar{c}^T(Y_0 - Y^*) - \bar{c}^T(Y_{new} - Y^*)\right]$$

or $\bar{c}^T(Y_{new} - Y^*) \le \left(1 - \frac{\alpha r}{R}\right)\bar{c}^T(Y_0 - Y^*)$

and since $\bar{c}^T(Y_0 - Y^*) > 0,$ we have

$$\frac{\bar{c}^T(Y_{new} - Y^*)}{\bar{c}^T(Y_0 - Y^*)} \le 1 - \frac{\alpha r}{R} = 1 - \frac{\alpha}{n-1} \qquad (22.29)$$

Now, by assumption (A2), $\bar{c}^T Y^* = 0$ and hence

$$\frac{\bar{c}^T Y_{new}}{\bar{c}^T Y_0} \le 1 - \frac{\alpha}{n-1} \qquad (22.30)$$

Thus in each step, there is a reduction in the value of the objective function $\bar{c}^T Y$, which however depends on the choice of α.

We have made the above analysis with the numerator of the objective function of the transformed problem (22.15) and hence does not guarantee that the objective function in the fractional programming problem (22.15) and therefore, in the original problem (22.11) will also decrease. We should note that a linear function is not invariant under a projective transformation (as can be seen from the objectives in (22.11) and (22.15)) but ratios of linear functions are transformed into ratios of linear functions.

Fortunately, there is a novel function known as the potential function which may be used to measure the progress of the algorithm toward optimality.

The potential function defined by Karmarkar is given by,

$$f(X) = \sum_{j=1}^{n} \ln\left(\frac{c^T X}{x_j}\right) \qquad (22.31\ )$$

which has the property that under the projective transformation (22.12), f(X) is transformed into a function of the same form and a reduction in the value of $C^T X$

is achieved by a reduction in the value of f(X) and hence assures convergence to optimality.

The transformed potential function in the Y-space at the kth iteration under (22.12) is

$$F(Y) = f\left(\frac{D_k Y}{e^T D_K Y}\right) = \sum_{j=1}^{n} \ln\left(\frac{\bar{c}^T Y}{x_{kj} y_j}\right)$$

$$= n \ \ln \bar{c}^T Y - \sum_{j=1}^{n} \ln y_j - \sum_{j=i}^{n} \ln x_{kj} \tag{22.32}$$

To measure the decrease in the value of the potential function f(X) or equivalently in F(Y) we compute $F(Y_{new}) - F(Y_0)$ from (22.32) and since Y_{0j} = l/n for j = 1,2,..n, we obtain

$$F(Y_{new}) - F(Y_0) = n \ \ln \left[\frac{\bar{c}^T Y_{new}}{c^T Y_0}\right] - \sum_{j=1}^{n} \ln\left[n\, Y_{new,j}\right]$$

$$\leq n \ln \left[1 - \frac{\alpha}{n-1}\right] - \sum_{j=1}^{n} \ln\left[n\, Y_{new,j}\right] \qquad \text{[from (22.30)]}$$

Since In (1–x) ≤ –x, we have

$$F(Y_{new}) - F(Y_0) \leq -\frac{n\alpha}{n-1} - \sum_{j=i}^{n} \ln\left[n\, Y_{new,j}\right] \tag{22.33}$$

Let, $\bar{\alpha} = \frac{n}{n-1}\alpha$, where $0 < \alpha < 1$ is sufficiently small so that

$$\sqrt{\frac{n-1}{n}}\,\bar{\alpha} = \sqrt{\frac{n}{n-1}}\,\alpha < 1. \tag{22.34}$$

We now make use of the following results from calculus.

Lemma 22.1. If $|\, x \,| \leq \beta < 1$, then $|\ln (1 + x) - x| \leq \frac{x^2}{2(1-\beta)}$

Proof: Follows directly from the expansion of In (l + x),

Lemma 22.2. If $\|\, n\, Y - e \,\| \leq \beta$, for $\beta = \sqrt{\frac{n}{n-1}}\,\alpha < 1$,

$$e^T Y = 1 \ , \ Y > 0, \ \ \text{then} \ \left|\sum_{j=i}^{n} \ln (n\, y_j)\right| \leq \frac{\beta^2}{2(1-\beta)}$$

Proof: $\| n Y - e \| \leq \beta$ implies that

$$\sum_{j=i}^{n} (n y_j - 1)^2 \leq \beta^2$$

$$\text{or } | n y_j - 1 | \leq \beta, \quad j = 1,2,..n$$

By Lemma 22.1, we have

$$\left| \ln[1 + (n y_j - 1)] - (n y_j - 1) \right| \leq \frac{(n y_j - 1)^2}{2(1-\beta)}$$

or
$$\sum_{j=1}^{n} \left| \ln n y_j - (n y_j - 1) \right| \leq \sum_{j=1}^{n} \frac{(n y_j - 1)^2}{2(1-\beta)} \leq \frac{\beta^2}{2(1-\beta)}$$

or
$$\left| \sum_{j=1}^{n} \ln n y_j - \sum_{j=1}^{n} (n y_j - 1) \right| \leq \frac{\beta^2}{2(1-\beta)}$$

and since = $\sum_{j=1}^{n} (n y_j - 1) = 0$, we have

$$\left| \sum_{j=1}^{n} \ln n y_j \right| \leq \frac{\beta^2}{2(1-\beta)} \tag{22.35}$$

Now,
$$\sum_{j=1}^{n} \ln n y_j = \ln n^n \prod_{j=1}^{n} y_j$$

$$= n \ln n + n \ln \left(\prod_{j=1}^{n} y_j \right)^{1/n}$$

$$= n \ln n \left(\prod y_j \right)^{1/n}$$

$$\leq n \ln n . \frac{\Sigma y_j}{n} = n \ln 1 = 0 \text{ since G.M.} \leq \text{A.M.} \tag{22.36}$$

Thus,
$$\sum_{j=1}^{n} \ln n y_j \leq 0$$

Hence from (22.35) we have

$$0 \leq -\sum_{j=1}^{n} \ln(n y_j) \leq \frac{\beta^2}{2(1-\beta)} \tag{22.37}$$

Now since $e^T Y_{new} = 1$, $Y_{new} > 0$ and by (22.21)

$$\|n Y_{new} - e\| = n \alpha r = \sqrt{\frac{n-1}{n}}\,\bar{\alpha} < 1 \quad \left(r = \frac{1}{\sqrt{n(n-1)}}\right),$$

Y_{new} satisfies the conditions of Lemma 22.2 and by (22.37) the equation (22.33) becomes

$$F(Y_{new}) - F(Y_0) \leq -\bar{\alpha} + \frac{\bar{\alpha}^2}{2(1-\bar{\alpha})^2}, \tag{22.38}$$

where $\beta = \sqrt{\frac{n-1}{n}}\,\bar{\alpha}.$

For $\bar{\alpha} = \frac{n\alpha}{n-1} = \frac{1}{3}$, we have

$$F(Y_{new}) - F(Y_0) \leq -1/5 \tag{22.39}$$

Hence the function 'F' and consequently the potential function 'f' decreases by 1/5 in every iteration and thus over k iterations

$$f(X_k) - f(X_0) = n \ln\left(\frac{c^T X_k}{c^T X_0}\right) - \sum_{j=1}^{n} \ln(nX_{kj}) \leq -\frac{k}{5}.$$

Since $\sum_{j=1}^{n} \ln(n X_{kj}) \leq 0$ [See (22.36)], we have

$$\ln\left(\frac{c_T X_k}{c^T X_{k0}}\right) \leq -\frac{k}{5n}$$

or $$c^T X_k \leq c^T X_0\, e^{-k/5n} \tag{22.40}$$

Therefore for $k = 10n\,L$, we have

$$c^T X_k \leq c^T X_0 e^{-2L}$$

where L, the lower bound on the input length for the problem is given by

$$L = [1 + \log(|D_{et\,max}|) + \log(1 + |c_{j\,max}|)] \tag{22.41}$$

where $|D_{et\,max}|$ is the largest numerical value of the determinant of any basis of the problem and $|c_{j\,max}|$ is the largest numerical value of any cost coefficient c_j.

Since $c^T x_0 \leq 2^L$, (See [517]), we have

$$c^T X_k \leq c^T X_0 e^{-2L} < 2^L \cdot 2^{-2L} = 2^{-L} \tag{22.42}$$

Further since at each iteration, the number of arithmetic operations in the worst case is $0(n^3)$ the polynomial complexity of the algorithm is $0\ (n^4L)$.

A modification to the algorithm

The computational effort of the algorithm discussed above is dominated by the computation of c_p, the projection of the gradient vector, where

$$c_P = [I - P^T (PP^T)^{-1}P]\ \bar{c}$$

Thus, in each iteration, we have to find the inverse of $(P\ P^T)$:

$$PP^T = \begin{bmatrix} AD_k^2A^T & 0 \\ 0 & n \end{bmatrix} \tag{22.43}$$

Note that the only change in this matrix, from step to step is in the elements of the diagonal matrix D_k. Taking this advantage Karmarkar has shown that instead of recomputing the inverse it can be obtained by an undating procedure so that the algorithm can run with a bound of $0(n^{2.5})$ operations per step. The overall polynomial complexity of the modified method then reduces to $0(n^{3.5}L)$.

22.7. Conversion of a General Linear Program into Karmarkar's Form

Consider the linear programming problem in the form

$$\begin{aligned} \text{Maximize} \quad & c^Tx \\ \text{Subject to} \quad & AX \leq b \\ & X \geq 0 \end{aligned} \tag{22.44}$$

where A is $m \times n$ of rank m and the data are all integers.

There are several ways by which this problem can be converted into the form considered by Karmarkar accompanying the assumptions Al and A2. We now present one such method.

The dual of the problem (22.44) is given by

$$\begin{aligned} \text{Minimize} \quad & b^TY \\ \text{Subject to} \quad & A^TY \geq c \\ & Y \geq 0 \end{aligned} \tag{22.45}$$

By duality theory, the system of inequalities (22.46) has a solution if and only if the original problem has a finite optimal solution.

$$\begin{aligned} & AX \leq b \\ & A^TY \geq c \\ & c^TX - b^TY = 0 \\ & X \geq 0,\ Y \geq 0 \end{aligned} \tag{22.46}$$

Adding slack and surplus variables x_{n+i}, $i = 1, 2 \ldots m$ and Y_{m+j}, $j = 1, 2..n$ in the problem, (22.46) we have

$$\sum_{j=1}^{n} a_{ij}x_j + x_{n+i} = b_i, \quad i = 1,2,..m$$

$$\sum_{i=1}^{m} a_{ij}y_i - y_{m+j} = c_j, \quad j = 1,2,..n$$

$$\sum_{j=1}^{n} c_j x_j - \sum_{i=1}^{m} b_i y_i = 0. \tag{22.47}$$

$$x_j \geq 0, j = 1,2\ldots n, n+1\ldots n+m$$
$$y_i \geq 0, i = 1,2\ldots m, m+1,\ldots m+n$$

Let us now define Q by

$$\sum_{j=1}^{n+m} x_j + \sum_{i=1}^{m+n} y_i \leq Q \tag{22.48}$$

where Q is a sufficiently large number so that any solution of (22.47) satisfies (22.48)

Introducing the bounding constraint

$$\sum_{j=1}^{n+m} x_j + \sum_{i=1}^{m+n} y_i + s_1 = Q$$

where $s_1 \geq 0$ is a slack variable, in the problem (22.47) we have

$$\sum_{j=1}^{n} a_{ij}x_j + x_{n+i} = b_i \quad i = 1,2,..m$$

$$\sum_{i=1}^{m} a_{ij}y_j - y_{m+j} = c_j, \quad j = 1,2,..n$$

$$\sum_{j=1}^{n} c_j x_j - \sum_{i=1}^{m} b_i y_i = 0 \tag{22.49}$$

$$\sum_{j=1}^{n+m} x_j + \sum_{i=1}^{m+n} y_i + s_1 = Q$$

all $x_j, y_i \geq 0$; $s_1 \geq 0$

Further to obtain the problem in the homogeneous form, we introduce an additional variable s_2 with the restriction $s_2 = 1$ and express the problem as

$$\sum_{j=1}^{n} a_{ij}x_j + x_{n+i} - b_i s_2 = 0, \quad i = 1,2..m$$

$$\sum_{i=1}^{m} a_{ij}y_i - y_{m+j} - c_j s_2 = 0, \quad j = 1,2..n$$

$$\sum_{j=1}^{n} c_j x_j - \sum_{i=1}^{m} b_i y_i = 0$$

$$\sum_{j=1}^{n+m} x_j + \sum_{i=1}^{m+n} y_i + s_1 - Qs_2 = 0$$

$$\sum_{j=1}^{n+m} x_j + \sum_{i=1}^{m+n} y_i + s_1 + s_2 = Q+1. \tag{22.50}$$

all x_j, $y_i \geq 0$ and s_1, $s_2 \geq 0$

where the last two constraints are equivalent to the constraint $\Sigma_j x_j + \Sigma_i y_i + s_1 = Q$ in (22.49) and $s_2 = 1$

Let
$$\begin{aligned} x_j &= (Q+1)v_j, \; j = 1,2,\ldots.n+m, \\ y_i &= (Q+1)v_{n+m+i}, \; i = 1,2\ldots\ldots m+n, \\ s_1 &= (Q+1)v_{2m+2n+1}, \\ s_2 &= (Q+1)v_{2m+2n+2}. \end{aligned} \tag{22.51}$$

The system is then reduced to

$$\sum_{j=i}^{n} a_{ij}v_j + v_{n+i} + b_i v_{2m+2n+2} = 0, \quad i = 1,2,\ldots.m$$

$$\sum_{i=1}^{m} a_{ij}v_{n+m+i} - v_{2m+n+j} - c_j v_{2m+2n+2} = 0, \quad j = 1,2,\ldots.n$$

$$\sum_{j=1}^{n} c_j v_j - \sum_{i=1}^{m} b_i v_{n+m+i} = 0 \tag{22.52}$$

$$\sum_{j=1}^{n+m} v_j + \sum_{i=1}^{m+n} v_{m+n+i} + v_{2m+2n+1} - Qv_{2m+2n+2} = 0$$

$$\sum_{j=1}^{n+m} v_j + \sum_{i=1}^{m+n} v_{m+n+i} + v_{2m+2n+1} - Qv_{2m+2n+2} = 1$$

$$v \geq 0$$

We now introduce the artificial variable λ in the constraints of (22.52) such that the sum of the coefficients in each homogeneous constraint and of λ is zero so that Karmakar's assumption A1 holds.

We then consider the problem

Minimize $\quad \lambda$

Subject to $\sum_{j=1}^{n} a_{ij}v_j + v_{n+i} - b_i v_{2m+2n+2} + K_i\lambda = 0, \; i = 1,2\ldots m$

$$\sum_{i=1}^{m} a_{ij}v_{n+m+i} - v_{n+2m+j} - c_j v_{2m+2n+2} + K_{m+j}\,\lambda = 0, \;\; j=1,2\ldots n$$

$$\sum_{j=i}^{n} C_j v_j - \sum_{i=1}^{m} b_i v_{n+m+i} + k_{m+n+1}\lambda = 0,$$

$$\sum_{j=i}^{2m+2n+1} v_j - Qv_{2m+2n+2} + (Q-(2m+2n+1))\lambda = 0$$

$$\sum_{j=i}^{2m+2n+2v} v_j + \lambda \;\; = 1. \tag{22.53}$$

All $V \geq 0, \lambda \geq 0$ and

K's are integers, positive or negative or zero.

If the sum of the coefficients in a particular constraint of (22.52) is already zero, the artificial variable λ need not be added to that constraint, that is, the corresponding K is of value zero. Note that the minimum value of λ in (22.53 is zero if and only if the system pf inequalitites in (22.46) has a solution.

The problem (22.53) is now in Karmarkar's form and assumptions Al and A2 are also satisfied. By solving this problem by Karmarkar's algorithm, we obtain optimal solutions to the primal and dual problems of the original linear program simultaneously. One disadvantage of this approach of combining the primal and the dual into a single problem is the increase in dimension of the system equations which must be solved at each iteration.

Karmarkar suggested some modifications to his algorithm in order to solve a general L.P. problem under the assumption Al only dispensing the assumption A2, that is, without the requirement that the optimal value of the objective function has to be zero.

By the use of the Big-M method, where we are to add artificial columns and/or artificial rows and a large penalty parameter M, the problem (22.44) can be converted into Karmarkar's form along with the assumption A1 only. Karmarkar's modified algorithm may then be applied to solve the problem.

Several other approachces are also proposed.

Ever since Karmarkar's algorithm was published, there was a spurt of research activities in polynomial-time algorithms for linear programming and a number of variants and extensions of his algorithm appeared in literature. For example, Charnes, Song and Wolfe [80], Anstreicher [12], Tomlin [470], Todd and Burrell [469], Kojima [275], Gay [192], Lusting [307] among others. A particularly perceptive paper is that of Gill et al [198] who have shown relationships of Karmarker's algorithm to projected Newton barrier methods.

22.8 Exercises

1. Show that solving the linear programming problem Minimize c^TX, subject to $AX \geq b$, $X \geq 0$ by Khachian's ellipsoid method is equivalent to finding a solution to a system of linear inequalities.
2. Consider the linear programming problem Minimize c^TX subject to $AX = b$, $X \geq 0$, where A is $m \times n$ and the coefficients of one of the constraints, say $a_{mj} > 0$, $j = 1, 2..n$ and $b_m > 0$. Show how the problem can be converted into Karmarkar's form.
3. Consider the linear programming problem Minimize c^TX subject to $AX = b$, $X \geq 0$, where A is $m \times n$ of rank m and the data are all integers. Show how the big M method can be used to convert the problem into Karmarkar's form where Assumption Al only is satisfied. Also, show from this converted problem, how one can determine whether the given linear program is feasible or unbounded.
   [Hint: Proceed in the same way as in Section 22.7 with the primal problem only. After reducing the system similar to (22.53) add $M\lambda$ to the objective function C^TX.]
4. Show that the matrix PP^T is nonsingular where P is defined as in (22.19).
5. Convert the following problem into Karmarkar's form which satisfies Assumption A,

$$\begin{aligned} \text{Maximize} \quad & z = x_1 + x_2 \\ \text{Subject to} \quad & 2x_1 + x_2 \leq 2 \\ & x_1, x_2\,^3 \geq 0 \end{aligned}$$

6. Perform three iterations of Karmarkar's algorithm on the following problem.

$$\begin{aligned} \text{Maximize} \quad & z = x_1 + x_2 - 4x_3 \\ \text{Subject to} \quad & x_1 - x_2 - 2x_3 + 2x_4 = 0 \\ & x_1 + x_2 + x_3 + x_4 = 1 \\ & x_j \geq 0, \; j = 1,2,3,4. \end{aligned}$$

7. Solve the following linear programming problem by Karmarkar's algorithm

$$\begin{aligned} \text{Maximize} \quad & z = 2x_1 - x_3 \\ \text{Subject to} \quad & 2x_1 - x_2 - x_3 = 0 \\ & x_1 + x_2 + x_3 = 1 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

8. Solve the following linear programming problem by Karmarkar's algorithm.

$$\begin{aligned} \text{Minimize} \quad & z = x_1 \\ \text{Subject to} \quad & x_1 - 2x_2 + x_3 = 0 \\ & x_1 + x_2 + x_3 = 1 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

PART – 3

NONLINEAR AND DYNAMIC PROGRAMMING

This Page Intentionally Left Blank

CHAPTER 23

Nonlinear Programming

23.1. Introduction

We recall that the general problem of mathematical programming problem can be stated as

$$\begin{aligned} &\text{Minimize} && f(X) \\ &\text{Subject} && g_i(X) \le 0,\ 1 = 1,2,\ldots m \\ &&& X \ge 0 \end{aligned}$$

where $f(X)$, $g_i(X)$, $i = 1,2\ldots m$ are real valued functions of $X \in R^n$

If the constraints and the function to be minimized are linear, this is known as a linear programming problem, otherwise this is said to be a nonlinear program.

We have noted that a very large class of decision problems could be formulated as linear programming problems and the simplex method is powerful enough to solve all of these problems. Very soon however, it was recognized that many a practical problem cannot be represented by linear programming model and interest in nonlinear programming developed almost simultaneously with the growing interest in linear programming. Therefore attempts were made to develop more general mathematical programming methods and many significant advances have been made in the area of nonlinear programming. The first major development was the fundamental paper by Kuhn and Tucker in 1951 [291] which laid the foundations for a good deal of later work in nonlinear programming.

In general, nonlinear programming problems present much greater difficulties to solve than linear programs, because for a nonlinear program, the optimal solution may occur at an interior point or on the boundary of the feasible set. Moreover, a local optimal may not be a global one, which makes the problem more difficult. Most of the computational techniques for nonlinear programming therefore, aim at finding a local solution of the problem. If however, f and g_i, $i = 1,2.m$ are convex then any local solution is global. Such problems form a special class of nonlinear programming and is called convex programming.

Before we consider the traditional nonlinear programming problems, we first present some results regarding optimization of problems without constraints.

23.2. Unconstrained Optimization

In this section we shall discuss the optimality conditions of a function f(X) over $X \in R^n$, when there are no constraints on the variables. Unconstrained problems do not arise frequently in practice. However, the optimality conditions for unconstrained problems are important not only for themselves but also for the fundamental role they play in the development of many techniques for solving constrained optimization problems.

The problem of our concern here, is to determine whether a point X_0 is a local or a global minimum point of f(X) over $X \in R^n$. A point $X_0 \in R^b$ is called a local minimum of the function if $f(X_0) \leq f(X)$ for each $X \in N_\epsilon(X_0)$. It is global minimum if $f(X_0) \leq f(X)$ for all $X \in R^n$.

Theorem 23.1. Let f be a real valued function defined on R^n and is differentiable at X_0. If there is a vector d such that

$$\nabla f(X_0)^T d < 0 \tag{23.1}$$

then there exists a $\delta > 0$, such that

$$f(X_0 + \lambda. d) < f(X_0) \tag{23.2}$$

for all $\lambda \in (0, \delta)$

Proof: Since f is differentiable at X_0, we have

$$f(X_0 + \lambda d) = f(X_0) + \lambda \nabla f(X_0)^T d + \lambda \|d\| \alpha(X_0; \lambda d).$$

where $\alpha(X_0; \lambda d) \to 0$ as $\lambda \to 0$.

Hence $$\frac{f(X_0 + \lambda d) - f(X_0)}{\lambda} = \nabla f(X_0)^T d + \|d\| \alpha(X_0; \lambda d)$$

and $$\lim_{\lambda \to 0} \frac{f(X_0 + \lambda d) - f(X_0)}{\lambda} = \nabla f(X_0)^T d < 0.$$

From the definition of limit, it then follows that there exists a $\delta > 0$ such that for all $\lambda \neq 0$, and $-\delta < \lambda < \delta$

$$\frac{f(X_0 + \lambda d) - f(X_0)}{\lambda} < 0.$$

Selecting $\lambda > 0$, we have

$$f(X_0 + \lambda d) - f(X_0) < 0, \text{ for all } \lambda \in (0, \delta)$$

Hence, it follows that d is a descent direction of the function f at X_0.

Theorem 23.2. (First order necessary conditions)

Let the function $f : R^n \to R^1$ be differentiable at X_0. If X_0 is a local minimum, then

$$\nabla f(X_0) = 0 \tag{23.3}$$

Proof : Suppose that $\nabla f(X_0) \neq 0$. Then for some j,

$$\frac{\partial f(X_0)}{\partial x_j} < 0 \quad \text{or} \quad \frac{\partial f(X_0)}{\partial x_j} > 0.$$

Then by selecting d_j with the appropriate sign, it is always possible to have

$$d_j \frac{\partial f(X_0)}{\partial x_j} < 0.$$

Thus, if we select $d = -\nabla f(X_0)$, we have

$$\nabla f(X_0)^T d = -\|\nabla f(X_0)\|^2$$

and by Theorem 23.1, there exists a $\delta > 0$ such that

$$f(X_0 + \lambda d) < f(X_0), \text{ for all } \lambda \in (0, \delta).$$

This contradicts the assumption that X_0 is a local minimum. Hence $\nabla f(X_0) = 0$.

Theorem 23.3. (Second order necessary conditions).

Let the function $f: R^n \to R^1$ be twice differentiable at X_0. If X_0 is a local minimum, then

(i) $\nabla f(X_0) = 0$, and

(ii) $H(X_0)$, the Hessian matrix evaluated at X_0 is positive semidefinite.

Proof: Since X_0 is a local minimum, from theorem 23.2, we have $\nabla f(X_0) = 0$ and since f is twice differentiable at X_0, we have for some non-zero vector d.

$$f(X_0 + \lambda d) = f(X_0) + \lambda \nabla f(X_0)^T d + 1/2\, \lambda^2 d^T H(X_0) d + \lambda^2 \|d\|^2 \alpha (X_0 ; \lambda d) \tag{23.4}$$

where $\alpha (X_0 ; \lambda d) \to 0$ as $\lambda \to 0$.

Thus
$$\frac{f(X_0 + \lambda d) + f(X_0)}{\lambda^2} = \frac{1}{2} d^T H(X_0) d + \| d \|^2 \alpha(X^0 ; \lambda d) \tag{23.5}$$

Now, since X_0 is a local minimum,

$$f(x_0 + \lambda d) \geq f(X_0), \text{ for } \lambda \text{ sufficiently}$$

small and we have

$$\frac{1}{2} d^T H(X_0) d + \| d \|^2 \alpha(X_0 ; \lambda d) \geq 0 \qquad 23.6)$$

By taking the limit as $\lambda \to 0$, we then have,

$d^T H(X_0) d \geq 0$ and hence

$H(X_0)$ is positive semidefinite

The necessary condition $\nabla f(X_0) = 0$ for X_0 to be a local minimum is not sufficient. The condition is also satisfied by a local maximum or by a saddle point (defined in section 23.7). The points satisfying $\nabla f(X) = 0$ are called stationary points or critical points of f. Theorem 23.4 gives a sufficient condition for a local minimum.

Theorem 23.4. (Sufficient conditions).

Suppose that the function $f: R^n \to R^1$ is twice differentiable at X_0. If $\nabla f(X_0) = 0$ and the Hessian matrix $H(X_0)$ is positive definite, then X_0 is a local minimum.

Proof: Since f is twice differentiable at X_0, by Taylor's theorem, we have,

$$f(X) = f(X_0) + \nabla f(X_0)^T (X - X_0) + \frac{1}{2}(X-X_0)^T H [\lambda X_0 + (1-\lambda)X] (X-X_0) \tag{23.7}$$

for some λ, $0 < \lambda < 1$.

Since by hypothesis $\nabla f(X_0) = 0$, (23.7) yields

$$f(X) - f(X_0) = \frac{1}{2}(X (X-X_0)^T H [\lambda X_0 + (1-\lambda)X] (X-X_0) \tag{23.8}$$

Now, since we have assumed the existence and continuity of the second partial derivates of $f(X_0)$, it is clear that the second partial derivates

$$\frac{\partial^2 f(X^0)}{\partial X_i \partial X_j}$$

will have the same sign as the second partial derivatives

$$\frac{\partial^2}{\partial X_i \partial X_j} f[\lambda X_0 + (1-\lambda)X]$$

in some neighbourhood $N_\delta(X_0)$.

Thus if $H(X_0)$ is positive definite,

$$f(X) - f(X_0) > 0 \text{ for } X \in N_\delta(X_0).$$

Hence X_0 is a local minimum.

The following theorem shows that the necessary condition $\nabla f(X_0) = 0$ is also sufficient for X_0 to be a global minimum if f is pseudoconvex.

Theorem 23.5. Let f be a pseudoconvex function on R^n. Then $\nabla f(X_0) = 0$ is a necessary and sufficient condition that X_0 is a global minimum of over R^n.

Proof: The necessity holds by theorem 23.2.

Now, suppose that $\nabla f(X_0) = 0$.

Since f is pseudoconvex,

$$\nabla f(X_0)^T (X - X_0) \geq 0 \Rightarrow f(X) \geq f(X_0), \text{ for all } X \in R^n$$

and since $\nabla f(X_0) = 0$, we have $\nabla f(X_0)^T (X - X_0) = 0$

and hence

$$f(X) \leq f(X_0), \text{ for all } X \in R^n$$

23.3. Constrained Optimization

Consider the nonlinear programming problem

$$\begin{aligned} &\text{Minimize} && f(X) \\ &\text{Subject to} && g_i(X) \leq 0, \ i = 1,2..m \end{aligned} \tag{23.9}$$

where the functions f and g_i are f: $R^n \to R^1$ and g_i: $R^n \to R^1$ are assumed to be differentiable.

Let the constraint set be denoted by S, that is

$$S = \{X | g_i(X) \leq 0, i = 1,2...m\}$$

To develop the optimality conditions for the problem, we first introduce the concept of feasible direction.

Feasible Directions

Let S be a nonempty set in R^n. A direction d is said to be a feasible direction at $X \in S$, if we do not leave the region S, when making a sufficiently small move in the direction d.

Mathematically, a direction d is feasible at $X \in S$, if there exists a $\delta > 0$, such that $X + \lambda d \in S$ for all $\lambda \in (0, \delta)$.

Let the set of all feasible directions at X be denoted by

$$D(X) = \{d | \text{ there exists } \delta > 0 \text{ such that } X + \lambda d \in s, \text{ for all } \lambda \in (0, \delta)\} \tag{23.10}$$

Theorem 23.6. If X_0 is a local optimal solution to the nonlinear program (23.9), then

$$\nabla f(X_0)^T d \geq 0, \text{ for all } d \in D(X_0). \tag{23.11}$$

Proof : Suppose there is a direction $d^0 \in D(X_0)$ such that

$$\nabla f(X_0)^T d^0 < 0.$$

Then by Theorem 23.1, a small movement from X_0 in the direction d^0 would decrease $f(X_0)$ contradicting the assumption that X_0 is a local optimal solution to the problem.

Let $\bar{D}$ $X_0)$ be the closure of $D(X_0)$, so that every point in $\bar{D}(X_0)$ is the limit of points in $D(X_0)$. Obviously, $D(X_0) \subset \bar{D}(X_0)$, but a point in $\bar{D}(x_0)$ need not be in $D(X_0)$ unless $D(X_0)$ is itself a closed set when $D(X_0) = \bar{D}(X_0)$.

We will now show that Theorem 23.6 also holds for all directions in $\bar{D}(x_0)$.

Corollary 23.1. If X_0 is a local optimal solution to the nonlinear programming problem (23.9), then

$$\nabla f(X_0)^T d \geq 0, \text{ for all } d \in \bar{D}(X_0) \tag{23.12}$$

Proof: Let $d \in \bar{D}(X_0)$, the closure of $D(X_0)$. Then the direction d can be expressed as a limit of directions d^k in $D(X_0)$.

$$\text{Thus, } d = \lim_{\lambda \to 0} d^k, \quad d^k \in D(X_0).$$

Since $d^k \in D(X_0)$, by Theorem 23.6

$$\nabla f(X_0)^T d^k \geq 0, \text{ for all k}.$$

By taking the limit as $k \to \infty$, we then have

$$\nabla f(X_0)^T d = \lim_{k\to\infty} f(X_0)^T d^k \geq 0, \quad d \in \overline{D}(X_0)$$

23.4. Kuhn–Tucker Optimality Conditions

It has been shown in Corollary 23.1 that a necessary conditions for X_0 to be a local optimal solution to the nonlinear program (23.9) is that $\nabla f(X_0)^T d \geq 0$ for all $d \in \overline{D}(X_0)$. But since $\overline{D}$ is not defined in terms of the constraints, it is not possible to express the conditions in more usable algebraic statement. We therefore formulate $\overline{D}(x_0)$ in terms of the constraints and then by making use of it, develop the Kuhn–Tucker optimality conditions.

Let $X \in S$ and suppose that for some i's, $g_i(X) < b_i$. Since g_i are continuous, a small move in any direction from X will not violate these constraints and hence the inactive constraints $g_i(X) < 0$ do not influence $\overline{D}(X)$. For our purpose, therefore, we need to consider only the active constraints $g_i(X) = 0$. At a feasible point X, the constraints are therefore divided into two sets, one consisting of active constraints and the other consisting of inactive constraints.

Let I (X) be the set of indices for which $g_i(X)$ are active (binding) at a feasible point X, that is

$$\text{for } X \in S, \ I(X) = \{i | g_i(X) = 0 \qquad (23.13)$$

This implies that for $i \notin I(X)$, $g_i(X) < 0$.

We now define the set

$$\mathscr{D}(X) = \{d | \nabla gi(X)^T d \leq 0, \text{ for all } i \in I(X)\}. \qquad (23.14)$$

It is clear that is a nonempty closed set.

Theorem 23.7. $\overline{D}(X) \subset \mathscr{D}(X)$.

Proof: Let $d \in D(X)$ and suppose that $g_i(X) = 0$. If $\nabla g_i(X)^T d > 0$, then from Theorem 23.1, it follows that there exists a $\delta > 0$, such that

$$g_i(X + \lambda d) > g_i(X) = 0, \text{ for all } \lambda \in (0,\delta)$$

which implies that the direction d is not feasible. Hence $\nabla g_i(X)^T d \leq 0$, for all $i \in I(X)$.

Consequently, $\mathscr{D}(X) \subset \mathscr{D}(X)$

Since $\mathscr{D}(X)$ is a closed set, it follows that $\overline{D}(X) \subset \mathscr{D}(X)$

In general however, the condition that $\nabla g_i(X)^T d \leq 0$ holds for all $i \in I(X)$ is not sufficient to guarantee that d is a feasible direction. In other words, there may be directions in $\mathscr{D}(X)$ that are not in $\overline{D}(X)$.

Example. Consider the constraint set generated by

$$g_1(X) = x_1^3 - x_2 \le 0$$
$$g_2(X) = x_1^3 + x_2 \le 0$$

At the point $X = (0, 0)^T$, both $g_1(X)$ and $g_2(X)$ are satisfied as equalities.

The gradient vectors are

$$\nabla g_1(X)^T = (3x_1^2, -1)$$
$$\nabla g_2(X)^T = (3x_1^2, +1).$$

Thus for $d = (1, 0)^T$, we have

$$\nabla g_1(0)^T d = 0$$
$$\nabla g_2(0)^T d = 0$$

Hence $d \in \mathscr{D}(X)$

But obviously d points in an infeasible direction as any point on the X_1 axis to the right of the origin violates the constraints (see Figure 23.1).

Thus for $\lambda > 0$, $X + \lambda d \notin S$ and therefore d is not a feasible direction.

Hence $\quad d \notin \overline{D}(X)$

In order to exclude such cases Kuhn and Tucker assumed certain conditions, known as the constraint qualification to hold at the boundaries of the constraint set.

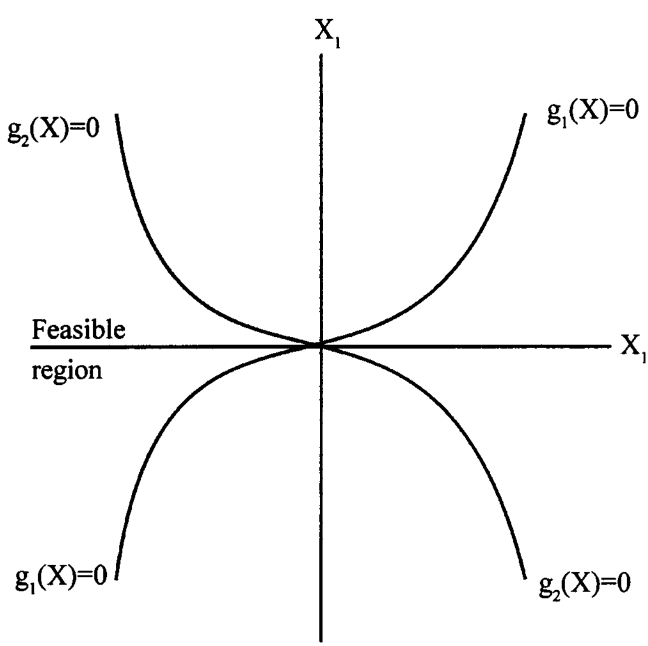


Figure 23.1.

23.5. Kuhn–Tucker Constraint Qualification

The K–T constraint qualification is said to hold at a feasible point X of the problem (23.9) if every vector d $\wp$(X) is tangent to a differentiable are[1] contained in the constraint set. In other words, the constraint qualification holds at X $\in$ S if, given d $\in \wp$ (X), there exists a differentiable are a(θ) with the properties:

(i) a (0) = X

(ii) there is a $\delta < 0$ such that for all θ, $0 \leq \theta \leq \delta$ a (θ)$\in$ S, and

(iii) $\lim\limits_{\theta \to 0^+} \dfrac{a(\theta(-a(0)}{\theta} = d$

Theorem 23.8. If X_0 is a local optimal solution to the nonlinear program (23.9) and the Kuhn–Tucker constraint qualification holds, then

$$\nabla f(X_0)^T d \geq 0, \text{ for all } d \in \mathscr{D}(x_0)$$

Proof: By Taylor's expansion about X_0, we have

$$f(X) = f(X_0) + \nabla (X_0 + \mu (X - X_0))^T (X - X_0) \text{ for some } \mu,\ 0 \leq \mu \leq 1$$

Since X_0 is a local minimum,

$$f(X) - f(X_0) \geq 0$$

for all feasible point X, sufficiently close to X_0

Hence $$\nabla f(X_0 + \mu (X - X_0))^T (X - X_0) \geq 0$$

Since the constraint qualification holds, there exists a differentiable are a(θ) such that a (0) = X_0 and for a small positive θ, a (θ) = X $\in$ S.

Hence $$\nabla f(X_0 + \mu \{a(\theta) - a(\theta)\})^T (a(\theta) - a(0)) \geq 0$$

Dividing both sides of the inequality by $\theta > 0$ and taking the limit as $\theta \to 0$,

we have $$\nabla f(X_0)^T \lim_{\theta \to 0} \frac{a(\theta) - a(0)}{\theta} \geq 0$$

or $$\nabla f(X_0)^T d \geq 0$$

as was to be proved.

Theorem 23.9. (Kuhn–Tucker Necessary Conditions)

If X_0 is a local optimal solution to the nonlinear program (23.9) and the K-T constraint qualification holds, then there exists scalars λ_i, i = 1,2,..m. such that

$$\nabla f(X_0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(X_0) = 0 \tag{23.15}$$

$$\lambda_i g_i(X_0) = 0, \qquad i = 1,2..m \tag{23.16}$$

$$\lambda_i \geq 0, \qquad i = 1,2..m \tag{23.17}$$

1. A continuous function 'a': $R^1 \to R^n$ is said to be an are in R^n. As the parameter θ varies, a (θ) describes a path in R^n.

Proof: If X_0 is a local optimal solution, then by Theorem 23.8

$$\nabla f(X_0)\, d \geq 0$$

for all d satisfying

$$\nabla g_i(X_0)\, d \leq 0, \text{ for all } i \in I\,(X_0)$$

Then by Farkas' lemma, there exit scalars $\lambda_i \geq 0$, $i \in I\,(X_0)$, such that

$$\nabla f(X_0) + \sum_{i \in I(X_0)}^{m} \lambda_i \nabla g_i(X_0) = 0$$

Now, by letting $\lambda_i = 0$ for all $i \notin I\,(X_0)$, we have,

$$\nabla f(X_0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(X_0) = 0$$

$$\lambda_i g_i(X_0) = 0, \quad i = 1,2..m$$

$$\lambda_i \qquad \geq 0, \quad i = 1,2..m$$

The conditions (23.15) – (23.17) are known as the Kuhn–Tucker conditions. The scalars λ_i are called the Lagrangian multipliers and the condition $\lambda_i g_i\,(X_0) = 0$, $i = 1,2,..m$ is referred to as the complementary slackness condition.

Theorem 23.10. (Kuhn–Tucker Sufficient Conditions)

Consider the nonlinear programming problem (23.9) where the functions f and g_i are differentiable. Let the objective function f be pseudoconvex and the functions g_i, $i = 1,2,..m$ be quasiconvex. Suppose that a feasible solution X_0 satisfies the Kuhn–Tucker conditions. Then X_0 is a global optimal solution to the problem.

Proof: Let X be any feasible solution to the problem Since g_i ($i = 1,2..m$) are quasiconvex, the constraint set S is convex and therefore $d = X - X_0$ is a feasible direction.

Hence by Theorem 23.7, $\quad d \in \mathscr{D}(X_0)$

and thus $\quad \nabla g_i(x_0)^T d \leq 0$, for $i \in I(X_0)$ (23.18)

Multiplying (23.18) by $\lambda_i \geq 0$, for $i \in I(X_0)$ and by $\lambda_i = 0$, for $i \notin I(X_0)$ and summing over i, we have,

$$\sum_{i=1}^{m} \lambda_i \nabla g_i(X_0)^T d \leq 0$$

Since K-T conditions are satisfied,

$$\nabla f(X_0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(X_0) = 0.$$

It then follows that

$$\nabla f(X_0)^T d \geq 0.$$

or $\quad \nabla f(X_0)^T (X - X_0) \geq 0$

From the pseudo convexity of f, we then have,

$$f(X) \geq f(X_0)$$

and hence X_0 is a global optimal solution.

Kuhn–Tucker Conditions for Problems with Inequality and Equality Constraints

Consider the problem,

$$\begin{aligned} &\text{Minimize} && f(X) \\ &\text{Subject to} && g_i(X) \leq 0,\ i = 1,2..m_1 \\ &&& h_i(X) = 0,\ i = m_1 + 1,..m \\ &&& X \geq 0 \end{aligned} \tag{23.19}$$

where f, g_i, $i = 1,2..m_1$, and h_i, $i = m_1 + 1..m$ are continuously differentiable functions defined on R^n.

Theorem 23.11: IF X^0 is a local optimal solution to the nonlinear program (23.19) and the constraint qualification holds, then there exists scalars u_i, $i = 1,2..m_1$ and V_i, $i = m_1 + 1, ..m$ such that

$$\nabla f(X^0) + \sum_{i=1}^{m_1} u_i \nabla g_i(X^0) + \sum_{i=m_1+1}^{m} v_i \nabla h_i(X^0) \geq 0.$$

$$\left[\nabla f(X^0) + \sum_{i=1}^{m_1} u_i \nabla g_i(X^0) + \sum_{i=m_1+i}^{m} v_i \nabla h_i(X^0)\right]^T X^0 = 0. \tag{23.20}$$

$$\begin{aligned} u_i g_i(X^0) &= 0, && i = 1,2..m_1, \\ u_i &\geq 0, && i = 1,2..m_i, \\ &v_i \text{ unrestricted}, && i = m_1 + 1,....m \end{aligned}$$

Proof: We first rewrite the problem in the form that enables us to apply the Kuhn–Tucker conditions directly. The problem then becomes

$$\begin{aligned} &\text{Minimize} && f(X) \\ &\text{Subject to} && g_i(X) \leq 0,\ i = 1,2..m_1 \\ &&& h_i(X) \leq 0,\ i = m_1 + 1..m \\ &&& -h_i(X) \leq 0,\ i = m_1 + 1,..m \\ &&& -X \leq 0 \end{aligned} \tag{23.21}$$

An application of Theorem 23.9 implies that there exist scalars u_i, $i = 1,2..m_1$, α_i, β_i $i = m_1 + 1,..m$ and μ_j $j = 1,2,..n$ such that

$$\nabla f(X^0) \sum_{i=1}^{m1} u_i \nabla g_i(X^0) + \sum_{i=m_1+1}^{m} \alpha_i \nabla h_i(X^0) - \sum_{m_1+1}^{m} \beta_i \nabla h_i(X^0) - \mu = 0$$

$$u_i g_i(X^0) = 0, \qquad i = 1,2..m_1$$

$$\alpha_i h_i(X^0) = 0, \qquad i = m_1 + 1..m \tag{23.22}$$
$$\beta_i\, h_i(X^0) = 0 \qquad i = m_1 + 1..m$$
$$-\mu_j X_j^0 = 0 \qquad j = 1 ...n$$
$u_i, \alpha_i, \beta_i, \mu_j$ are all nonnegative

On setting $\alpha_i - \beta_i = v_i$, the conditions reduce to

$$\nabla f(X^0) + \sum_{i=1}^{m_1} u_i \nabla g_i(X^0) + \sum_{i=m_1+1}^{m} V_i \nabla h_i(X^0) \geq 0$$

$$\left[\nabla f(X^0) + \sum_{i=1}^{m_1} u_i \nabla g_i(X^0) + \sum_{i=m_1+1}^{m} V_i h_i(X^0)\right]^T X^0 = 0.$$

$$u_i g_i(X^0) = 0, \qquad i = 1,2..m_1$$
$$u_i \geq 0, \qquad i = 1,2..m_1 \tag{23.23}$$
$$v_i \text{ unrestricted, } i = m_1 + 1,....m$$

23.6. Other Constraint Qualifications

In 1948, Fritz John [257] developed the necessary optimal conditions, that if X_0 is a local optimal solution of the nonlinear program (23.9) then there exist scalars u_0 and u_i for $i = 1,2..m$ such that.

$$u_0 \nabla f(X_0) + \sum_{i=1}^{m} u_i \nabla g_i(X_0) = 0$$
$$u_i g_i(X_0) = 0, \qquad i = 1,2,..m$$
$$u_0, u_i \geq 0, \qquad i = 1,2,..m$$

where not all u_0, u_i, $i = 1,2..m$ are equal to zero.

It should be noted that in the above conditions, there is no guarantee that $u_0 > 0$. If $u_0 = 0$, Fritz John conditions fail to provide any information in locating an optimal solution of the problem because then the term $u_0 \nabla f(X_0)$ disappears from the above system. In order to exclude such cases, we need to introduce some regularity conditions.

In 1951, Kuhn and Tucker [291] independently developed necessary optimality conditions under certain regularity conditions known as constraint qualifications, which ensures that $u_0 > 0$.

Since then many authors have developed the Kuhn–Tucker conditions under different constraint qualifications.

Zangwill Constraint Qualification [550]

It has been shown in Theorem 23.7 that $\overline{D}(X) \subseteq \mathscr{D}(X)$. Situations may then exist for which there are directions in $\mathscr{D}(X)$ that are not in $\overline{D}(X)$. In order to exclude such cases Zangwill assumed that

$$\mathscr{D}(X) = \overline{D}(X)$$

as the constraint qualification.

It then follows from Corollary 23.1 that for X_0 to be a local optimal solution to the problem,

$$\nabla f(X_0)^T d \geq 0, \text{ for all d satisfying}$$

$$\nabla g_i (X_0^T d \leq 0, \text{ for all } i \in I(X_0)$$

An application of Farkas' lemma, then yields the Kuhn–Tucker conditions.

Slater's Constraint Qualification [437]

Let the functions $g_i: R^n \to R^1$, $i = 1,2..m$ of the problem (23.9) be convex which define the convex constraint set

$$S = \{X | g_i(X) \leq 0, i = 1,2..m\}$$

Then the functions are said to satisfy Slaters' constraint qualification, if there exists an $\overline{X} \in S$, such that

$$g_i(\overline{X}) < 0, i = 1,2,..m.$$

Slater's Constraint Qualification for Problems with Inequality and Equality Constraints

Let in the problem (23.19), the function g_i be pseudoconvex at $\overline{X}$ for $i \in I(\overline{X}) = \{i | g_i(\overline{X}) = 0\}$, h_i for $i = m_1 + 1,..m$ is quasiconvex, quasiconcave and $\nabla h_i(\overline{X})$ for $i = m_1 \; 1,..m$ are linearly independent, then the functions are said to satisfy Slater's constraint qualification at $\overline{X}$ if there exists a feasible point X such that

$$g_i(X) < 0, \text{ for } i \in I(\overline{X})$$

$$\text{and} \quad h_i(X) = 0, \text{ for } i = m_1 + 1,..m.$$

Linear Independence Constraint Qualification

For the problem (23.19) $\nabla g_i(\overline{X})$ for $i \in I(\overline{X})$ and $\nabla h_i(\overline{X})$ for $i = m_1 + 1,..m$ are linearly independent.

Note that Slater's constraint qualification and linear independence constraint qualification imply Kuhn–Tucker constraint qualification.

Constraint qualifications are also suggested by Karlin [268], Arrow, Hurwicz and Uzawa [17], Cottle [82], Mangasarian and Fromovitz [322], Abadie [2], Evans [149] and others. The readers are referred to their individual work.

23.7. Lagrange Saddle Point Problem and Kuhn–Tucker Conditions

In this section, we show that a saddle point of the Lagrangian function associated with the nonlinear programming problem (if it exists) yields an optimal solution to the problem and that under certain conditions an optimal solution to the problem provides a saddle point of the associated Lagrangian function. We then derive the conditions for existence of a saddle point and its relation with the Kuhn–Tucker conditions.

Consider the problem,

$$\begin{aligned} &\text{Minimize} \quad f(X) \\ &\text{Subject to} \quad g_i(X) \leq 0,\ i = 1,2..m \\ &\qquad\qquad\qquad X \geq 0 \end{aligned} \tag{23.24}$$

where f(X), $g_i(X)$ i = 1,2.m are nonlinear differentiable functions of $X \in R^n$.

It can be shown that under suitable assumptions, the nonlinear program (23.24) can be transformed into an equivalent saddle point problem.

Saddle Point Problem: Let ϕ (X,U) be a real valued function of $X \in \mathfrak{X} \subset R^n$ and $U \in \mathcal{U} \subset R^m$. A point (X^0, U^0) is said to be a saddle point of ϕ (X, U) if

$$\phi(X^0, U) \leq \phi(X^0, U^0) \leq \phi(X, U^0), \text{ for all } X \in \mathfrak{X} \text{ and } U \in \mathcal{U} \tag{23.25}$$

In other words, a saddle point is a point (X^0, U^0) that minimizes the function ϕ (X, U) in $\mathfrak{X}$ for fixed $U^0 \in \mathcal{U}$ and maximizes the function in $\mathcal{U}$ for fixed $X^0 \in \mathfrak{X}$ simultaneously.

$v = \phi(X^0, U^0)$ is then called a saddle value, of ϕ (X, U)

The Lagrangian function associated with the nonlinear program (23.24) is given by,

$$\phi(X, U) = f(X) + U^T g(X) \tag{23.26}$$

where the vector $U \in R^m$ is called the vector of Lagrange multipliers and g(X) = $[g_1(X), g_2(X),...g_m(X)]^T$.

The corresponding saddle point problem is to find a pair (X^0, U^0), $X^0 \geq 0$, $U^0 \geq 0$ such that

$$f(X^0) + U^T g(X^0) \leq f(X^0) + U^{0T} g(X^0) \leq f(X) + U^{0T} g(X)$$
$$\text{for all } X \geq 0,\ U \geq 0 \tag{23.27}$$

Theorem 23.12. If (X^0, U^0) is a saddle point of the Lagrangian function ϕ (X, U), associated with the nonlinear program (23.24), then X_0 is an optimal solution to the problem.

Proof: Since (X^0, U^0) is a saddle point, from the left hand inequality in (23.27), we have,

$$U^T g(X^0) \leq U^{0T} g(X^0), \text{ for all } U \geq 0 \tag{23,28}$$

and hence the inequality is true for

$$U = U^0 + e_i, \text{ where } e_i \text{ is the ith mint m-vector}$$

Thus $$g_i(X^0) \leq 0. \tag{23.29}$$

Repeating this process for all i, we get

$$g(X^0) \leq 0. \tag{23.30}$$

Since $X^0 \geq 0$, (23.30) implies that X^0 is a feasible solution to (23.24)

Now, since $U^0 \geq 0$, from (23.30), we have

$$U^{0T} g(X^0) \leq 0. \tag{23.31}$$

But from (23.28) with U = 0, we have

$$U^{0T} g(X^0) \geq 0. \tag{23.32}$$

(23.31) and (23.32), then implies that

$$U^{0T} g(X^0) = 0. \tag{23.33}$$

The right hand inequality of (23.27) then becomes

$$f(X^0) \leq f(X) + U^{0T} g(X), \text{ for all } X \geq 0$$

and hence for all X feasible to (23.24).

Now, since $U^0 \geq 0$ and for feasible X, $g(X) \leq 0$,

we have $U^{)T} g(X) \leq 0$,

and hence $f(X^0) \leq f(X)$, for all X feasible to (23.24)

Thus, a saddle point of the Lagrangian is sufficient to locate an optimal solution to the nonlinear program, it is associated with. While no assumption was required to establish the sufficiency, we need to impose certain conditions on the functions in order to prove the necessary part.

Theorem 23.13. Let X^0 be an optimal solution to the nonlinear program (23.24) and assume that the Kuhn–Tucker constraint qualification holds. Further, let the differentiable functions f (X) and g_i (X), i = 1,2,..m are convex. Then there exists $U^0 \geq 0$ such that (X^0, U^0) is a saddle point of the associated Lagrangian function ϕ $(X, U) = f(X) + U^T g(X)$ so that

$$\phi(X^0, U) \leq (X^0, U^0) \leq \phi(X, U^0), \text{ for all } X \geq 0, U \geq 0$$

Proof: Since X^0 is an optimal solution to the nonlinear program and the Kuhn–Tucker constraint qualification is satisfied, Kuhn–Tucker conditions are applicable which implies that there exists $U^0 \geq 0$ such that

$$\nabla f(X^0) + U^{0T} \nabla g(X^0) \geq 0 \tag{23.34}$$

$$[\nabla f(X^0) + U^{0T} \nabla g(X_0)]^T X^0 = 0 \tag{23.35}$$

$$U^{0T} g(X^0) = 0 \tag{23.36}$$

Now, $\phi(X, U^0) = f(X) + U^{0T} g(X)$ is convex and

hence $\phi(X, U^0) \geq \phi(X^0, U^0) + \nabla x\, \phi(X^0, U^0)^T (X - X^0)$ for $X \geq 0$ (23.37)

But $\nabla x\phi(X^0, U^0) = \nabla f(X^0) + U^{0T} \nabla g(X^0) \geq 0$, by (23.34) (23.38)

Hence, $\nabla x..(X^0, U^0)^T (X - X^0) = \nabla x\phi(X^0, U^0)^T X - \nabla x\phi(X^0, U^0)^T X^0$

$= \nabla x\phi(X^0, U^0)^T X - 0$, by (23.35)

$= \nabla x\phi(X^0, U^0)^T X$

Now by (23.38), $\nabla x\phi(X^0, U^0) \geq 0$ and since $X \geq 0$, we have

$$\nabla x\phi(X^0, U^0)^T (X - X^0) \geq 0 \tag{23.39}$$

Hence by (23.37) $\phi(X, U^0) \geq \phi(X^0, U^0)$ (23.40)

Now since $\phi(X^0, U)$ is linear in U,

$$\phi(X^0, U) = \phi(X^0, U^0) + \nabla_0\phi\,(X^0, U^0)^T (U - U^0) \tag{23.41}$$

and since $\nabla_U\phi(X^0, U^0) = g(X^0)$

and $U^{0T} g(X^0) = 0$, by (2336)

$$\nabla_U\phi\,(X^0, U^0)^T (U - U^0) = \nabla_U\phi(X^0, U^0)^T U = g(X^0)^T U$$

Since $g(X^0) \leq 0$ and $U \geq 0$, $\nabla_U\phi(X^0, U^0)^T (U - U^0) \leq 0$ (23.42)

Hence $\phi\,(X^0, U) \leq (X^0, U^0)$ (23.43)

From (23.40) and (23.43), we then have

$$\phi(X^0, U) \leq \phi(X^0, U^0) \leq \phi(X, U^0)$$

and (X^0, U^0) is a saddle point.

Existence of a Saddle Point

We now derive the conditions for existence of a saddle point for a function $\phi(X, U)$, $X \geq 0$, $U \geq 0$.

Theorem 23.14. (Necessity) The conditions

$$\phi_x^0 = \left[\frac{\partial\phi}{\partial x_j}\right]_{x^0, U^0} \geq 0, \quad \phi_x^{0T} X^0 = 0, \quad X^0 \geq 0. \tag{23.44}$$

$$\phi_U^0 = \left[\frac{\partial\phi}{\partial u_i}\right]_{x^0, U^0} \leq 0, \quad \phi_U^{0T} U^0, \quad U^0 \leq 0 \tag{23.45}$$

are necessary for (X^0, U^0) to be a saddle point for any continuously differentiable function $\phi\,(X, U)$, $X \geq 0$, $U \geq 0$.

Proof: Since (X^0, U^0) to be a saddle point for $\phi(X, U)$, $\phi(X, U^0)$ has a local minimum at $X = X^0$. It may either be an interior point or a boundary point.

Hence the components of ϕ_x^0 must vanish except possibly when the corresponding components of X^0 vanish, in which case the conditions (23.44) must be satisfied.

A similar argument shows that the conditions (23.45) must also hold.

Theorem (23.15). (Sufficiency). If $\phi\,(X, U)$ is continuously differentiable function and $\phi(X, U^0)$ is convex in X and $\phi\,(X^0, U)$ is concave in U, then the conditions (23.44) and (23.45) are both necessary and sufficient for (X^0, U^0) to be a saddle point of $\phi\,(X, U)$.

Proof: Since $\phi(X, U^0)$ is convex in X, we have

$$\phi\,(X, U^0) \geq \phi\,(X^0, U^0) + \phi_x^{0T}\,(X - X^0)$$

$$\geq \phi(X^0, U^0) \text{ (since } \phi_x^0 \geq 0, X \geq 0, \phi_x^{0T} X^0 = 0)$$

$$\geq \phi(X^0, U^0) + \phi_x^{0T} (U - U^0) \text{ (since } \phi_U^0 \leq 0, U \geq, 0, \phi_x^{0T} U^0 = 0)$$

$$\geq \phi(X^0, U), \qquad \text{for all } X \geq 0, U \geq 0$$

since $\phi(X^0, U)$ is concave in U.

Hence $\phi(X^0, U) \leq \phi(X^0, U^0) \leq \phi(X, U^0)$, for all $X \geq 0$, $U \geq 0$.

From the results obtained above, it should be noted that the Kuhn–Tucker necessary and sufficient conditions for X_0 to be an optimal solution to the problem

$$\begin{array}{ll} \text{Minimize} & f(X) \\ \text{Subject to} & g_i(X) \leq 0, \quad i = 1,2..m \\ & X \geq 0 \end{array}$$

where f(X) and $g_i(X)$ are continuously differentiable convex functions can be expressed as

$$\phi_x^0 \geq 0, \ \phi_x^{0T} X^0 = 0, \ X^0 \geq 0.$$

$$\phi_U^0 \leq 0, \ \phi_U^{0T} U^0 = 0, \ U^0 \geq 0.$$

where $\phi(X, U) = f(X) + U^T g(X)$, $X \geq 0$, $U \geq 0$

23.8. Exercises

1. Find all local maxima and minima and the global maximum and minimum of

   (i) $f(x_1, x_2) = x_1^2 + 3x_1x_2 + x_2^2$

   (ii) $f(x_1, x_2) = (x_1 - 1)^2 e^{x_2} + x_1$.

2. Show that for the problem minimize $f(x) = (x^2 - 1)^3$, $x = 0$ is the global minimum.

3. Determine the values of a and b so that $f(x) = x^3 + ax^2 + bx$, has

   (i) a local maximum at $x = -1$ and a local minimum at $x = +1$

   (ii) a local maximum at $x = 1$ and a local minimum at $x = 0$.

4. Consider the constrained optimization problem and the sets D, $\overline{D}$ and $\mathscr{D}$ as defined in sections 23.3 and 23.4. Show that Zangwill's constraint qualification $\mathscr{D}(x) = \overline{D}(X)$ implies Kuhn–Tucker constraint qualification.

5. Consider the constraint set

$$x_1 + x_2 \leq 1$$
$$x_1 \geq 0, x_2 \geq 0$$

   Is $\overline{D}(X_0) = \mathscr{D}(x_0)$ at the feasible point $X_0 = \left(\frac{1}{2}, \frac{1}{2}\right)^T$?

6. Show that Slater's constraint qualification and linear independence constraint qualification imply Kuhn–Tucker constraint qualification.
7. Prove that if in a constrained optimization problem, all constraints are linear, then Kuhn–Tucker constraint qualification holds.
8. Solve the problem

$$\text{Minimize} . \sum_{j=i}^{n} \frac{c_j}{x_j}$$

$$\text{Subject to } \sum_{j=i}^{n} a_j x_j = b$$

$$x = 0, \qquad j = 1,2,..n$$

where c_j, a_j and b are positive constants, by finding a point X_0 that satisfies the Kuhn–Tucker conditions.

9. By using the Kuhn–Tucker conditions, establish the well-known result: If $x_1, x_2,..x_n \geq 0$, then

$$\left(\prod_{j=l}^{n} x_j\right)^{\frac{1}{n}} \leq \frac{1}{n}\sum_{j=i}^{n} x_j$$

[Hint: Consider the problem: Minimize $\prod_{j=l}^{n} x_j$,

Subject to $\sum_{j=i}^{n} x_j = 1$, $x_j \geq 0$, $j = 1,2..n$]

10. Obtain the Kuhn–Tucker conditions for the solution of the problem

$$\begin{aligned} &\text{Minimize} && c^T X + 1/2\, X^T B X \\ &\text{Subject to} && AX = b. \\ & && X \geq 0 \end{aligned}$$

where c is an n vector, b an m vector, A an m x n matrix and B is an n x n symmetric positive semidefinite matrix.

11. Show that the solution of the pair of dual linear programs

$$\begin{aligned} & \text{Maximize} && c^T X \\ & \text{Subject to} && AX \leq b \\ & && X \geq 0 \\ \text{and} \quad & \text{Minimize} && b^T U \\ & \text{Subject to} && A^T U \geq c \\ & && U \geq 0 \end{aligned}$$

where c is an n vector, b an m vector and A is m x n matrix is equivalent to the solution of the saddle point problem for the bilinear function ϕ (x, u).

CHAPTER 24

Quadratic Programming

24.1. Introduction

Quadratic programming is an important class of convex programming in which a convex quadratic function is to be minimized subject to linear constraints. Such problems arise in various contexts. Some of them referred by Wolfe [524] are as follows.

(a) Regression [554]

To find the best least-square fit to given data, where certain parameters are known a prior to satisfy linear inequality constraints.

Consider the regression model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k ,$$

where β_j (j = 0,1,...k) are the regression parameters.

Let y_i be the ith observation on the dependent variable corresponding to the given values x_{ij} of the independent variables, so that

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_k x_{ik} + e_i , (i = 1,2..n)$$

where e_i is the error by which an observation falls off the true regression.

In order to have a best fit to the model, the estimates of β's are to be so determined that these errors are as small as possible. A satisfactory method of determining the estimates is the method of least squares which consists in minimizing the sum of squares of the errors of estimation. Thus, we need to

$$\text{Minimize} \quad \sum_{j=0}^{n} e_i^2 = \sum_{i}^{n} (y_i - \beta_0 - \beta_1 x_{i1} + ... + \beta_k x_{ik})^2$$

which is a quadratic function in the variables $\beta_0, \beta_1, .. \beta_k$.

It is however quite common in econometrics and in many other practical applications to impose certain linear constraints on the parameters of β's.

$$\sum_{j=0}^{k} a_{rj}\beta_j = b_r, \qquad r = 1, 2...m$$

$$\beta_j \geq 0, \qquad j = 0, 1, 2, \ldots k.$$

We then minimize a quadratic function subject to linear constraints, which is a quadratic programming problem.

(b) Efficient Production [127]

Suppose that an entrepreneur wants to maximize his profit from the sale of the n commodities he produces. Suppose that p_j, the price per unit for the jth product, decreases linearly as the output x_j of the jth product increases.

Thus $\qquad p_j = a_j - b_j x_j, \; j = 1,2 \ldots n.$

where $a_j > 0$, $b_j > 0$.

With the usual linear constraints regarding availability of resources employed, the problem reduces to

$$\text{Maximize} \quad \sum_{j=1}^{n} (a_j - b_j x_j) x_j .$$

$$\text{Subject to} \sum_{j=1}^{n} a_{ij} x_j \leq b_i \;, \quad i = 1,2..m$$

$$x_j \geq 0, \quad j=1,2..n.$$

which is again a quadratic programming problem

(c) Portfolio Selection [328]

Suppose that an investor wishes to invest a fixed amount of money in n different securities and is faced with the problem of selecting a portfolio (combination of securities) which will yield at least a given expected return α with minimum risk.

Let x_j, be the proportion of money invested in the jth security

p_j, be the variable return from the jth security with the estimated mean μ_j and the variances and covariances σ_{jk} (j, k = 1,2..n)

The problem then is to determine an investment portfolio (x_1, x_2,. x_n), which

$$\text{Minimize} \quad \sum_{j=1}^{n} \sum_{k=1}^{n} \sigma_{jk} x_j x_k$$

$$\text{Subject to} \quad \sum_{j=1}^{n} x_j = 1$$

$$\sum_{j=1}^{n} \mu_j x_j \geq \alpha$$

$$x_j \geq 0, \qquad j=1,2..n$$

which is a quadratic programming problem.

(d) Convex Programming [514]

To find the minimum of a general convex function under linear constraints using a quadratic approximation

Several methods have been developed for solving quadratic programming problems. An exposition of them has been given in the books by Boot [60]; Kunzi, Krelle and Oettli [294].

It was Barankin and Dorfman [26] who first pointed out that if the quadratic program has a solution then beginning with an arbitrary basic solution of the linear system of the Kuhn–Tucker conditions for the problem, it is possible to find one which satisfy the complimentary condition and hence provides a solution of the problem. Markowitz [327] on the other hand, suggested a method which begins with finding a solution of a modified linear system and the complementary condition and then retaining the complementary condition, alter the variables until the linear system is also satisfied, thus giving the desired solution. Finally, Wolfe [524] showed that this can be achieved by slightly modifying the simplex algorithm.

24.2. Wolfe's Method

Wolfe [524] modified the simplex method to solve the quadratic programming problem

$$\begin{aligned} \text{Minimize } f(\lambda, X) = \lambda\, p^T X + 1/2\, X^T CX \\ AX = b. \\ X \geq 0 \end{aligned} \tag{24.1}$$

where λ is a single nonnegative parameter that can be chosen as convenient, p,X are n–vectors, C is an $n \times n$ symmetric positive semidefinite matrix, A an $m \times n$ matrix and $b \geq 0$, is an m–vector.

Without loss of generality, equality constraints have been assumed in (24.1) as inequality constraints can easily be converted into equality constraints by addition of additional variables. Elements of b can then be assumed to be nonnegative.

Lemma 24.1: The function $f(\lambda, X)$ is convex

Proof: Since C is positive semidefinite X^TCX is Convex, λp^TX being linear may be considered to be convex. Further, since the sum of two convex functions is convex, $f(\lambda,X)$ is convex. This ensures that any local minimum encountered in the problem is a global minimum as the constraint set is convex.

Theorem 24.1. The necessary and sufficient conditions for an $X \geq 0$ to be a solution to the quadratic programming problem (24.1) are that there exist a U and a $V \geq 0$, such that X, U, V satisfy

$$\begin{aligned} V^T X = 0 \\ AX = b \\ CX - V + A^T U + \lambda p = 0 \end{aligned} \tag{24.2}$$

Proof: Since the constraints of the problem (24.1) are linear, the Kuhn–Tucker constraint qualification is automatically satisfied and since $f(\lambda, X)$ is convex, the Kuhn–Tucker conditions for a solution to the problem are both necessary and sufficient.

Now, following the notations of Section 23.7 the Kuhn–Tucker conditions can be stated as

$$\phi_x \geq 0, \quad \phi_x^T X = 0, X \geq 0$$

$$\phi U_1 \leq 0, \phi U_2 \leq 0 \phi_{U1}^T U_1 = 0, \phi_{U_3}^T U_2 = 0, U_1 \geq 0, U_2 \geq 0.$$

where $\phi(X, U_1 U_2) = \lambda P^T X + \frac{1}{2} X^T C X + U_1^T (AX - b) + U_2^T (-AX + b)$

We then have

$$\phi_X = \lambda p + CX + A^T U_1 - A^T U_2 \geq 0 \tag{24.3}$$

$$\phi_x^T X = [\lambda p + CX + A^T U_1 - A^T U_2]^T X = 0 \tag{24.4}$$

$$X \geq 0. \tag{24.5}$$

$$\phi U_1 = AX - b \leq 0 \tag{24.6}$$

$$\phi U_2 = -AX + b \leq 0 \tag{24.7}$$

$$U_1 \geq 0, U_2 \leq 0 \tag{24.8}$$

(24.6), (24.7) imply that $AX = b$

and then $\phi_{U1}^T U_1 = \phi_{U2}^T = 0.$

are automatically satisfied.

Let $V = \phi_X$ and $U = U_1 - U_2$.

The conditions are then reduced to

$$\begin{gathered} V^T X = 0 \\ AX = b \\ CX - V + A^T U + \lambda p = 0 \\ X \geq 0, V \geq 0. \end{gathered} \tag{24.9}$$

24.2.1. The Computation

Wolfe developed the computational algorithms for solving the quadratic programming problem in two separate forms: first the "short form", when λ is fixed and the next the "long form", when the problem is solved parametrically for all $\lambda \geq 0$. While the convergence of the process in the short form requires that either $\lambda = 0$ or that C is positive definite, for the long form no such restriction is needed

The Short Form: λ fixed

Consider the problem

$$\begin{aligned} \text{Minimize} \quad & e^T_{(m)}W \\ \text{Subject to} \quad & AX + W = b. \\ & CX - V + A^TU + Z_1 - Z_2 = -\lambda p. \\ & X, V, W, Z_1, Z_2 \geq 0 \end{aligned} \tag{24.10}$$

where $e_{(m)}$ is an m-component vector with each element equal to unity and m-component vector W and n-component vectors Z_1, Z_2 are vectors of additional variables introduced in the system.

It is assumed that the constraints in (24.10) are nondegenerate.

Since $b \geq 0$, an initial basis for the system (24.10) can be formed from the coefficients of W, Z_1 and Z_2. The problem (24.10) with U = 0, V = 0, that is the problem

$$\begin{aligned} \text{Minimize} \quad & e^T_{(m)}W \\ \text{Subject to} \quad & AX + W = b \\ & CX + Z_1 - Z_2 = -\lambda P \\ & X, W, Z_1, Z_2 \geq 0 \end{aligned} \tag{24.11}$$

is then solved by the simplex method. If the quadratic program (24.1) is feasible, the minimum obtained for $e^T_{(m)}W$ will be equal to zero and the condition $V^TX = 0$ is also satisfied.

From the optimal solution of (24.11) we discard W and the unused components of Z_1 and Z_2 and denote the remaining n components by Z and their coefficients by E. E is then a diagonal matrix with elements +1 or −1 depending on whether $z_j = z_{1j}$ or $z_j = z_{2j}$.

We thus have a basic feasible solution of the system

$$\begin{aligned} AX &= b \\ CX - V + A^TU + EZ &= -\lambda p \\ X, V, Z &\geq 0 \end{aligned} \tag{24.12}$$

with U = 0, V = 0 so that the condition $V^TX = 0$ is satisfied.

With this basic solution, we now initiate the simplex method to

$$\begin{aligned} \text{Minimize} \quad & e^T_{(n)} Z \\ \text{Subject to} \quad & (24.12) \end{aligned} \tag{24.13}$$

under the side condition that if x_k is in the basis, v_k is not to be entered into the basis and if v_k is in the basis, x_k is not to be entered into the basis (k = 1, 2,...n).

(24.14)

The side condition (24.14) ensures that at each stage of the simplex iteration we have $V^TX = 0$.

If the iterative process terminates with $e^T_{(n)} Z = 0$, we have a solution of (24.2) and then the X-part of the solution is a solution for the quadratic programming problem for a given λ.

However, it may happen that under the side conditions (24.14), it is not possible to reduce $e^T_n Z$ to zero unless it is assumed that either C is positive definite or $\lambda = 0$. To establish this statement we prove the following theorem.

Theorem 24.2: Let A be an $m \times n$ matrix, b an $m\times 1$ and C be an $n \times n$ symmetric positive semidefinite matrix. Let Q be $n \times h$ matrix, q be an h component vector and g an n–vector. Let $X \geq 0$, $V \geq 0$ such that $V^TX = 0$ be given. Further, let X_x denote the vector whose components are the positive components of X and V_x, the vector whose components are the corresponding components of V ($V_x = 0$). Similarly, let V_v denote the vector whose components are the positive components of V and X_v, the vector whose components are the corresponding components of $X(X_v = 0)$.

If the linear form

$$q^TW \tag{24.15}$$

is minimal under the linear constraints

$$V_x = 0 \tag{24.16}$$

$$\begin{aligned} &X_v = 0 \\ &AX \quad = b \\ &CX - IV + A^TU + QW = g \\ &X \geq 0,\ V \geq 0,\ W \geq 0. \end{aligned} \tag{24.17}$$

then there exists an no vector r such that

$$Cr = 0\ ,\ Ar = 0 \text{ and } q^TW = r^Tg.$$

Proof: We have already distinguished in the vectors X,V, the corresponding parts $X_x>0$, $V_x = 0$ and the corresponding parts $X_v = 0$, $V_v > 0$. After a possible reordering of indices, the vectors X and V are partitioned as follows:

$$X^T = (X^T_x, X^T_\delta, X^T_v)$$

$$V^T = (V^T_x, V^T_\delta, V^T_v) \tag{24.18}$$

where X_δ and V_δ are the remaining parts of X and V respectively.

Similarly, A, C, Q and g are partitioned as

$$A=(A_xA_\delta A_v), \qquad 24.19)$$

$$C = \begin{pmatrix} C_{xx} & C_{x\delta} & C_{xv} \\ C^T_{x\delta} & C_{\delta\delta} & C_{\delta v} \\ C^T_{xv} & C^T_{\delta v} & C_{vv} \end{pmatrix} \tag{24.20}$$

$$g^T = \left(g^T_x, g^T_\delta, g^T_v\right) \tag{24.21}$$

$$Q^T = (Q^T_x, Q^T_\delta, Q^T_v) \tag{24.22}$$

Note that X_x and V_v contain all the positive elements of X and V respectively and thus X_δ and V_δ consist of variables with zero values, i.e. $X_x = V_v = 0$ in the solution of the program.

The constraints (24.17) can then be expressed as

$$
\begin{aligned}
&A_xX_x+A_\delta X_\delta = b\\
&C_{xx}X_x+C_{x\delta}X_\delta+A_x^TU+Q_xW=g_x\\
&C_{x\delta}^TX_x+C_{\delta\delta}X_\delta-I_\delta V_\delta+A_\delta^TU+Q_\delta W=g_\delta \qquad (24.23)\\
&C_{xv}^TX_x+C_{\delta x}^TX_\delta-I_vV_v+A_v^TU+Q_vW=g_v\\
&X_x\geq 0,\ X_\delta\geq 0,\ V_\delta\geq 0,\ V_v\geq 0,\ W\geq 0
\end{aligned}
$$

X_v and V_x are not included in the formulation above since they have zero values.

According to the hypothesis of the theorem, the values of the variables $X_x > 0$, $X_\delta \geq 0$, $V_\delta \geq 0$, $V_v > 0$, $W \geq 0$ and U (Unrestricted) minimizes the linear form q^TW.

By the duality theorem in linear programming then, there exists a solution of the dual to this problem and their objective functions have the same value.

The dual program can be stated as,

$$
\begin{aligned}
\text{Maximize}\quad & b^TY+g_x^Tr_x+g_\delta^Tr_\delta+g_v^Tr_v.\\
\text{Subject to}\quad & A_x^TY+C_{xx}r_x+C_{x\delta}r_\delta+C_{xv}r_v=0 \quad (a)\\
& A_\delta^TY+C_{x\delta}^Tr_x+C_{\delta\delta}r_\delta+C_{\delta v}r_v\leq 0 \quad (b)\\
& -I_\delta r_\delta \leq 0 \quad (c) \qquad (24.24)\\
& -I_vr_v = 0 \quad (d)\\
& A_xr_x+A_\delta+r_\delta+A_vr_v = 0 \quad (e)\\
& Q_x^Tr_x+Q_\delta^Tr_\delta+Q_v^Tr_v\leq q \quad (f)
\end{aligned}
$$

and we have $$q^TW=b^TY+g_x^Tr_x+g_\delta^Tr_\delta+g_v^Tr_v \qquad (24.25)$$

The dual variables are unrestricted in sign, since the constraints (24.23) are all equations. Moreover, when a variable in (24.23) is positive as in X_x, V_v or is unrestricted as in U, the corresponding dual constraint is an equality.

From (24.24c) and (24.24d) it immediately follows that $r_\delta \geq 0$ and $r_v = 0$. Now, multiplying (24.24a) on the left by r^T_x and (24.24b) by r^T_δ, we obtain

$$
\begin{aligned}
&r_x^TA_x^TY+r_x^TC_{xx}r_x+r_x^TC_{x\delta}r_\delta=0\\
&r_\delta^TA_\delta^TY+r_\delta^TC_{\delta x}^Tr_x+r_\delta^TC_{\delta\delta}r_\delta\leq 0 \qquad (24.26)
\end{aligned}
$$

Then adding these two expressions and taking the transpose, we get

$$Y^T(A_xr_x+A_\delta r_\delta)+\left(r_x^Tr_\delta^T\right)\begin{bmatrix}C_{xx} & C_{x\delta}\\ C_{x\delta}^T & C_{\delta\delta}\end{bmatrix}\begin{pmatrix}r_x\\ r_\delta\end{pmatrix}\leq 0 \qquad (24.27)$$

The first term in (24.27) vanishes by (24.24e) and the matrix in the second term being a principal submatrix of the positive semidefinite matrix C is itself positive semidefinite, so that the second term cannot be negative. The only possibility is that

$$\left(r_x^T r_\delta^T\right)\begin{bmatrix} C_{xx} & C_{x\delta} \\ C_{x\delta}^T & C_{\delta\delta} \end{bmatrix}\begin{pmatrix} r_x \\ r_\delta \end{pmatrix} = 0.$$

which implies that

$$\begin{bmatrix} C_{xx} & C_{x\delta} \\ C_{x\delta}^T & C_{\delta\delta} \end{bmatrix}\begin{pmatrix} r_x \\ r_\delta \end{pmatrix} = 0. \tag{24.28}$$

or

$$C_{xx} r_x + C_{x\delta} r_\delta = 0 \tag{24.29}$$

$$C_{x\delta}^T r_x + C_{\delta\delta} r_\delta = 0 \tag{24.30}$$

From (24.29) and (24.24a), we obtain

$$A^T{}_x Y = 0 \tag{24.31}$$

and therefore,

$$b^T Y = \left[X_x^T A_x^T + X_\delta^T A_\delta^T + X_v^T A_v^T\right] Y \tag{24.32}$$

$= X_x^T A_x^T Y = 0$, (since $X_\delta = 0$ by definition)

Since $r_v = 0$, we may generalize (24.28) to have

$$\begin{bmatrix} C_{xx} & C_{x\delta} & C_{xv} \\ C_{x\delta}^T & C_{\delta\delta} & C_{\delta v} \\ C_{xv}^T & C_{\delta v}^T & C_{vv} \end{bmatrix}\begin{pmatrix} r_x \\ r_\delta \\ C_{vv} \end{pmatrix} = 0 \tag{24.33}$$

Setting $r^T = \left(r_x^T, r_\delta^T, r_v^T\right)$ we have from (24.33), (24.24e) and (24.32), (24.25)

$$Cr = 0,\ Ar = 0,\ q^T W = r^T g. \tag{24.34}$$

This proves the theorem.

Coming back to the short form computation, we note that if in problem (24.13), it is not possible to reduce $e^T{}_n Z$ to zero under the side condition, the hypothesis of theorem 24.2 will be satisfied with

$$Q = E,\ q^T = (1.1....1),\ W = Z \text{ and } g = -\lambda p.$$

and we will have

$$e^T{}_{(n)} Z = q^T W = r^T g = -\lambda\, r^T p \tag{24.35}$$

with $\quad Cr = 0.$

In order to have Min $e^T{}_{(n)} Z = 0$, therefore, we should either have $\lambda = 0$ or C positive definite, when necessarily $r = 0$ and then by theorem 24.1 X-part of the terminating solution will be a solution of the quadratic program (24.1)

We however, observe that the short form computation can work with positive semidefinite quadratic form if the diagonal elements of C are perturbed to change C to $(C+\epsilon I)$ for a small $\epsilon > 0$ so that $X^T (C + \epsilon I) X > 0$ and the algorithm can be operated as if a positive definite form were employed. The perturbation can be made small enough so that the numerical results obtained are not affected.

The Long Form

In order to obtain solutions for the quadratic programming problem (24.1) for all values of $\lambda \geq 0$, we first obtain a solution for $\lambda = 0$.

Having performed the short form computation for $\lambda = 0$, we obtain a basic feasible solution of the system

$$\begin{aligned} &AX && = b \qquad (24.36)\\ &CX - V + A^T U + EZ && = 0 \\ &X \geq 0, v \geq 0 \end{aligned}$$

with $\qquad Z = 0$ and $V^T X = 0$

We then apply the simplex method to find a solution of the problem,

$$\begin{aligned} \text{Minimize} \quad & -\lambda && (24.37)\\ \text{Subject to} \quad & AX = b \\ & CX - V + A^T U + \lambda p = 0 && (24.38)\\ & X, V \geq 0, \lambda \geq 0. \end{aligned}$$

with the side condition $V^T X = 0$

As indicated in (24.36), an initial basic feasible solution of the problem (24.37), (24.38) having $\lambda = 0$ and $V^T X = 0$ is provided by the short form computation with $\lambda = 0$.

Two cases may arise

Case (i): $-\lambda$ has a finite minimum

Case (ii): $-\lambda$ is unbounded below.

Case (i): If $-\lambda$ has a finite minimum, the hypothesis of theorem 24.2 is satisfied with $W = \lambda$, $Q = p$, $q = -1$ and $g = 0$. We then have

$$\text{Min } (-\lambda) = q^T W = r^T g = 0$$

which means that $-\lambda$ has infact not been reduced.

Now, under the assumption of nondegeneracy of the constraints of (24.38), every basic solution of this system has exactly (m + n) positive variables. Since $\lambda = 0$, m variables in U are always in the basis, the remaining n variables are in X_x and V_v (X_δ, V_δ are empty). From Theorem 24.2 we know that there exists a vector r such that $Ar = 0$, $Cr = 0$ so that for any t, we have

$$\begin{aligned} &A(X + tr) = b. && (24.39)\\ &C(X + tr) - V + A^T U = 0 \end{aligned}$$

It follows from nondegeneracy that $r = (r_X, r_\delta, r_v) \geq 0$ for otherwise if at least one element of $r_X (r_\delta, r_v = 0)$ is negative, there would exist a $t > 0$ such that

$X + tr$ will vanish for at least one more component than X does. In that case, $(X + tr, V, U, \lambda = 0)$ would also be a solution of (24.38) with fewer than $(m + n)$ positive variables, violating the nondegeneracy assumption.

Thus $(X + tr)$ is feasible for the quadratic program (24.1) for all $t \geq 0$ and

$$f(\lambda, X + tr) = \lambda p^T(X + tr) + 1/2(X + tr)^T C(X + tr)$$
$$= \lambda p^T X + \lambda^t p^t r + 1/2\, X^T CX$$

Since by (24.24f), $p^T r \leq -1$,

$$f(\lambda, X + tr) \to -\infty \text{ as } t \to \infty \text{ for any } \lambda > 0.$$

Case (ii): Since λ is unbounded, by the theory of the simplex method (see Chapter 11) we have a sequence of basic solutions $(X^i, V^i, U^i, \lambda^i)$, $i = 1,2,..k$ for the system (24.38) and finally $(X^{k+1}, V^{k+1}, U^{k+1})$ such that $(X^k + tX^{k+1}, V^k + tv^{k+1}, U^k + tU^{k+1}, \lambda^k + t)$ is a solution for all $t \geq 0$, where the side condition ensures that

$$0 = V^i X^i = V^i X^{i+1} = V^{i+1} X^i = V^{i+1} X^{i+1} \text{ and } \lambda^i < \lambda^{i+1} \qquad (24.40)$$
$$i = 1,2,..k.$$

Thus X^i, V^i, U^i, λ^i is also a solution of the system (24.2) and hence X^i is a solution of the original quadratic programming problem for $\lambda = \lambda^i$, $(i = 1, 2..k)$

Now, for any λ, $\lambda^i \leq \lambda \leq \lambda^{i+1}$, $i = 1, 2, ...k-1$

$$X = \frac{\lambda^{i+1} - \lambda}{\lambda^{i+1} - \lambda^i} X^i + \frac{\lambda - \lambda^i}{\lambda^{i+1} - \lambda^i} X^{i+1} \qquad (24.41)$$

being a convex combination of X^i and X^{i+1} and letting V and U the same convex combination of V^i, V^{i+1} and U^i, U^{i+1} respectively, (X,V, U) is feasible for (24.38) and therefore also satisfies (24.2) and hence X is a solution of the quadratic program (24.1) for the corresponding λ.

For $\lambda > \lambda^k$,

$$\begin{aligned} X &= X^k + (\lambda - \lambda^k)\, X^{k+1}, \\ V &= V^k + (\lambda - \lambda^k)\, V^{k+1} \\ U &= U^k + (\lambda - \lambda^k)\, U^{k+1} \end{aligned} \qquad (24.42)$$

satisfies Theorem 24.2 and X is therefore a solution of the quadratic program for the corresponding λ $(\lambda > \lambda^k)$.

24.3. Dantzig's Method

Dantzig has suggested a variant of Wolfe's method for quadratic programming. The chief difference is that Dantzig's algorithm is more nearly a strict analogue of the simplex method. It has a tighter selection rule and the value of the objective function decreases monotonically. Therefore, no solution can reappear and hence the solution is obtained in a finite number of steps. It is believed to be computationally more efficient because there can be a greater decrease in the value of the quadratic function in each iteration.

Consider the problem

$$\begin{aligned} &\text{Minimize} && 1/2\, X^TCX \\ &\text{Subject to} && AX = b \\ &&& X \geq 0 \end{aligned}$$

where C is an n × n symmetric positive semidefinite matrix.

The Kuhn–Tucker conditions for an optimal solution of the problem are

$$\begin{aligned} AX &= b \\ CX - V + A^TU &= 0 \\ V^TX &= 0 \\ X, V &\geq 0 \end{aligned} \tag{24.44}$$

It is assumed that all basic solution of the above system are nondegenerate.

The procedure initiates with a basic feasible solution to the first two conditions of the system with $x \geq 0$ only (V, U not restricted) such that $V^TX=0$ at each iteration. The alogorithm then proceeds to find a solution which satisfies the restriction $V \geq 0$ and hence it yields an optimal solution to the given problem.

The Algorithm

Step 1. Let $X^T = [X^T_B X^T_N]$ be a basic feasible solution to $AX = b$, $X \geq 0$ and let $V^T = [V^T_B, V^T_N]$. Consider the basic solution to the enlarged system (24.44) with the basic vectors X_B, V_N, U which satisfies all the constraints except possibly $V \geq 0$.

Step 2. If $V \geq 0$, stop. The current solution is optimal. Otherwise, determine minimum $\{v_i \mid v_i < 0\} = v_j$. Go to step 3.

Step 3. Introduce x_j into the basis. If v_j drops, repeat step 2. Otherwise, if x_r drops, go to step 4.

Step 4. Introduce v_r into the basis. If v_j drops, go to step 2. If another variable x_k drops, repeat step 4 with v_r replaced by v_k.

The algorithm terminates in a finite number of steps.

For details see Dantzig [108, 109].

24.4. Beale's Method

Beale [37, 39] has developed a method for solving quadratic programming problems which is based on the basic principles of the simplex method rather than the Kuhn–Tucker conditions.

Consider the problem

$$\begin{aligned} &\text{Minimize} && F(X) = p^TX + X^TCX. \\ &\text{Subject to} && AX = b. \\ &&& X \geq 0 \end{aligned} \tag{24.45}$$

where $A = (a_{ij})$ is an m×n matrix, b is m×1, p is n × 1, X is n × 1 and C is an n × n symmetric matrix.

The symmetric matrix C however need not be positive definite or positive semi-definite. Thus Beale's method finds a local minimum of a nonconvex quadratic function, of course if the objective function is convex, the local solution obtained will be global. It is assumed that the problem is nondegenerate.

Beale's method is an iterative procedure and begins with any basic feasible solution of (24.45). Let A be partitioned as

$$A = (B, N)\,, \tag{24.46}$$

where B is the basis matrix which for convenience is assumed to consist of the first m-columns of A (by reordering the columns of A, if necessary) and N is the matrix consisting of the remaining n–m columns of A. Also let X_B and X_N be the vectors of basic and nonbasic variables.

The constraints AX = b, can then be written as

$$BX_B + NX_N = b.$$

and

$$X_B = B^{-1}b - B^{-1}NX_N \tag{24.47}$$

The basic variables can thus be expressed as

$$x_h = \alpha_{h0} + \sum_{q-1}^{n-m} \alpha_{hq} z_q, \quad h = 1,2..m \tag{24.48}$$

where $z_q = x_{m+q}$, q = 1,2.., n – m are the nonbasic variables.

In the present trial solution of the problem, therefore the basic variables x_h are equal to $\alpha_{h0} > 0$ and the nonbasic variables are all zero.

Using the equations (24.48), the objective function F(X) can now be expressed in terms of the nonbasic variables and to examine the effect on F of changing the value of the nonbasic variables, the partial derivatives with respect to each of the nonbasic variables, assuming that all the other nonbasic variables remain fixed and equal to zero, are considered.

For convenience, F(X) is expressed in the symmetric form

$$F(X) = \sum_{k=0}^{n-m}\sum_{q=0}^{n-m} \gamma_{kq} z_k z_q$$

$$= \gamma_{00} + 2\sum_{k=1}^{n-m} \gamma_{k0} z_k + \sum_{k=1}^{n-m}\sum_{q=1}^{n-m} \gamma_{kq} z_k z_q \tag{24.49}$$

where $z_0 = 1$ and $\gamma_{kq} = \gamma_{qk}$, k, q = 0,1,2..n–m.

Then $$\frac{1}{2}\frac{\partial F}{\partial z_k} = \gamma_{k0}, \quad k = 1,2,..n-m \tag{24.50}$$

If $\gamma_{k0} \geq 0$, a small increase in z_k with the other nonbasic variables held equal to zero will not reduce F.

However, if $\gamma_{k0} < 0$, the value of F can be reduced by a small increase in z_k. It is profitable to go on increasing z_k until either

(a) one of the basic variables becomes zero (a further increase in z_k will make the basic variable negative) or

(b) the partial derivative $F/\partial z_k$ changes its sign that is, when it vanishes and is about to become positive.

Case (a): Suppose that z_k is increased and one of the basic variables say x_v vanishes before $\partial F/\partial z_k$ does (x_v becomes nonbasic in place of x_k.) Then the equation

$$x_v = \alpha_{v0} + \sum_{q=1}^{n-m} \alpha_{vq} z_q \tag{24.51}$$

is used to express z_k in terms of x_v and the other nonbasic variables and is then substituted in (24.48) and (24.49) to express other basic variables and the objective function F in terms of the new set of nonbasic variables.

Case (b): If $\partial F/ z_k$ vanishes before any basic variable becomes negative, a new nonbasic variable

$$u_t = \gamma_{k0} + \sum_{q=1}^{n-m} \gamma_{kq} z_q = \frac{1}{2}\frac{\partial F}{\partial z_k} \tag{24.52}$$

is introduced into the problem, where the subscript 't' indicates that this is the tth such variable introduced during the iterative process. Using equation (24.52), we now express the new basic variable z_k in terms of u_t and the other nonbasic variables and is then substituted in (24.48) and (24.49) to express other basic variables and the objective function F in terms of the new set of nonbasic variables. Since z_k is now a basic variable and no former basic variable leaves the basis, there will be one more basic variable than before.

Since u_t is not restricted to have nonnegative values, it is called a free variable to distinguish it from the original x-variables which are called restricted variables.

Thus, if $\partial F/\partial u_t > 0$, then F can be reduced by making u_t negative. If a free variable enters the basis, it can be disregarded, as soon as it has been eliminated from the equations for the basic variables and for F. The process is repeated until a point is reached where it is not possible to change any of the nonbasic variables to have a further decrease in the value of the objective function. Thus, the conditions for termination of the iterative process are.

$$\frac{\partial F}{\partial z_k} \geq 0. \quad \text{for all k} \tag{24.53}$$

$$\frac{\partial F}{\partial u_t} = 0 \quad \text{for all free variable .}$$

which constitute necessary conditions for a local minimum of the problem. If however the objective function F(X) is convex, that is the matrix C is positive or positive semidefinite, the necessary conditions are also sufficient for a global minimum.

For the proof that the process terminates in a finite number of steps, see Beale [37,39]

24.4.1. Summary of Beale's Algorithm

The algorithm deduced from the preceding discussion on Beale's method for solving quadratic programming problems may now be summarized as follows:

Step 1. Obtain a basic feasible solution of the problem.

Step 2. Express the basic variables x_h, h = 1,2...m in terms of the nonbasic variables $z_q = x_{m+q}$, q = 1,2.. n−m

We will then have,

$$x_h = \alpha_{h0} + \sum_{q=1}^{n-m} \alpha_{hq} z_q.$$

Step 3. Express the objective function F(X) in terms of the nonbasic variables in the symmetric form

$$F(X) = \sum_{k=0}^{n-m}\sum_{q=0}^{n-m} \gamma_{kq} z_k z_q$$

$$= \gamma_{00} + 2\sum_{k=1}^{n-m} \gamma_{kq} z_k + \sum_{k=1}^{n-m}\sum_{q=1}^{n-m} \gamma_{kq} z_k z_q$$

where $z_0 = 1$ and $\gamma_{kq} = \gamma_{kq}$.

Step 4. Consider the partial derivative of F(X) with respect to any one of the nonbasic variable say z_k,

$$\frac{1}{2}\frac{\partial F}{\partial z_k} = \gamma_{k0} + \sum_{q=1}^{n-m} \gamma_{kq} z_q$$

with all other nonbasic variables held at zero

(a) If

$$\frac{\partial F}{\partial z_k} \geq 0,$$

for all k, the current solution is optimal.

(b) If

$$\frac{\partial F}{\partial z_k} < 0,$$

for at least one k, increase z_k until either

i) some basic variable becomes zero, or

ii) $\partial F/\partial z_k$ vanishes and is about to become positive.

Step 5. To determine the change in the basis, calculate

$$\text{Min}\left[\frac{\alpha_{h0}}{|\alpha_{hk}|}, h=1,2..m, \frac{|r_{k0}|}{r_{kk}}\right]$$

$$\text{for } \alpha_{hk} < 0 \quad \text{and } \gamma_{kk} > 0.$$

(i) If the minimum occurs for some h = v, then x_v becomes nonbasic, and

(ii) if the minimum occurs for the second term, introduce a new nonbasic variable u_t, defined by

$$u_t = \frac{1}{2}\frac{\partial F}{\partial z_k} = \gamma_{k0} + \sum_{q=1}^{n-m} \gamma_{kq} z_q$$

u_t is unrestricted and is called a free variable. This will lead to one additional equation and one more basic variable than before.

Step 6. Go to step 2 and repeat the process until

$$\frac{\partial F}{\partial z_k} \geq 0, \text{ for all k and}$$

$$\frac{\partial F}{\partial u_t} = 0, \text{ for all free variables.}$$

Step 7. Obtain the local (global if F is convex) optimal solution and the value of Min F by setting nonbasic variables equal to zero in their expressions.

24.4.2. Example

Consider the problem

$$\begin{aligned} \text{Maximize} \quad & x_1 + 2x_2 - x_2^2. \\ \text{Subject to} \quad & x_1 + 2x_2 \leq 4 \\ & 3x_1 + 2x_2 \leq 6 \\ & x_1, x_2 \leq 0. \end{aligned}$$

By introducing slack variables, we have the problem

$$\begin{aligned} \text{Minimize} \quad & F(X) = -x_1 - 2x_2 + x_2^2 \\ \text{Subject to} \quad & x_1 + x_2 + x_3 = 4 \\ & 3x_1 + 2x_2 + x_4 = 6 \\ & x_1, x_2, x_3, x_4 \leq 0 \end{aligned}$$

Let us take x_3 and x_4 as basic variables and express them in terms of nonbasic variables

$$x_3 = 4 - x_1 - 2x_2$$
$$x_4 = 6 - 3x_1 - 2x_2$$

and $$F(X) = -x_1 - 2x_2 + x_2^2$$

Now, $$\frac{1}{2}\frac{\partial F}{\partial x_1} = -\frac{1}{2} \text{ and } \frac{1}{2}\frac{\partial F}{\partial x_2} = -1.$$

Hence it is profitable to increase x_2 and we calculate

Min $$\left\{\frac{\alpha_{30}}{|\alpha_{32}|}, \frac{\alpha_{40}}{|\alpha_{42}|}, \frac{|\gamma_{20}|}{\gamma_{22}}\right\} = \text{Min}\left\{\frac{4}{2}, \frac{6}{2}, \frac{1}{1}\right\}$$

$$= \text{Min } (2,3,1) = 1.$$

We therefore introduce a free nonbasic variable

$$u_1 = \frac{1}{2}\frac{\partial F}{\partial z_2} = -1 + x_2$$

We again express basic variables in terms of nonbasic variables

$$x_2 = 1 + u_1$$
$$x_3 = 4 - x_1 - 2(1 + u_1)$$
$$= 2 - x_1 - 2u_1$$
$$x_4 = 6 - 3x_1 - 2(1 + u_1)$$
$$= 4 - 3x_1 - 2u_1$$
$$F(X) = -x_1 - 2(1 + u_1) + (1 + u_1)^2$$
$$= -1 - x_1 + u^2_1.$$

Now, $$\frac{1}{2}\frac{\partial F}{\partial u_1} = 0, \quad \frac{1}{2}\frac{\partial F}{\partial x_1} = -\frac{1}{2}$$

x_1 is therefore increased and we next calculate

Min $$\left\{\frac{\alpha_{30}}{|\alpha_{31}|}, \frac{\alpha_{40}}{|\alpha_{41}|}, \frac{|\alpha_{20}|}{\gamma_{21}}, \frac{|\gamma_{10}|}{\gamma_{11}}\right\}$$

$$= \text{Min}\left\{\frac{2}{1}, \frac{4}{3}\right\}, (\alpha_{21} = 0\ \gamma_{11} = 0)$$

$$= \frac{4}{3}.$$

Thus x_4 leaves the basis and x_1 enters.

Then

$$x_1 = \frac{4}{3} - \frac{2}{3}u_1 - \frac{1}{3}x_4$$

$$x_2 = 1 + u_1$$

$$x_3 = \frac{2}{3} - \frac{4}{3}u_1 + \frac{1}{3}x_4.$$

$$F(X) = -\frac{7}{3} + \frac{2}{3}u_1 + \frac{1}{3}x_4 + u_1^2$$

$$\frac{1}{2}\frac{\partial F}{\partial u_1} = \frac{1}{3}, \ \frac{1}{2}\frac{\partial F}{\partial x_4} = \frac{1}{6}.$$

Hence it is profitable to decrease u_1

We therefore introduce another free variable u_2,

$$u_2 = 1/3 + u_1$$

and thus

$$u_1 = -\frac{1}{3} + u_2$$

$$x_1 = \frac{14}{9} - \frac{2}{3}u_2 - \frac{1}{3}x_4.$$

$$x_2 = \frac{2}{3} + u_2$$

$$x_3 = \frac{10}{9} - \frac{4}{3}u_2 + \frac{1}{3}x_4$$

$$F(X) = -\frac{22}{9} + \frac{1}{3}x_4 + u_2^2$$

Now, $\quad \frac{1}{2}\frac{\partial F}{\partial u_2} = 0 \ \text{ and } \ \frac{1}{2}\frac{\partial F}{\partial x_4} = \frac{1}{6} > 0$

and hence the minimum solution is achieved.

$$x_1 = \frac{14}{9}, x_2 = \frac{2}{3} \text{ and } \operatorname{Min} F = -\frac{22}{9}.$$

24.5. Lemke's Complementary Pivoting Algorthm

In 1968, Lemke [301a] proposed a complementary pivoting algorithm for solving linear complementarity problems. Since the Kuhn–Tucker conditions for quadratic programming problems can be written as a linear complementarity problem (see section 24.5.3), Lemke's algorithm can be used to solve quadratic programs.

We therefore, first briefly discuss the linear complementarity problem and then present the complementary pivoting algorithm suggested by Lemke.

24.5.1. The Linear Complementarity Problem

In a complementarity problem, we are to find a $Z \in R^n$ satisfying

$$f(Z)\ ^{3}\ 0,\ Z\ ^{3}\ 0,\ Z^T f(Z) = 0 \tag{24.54}$$

where f is a given vector-valued function from R^n to R^n.

If $$f(Z) = MZ + q \tag{24.55}$$

where M is a given $n \times n$ matrix and q is a given n-vector, the problem (24.54) is called a linear complementarity problem.

The problem has applications in many areas, such as bimatrix games and engineering optimization. In the recent years, a large number of papers, dealing with important results and generalizations of the complementarity problem, have appeared in the literature.

Thus, the linear complementarity problem is to find vectors W and z such that

$$W - MZ = q \tag{24.56}$$

$$w_j \geq 0,\ z_j \geq 0, \quad j = 1, 2, \ldots n \tag{24.57}$$

$$w_j z_j = 0, \quad j = 1, 2, \ldots n \tag{24.58}$$

where M is a given $n \times n$ matrix, q is a given n-vector.

The pair $(w_j\ z_j)$ is said to be a pair of complementary variables.

A solution (W^T, Z^T) to the above system is called a complementary basic feasible solution, if (W^T, Z^T) is a basic feasible solution to (24.56) and (24.57) and variable of the pair (w_j, z_j) is basic for $j = 1, 2, \ldots n$.

If $q \geq 0$, we immediately see that $W = q$, $Z = 0$ is a solution to the linear complementarity problem. If however, $q \not\geq 0$, we consider the related system

$$W - MZ - ez_0 = q \tag{24.59}$$

$$w_j \geq 0,\ z_j \geq 0, \quad j = 1, 2, \ldots n \tag{24.60}$$

$$w_j\ z_j = 0, \quad j = 1, 2, \ldots n \tag{24.61}$$

where z_0 is an artificial variable and e is an n-vector with all components equal to one.

It should be noted that any solution to the related system with $z_0 = 0$, provides a solution to the linear complementarity problem ((24.56) – (24.58)). Lemke's algorithm attempts to drive z_0 to zero, thus obtaining a solution to the linear complementarity problem.

A feasible solution (w^T, Z^T, z_0) to the system (24.59) – (24.61) is called an almost complementary basic feasible solution if

(1) (W^T, z^T, z_0) is a basic feasible solution to (24.59), (24.60)

(2) neither w_s nor z_s are basic for some $s \in \{1,2,\ldots n\}$, and

(3) z_0 is basic, and exactly one variable from each complementry pair (w_j, z_j), $j = 1,2,\ldots n$, $j \neq s$, is basic.

Given an almost complementary basic feasible solution (W^T, Z^T, z_0), where w_s and z_s are both nonbasic, an adjacent almost complementary basic feasible solution is obtained by introducing either w_s or z_s in the basis replacing a variable other than z_0 from the basis.

Lemke's algorithm moves among the adjacent almost complementary basic feasible solutions until either a complementary basic feasible solution is obtained or a direction indicating unboundedness of the region defined by (24.59) – (24.61) is found.

Summary of Lemke's Algorithm

We now summarize Lemke's complementary pivoting algorithm for solving the linear complementarity problem.

Step 1. Introduce the artificial variable z_0 and consider the system (24.59) – (24.61)

(a) If $q \geq 0$, stop; $(W, Z) = (q, 0)$ is a complementary basic feasible solution.

(b) If $q \geq 0$, express the system (24.59), (24.60) in a tableau format as in the simplex method. Let qs = minimum {qi, Isi≤n} and update the tableau by pivoting at rows and the zo column, i.e. replace ws in the basis by z_0, thus, the basic variables z_0 and w_j for $j = 1, 2, \ldots n$, $j \neq s$, become nonnegative and yields an almost complementary basic feasible solution to start with. Let $y_s = z_s$.

Step 2. In the updated tableau, let $\bar{a}_s$ be the column under the variable y_s and $\bar{q}$ be the right-hand side column of constants densting the values of the basic variables. If $\bar{a}_S \leq 0$, go to step 5. Otherwise, determine a row index r by the minimum ratio test

$$\frac{\bar{q}_r}{\bar{a}_{rs}} = \underset{1 \leq i \leq n}{\text{minimum}} \left\{ \frac{\bar{q}_i}{\bar{a}_{is}} : \bar{a}_{is} > 0 \right\}$$

If the basic variable at the r is z_0, go to step 3. Otherwise, go to step 4.

Step 3. Replace z_0 by y_s and update the current tableau by pivoting at the y_s column and the z_0 row. stop, because a complementary basic feasible solution is obtained.

Step 4. The basic variable at row r is either w_l or z_l for some $l \neq s$. The variable y_s enters the basis and the tableau is updated by pivoting at row r and the y_s column, where

$$Y_s = \begin{cases} z_l, \text{ if } w_l \text{ leaves the basis.} \\ w_l \text{ if } z_l \text{ leaves the basis.} \end{cases}$$

Go to step 2.

Step 5. Stop A ray $D = \left\{ \left(W^T, Z^T, z_0 \right)^T + \lambda d : \lambda \geq 0 \right\}$ is found such that every point in D satisfies (24.60) and (24.61), (W^T, Z^T, z_0) is the almost complementary basic feasible solution associated with the current tableau and d has 1 at the row corresponding to Y_s, $-\ \bar{a}_s$ at the rows of the current basic variables and zero everywhere else. Stopping the algorithm at this step is termed ray termination.

Convergence of the Algorithm

The following theorems show that the algorithm stops in a finite number of steps, either with a complementary basic feasible solution or with ray termination. Under certain conditions of the matrix M, the algorithm stop with a complementary basic feasible solution. For the proof of the theorems see [301a].

Theorem 24.3: Suppose that each almost complementary basic feasible solution to the system (24.59) – (24.61) is nondegenerate. Then none of the points generated by the complementary pivoting algorithm is repeated and the algorithm stops in a finite number of steps.

We now definite the following:

A square matrix C is said to be copositive if $x^T cx \geq 0$ for each $x \geq o$. Further, C is said to be copositive-plus if it is copositive and $x \geq 0$ and $x^T cx = 0$ imply that $(C+C^T)\, x = 0$.

Lemma 24.2: If a square matrix C has nonnegative entries, then C is copositive. Further, if C has nonnegative entries with positive diogonal elements, then C is copositive-plus.

Theorem 24.4: Suppose that each almost complementary basic feasible solution to the system (24.59) – (24.61) is nondegenerate, and suppose that M is copositive-plus. Then, Lemke's complementary pivoting algorithm terrminates in a finite number of steps. In particular, if the system (24.56), (24.57) is consistent, then the algorithm terminates with a complementary basic feasible solution to the system (24.56) – (24.58). On the other hand, if the system (24.56), (24.57) is inconsistent, then the algorithm stops with ray termination.

Corollary 24.1: If the matrix M has nonnegative entries, with positive diagonal elements, then Lemke's complementary pivoting algorithm terminates in a finite number of steps with a complementary basic feasible solution.

Proof: By lemma 24.2, M is copositive-plus. Moreover, under the assumption on M, the system $W-MZ = q$, $W, Z \geq 0$ has a solution since we can choose Z sufficiently large so that $W = MZ + q \geq 0$. The result then follows from the theorem.

24.5.2. Example

Consider the linear complementarity problem of finding W and Z satisfying

$$W - Mz = q,\ W^T x = 0,\ W \geq 0,\ Z \geq 0,$$

where

$$M = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 2 & -1 \\ -2 & 1 & 0 \end{bmatrix}, q = \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix}$$

Introducing the artificial variable z_0, w_c form the following tableau

Tableau 1

Basic variable	w_1	w_2	w_3	z_1	z_2	z_3	z_0	R.H.S.
w_1	1	0	0	–1	0	–2	–1	–1
w_2	0	1	0	–3	–2	1	–1	2
w_3	0	0	1	2	–1	0	–1	–3

Since min $\{q_i : 1 \leq i \leq 3\} = q_3$, we pivot at row 3 and the z_0 column and for the next tableau we have $y_s = z_3$.

Tableau 2

Basic variable	w_1	w_2	w_3	z_1	z_2	z_3	z_0	R.H.S.
w_1	1	0	–1	–3	1	–2	0	2
w_2	0	1	–1	–5	–1	1	0	5
z_0	0	0	–1	–2	1	0	1	3

Here, $y_s = z_3$ enters the basis. By the minimum ratio test w_2 leaves the basis and for the next iteration $y_s = z_2$. We pivot at the w_2 row and the z_3 column.

Tableau 3

Basic variable	w_1	w_2	w_3	z_1	z_2	z_3	z_0	R.H.S.
w_1	1	2	–3	–13	–1	0	0	12
z_2	0	1	–1	–5	–1	1	0	5
z_0	0	0	–1	–2	1	0	1	3

Here, z_2 enters the basis and by the minimum ratio test z_0 leaves the basis. Pivoting at the z_0 row and the z_2 column, we get the next tableau and obtain the complementary basic feasible solution.

Tableau 4

Basic variable	w_1	w_2	w_3	z_1	z_2	z_3	z_0	R.H.S.
w_1	1	2	–4	–15	0	0	1	15
z_3	0	1	–2	–7	0	1	1	8
z_2	0	0	–1	–2	1	0	1	3

Thus the solution we obtain is

$z_1 = 0,\ z_2 = 3,\ z_3 = 8,\ w_1 = 15,\ w_2 = 0,\ w_3 = 0$

24.5.3. Solving Quadratic Programs by Complementary Pivoting Algorithm

In this selection, we show that the Kuhn–Tucker conditions for a quadratic programming problem reduce to a linear complementary problem. Thus, the complentary pivoting algorithm can be used to solve a quadratic programming problem.

Consider the quadratic programming problem

$$\begin{aligned} \text{Minimize} \quad & f(x) = p^T x + \tfrac{1}{2} x^T c x \\ & Ax \le b, \\ & x \ge 0 \end{aligned} \tag{24.62}$$

where A is an mxn matrix, p is an n-vector, b is an m-vector and c is an nxn symmetric matrix.

Let Y denote the vector of slack variables and U be the vector of Lagrangian multiplier associated with $A\,x \le b$.

The Kuhn–Tucker conditions can then be written as

$$\begin{aligned} & Ax + Y = b \\ & -Cx - A^T U + V = p \\ & V^T x = 0,\ U^T Y = 0 \\ & X, Y, U, V \ge 0 \end{aligned} \tag{24.63}$$

If, we now take

$$M = \begin{bmatrix} O & -A \\ A^T & C \end{bmatrix}, q = \begin{bmatrix} b \\ p \end{bmatrix}, W = \begin{bmatrix} Y \\ V \end{bmatrix} \text{and } Z = \begin{bmatrix} U \\ X \end{bmatrix}, \tag{24.64}$$

The Kuhn–Tucker conditions can be expressed as the linear complementary problem

$$W - MZ = q,\ W^T Z = 0,\ W, Z, \ge 0 \tag{24.65}$$

Thus, the complementary pivoting algorithm discussed above can be used to find a solution of the Kuhn–Tucker conditions for the quadratic programming problem.

Convergence of the Procedure

In order to prove the convergence of Lemke's complementary pivoting algorithm to solve the Kuhn–Tucker conditions for the quadratic programming problem we need the result given by Lemma.

Lemma 24.3: Let A be an mxn matrix and c be an nxn symmetric matrix. If c is positive semidefinite, then the matrix $M = \begin{bmatrix} O & -A \\ A^T & C \end{bmatrix}$ is copositive-plus.

Proof: Let $Z^T = (X^T, Y^T) \ge 0$ then

$$Z^T MZ = (X^T, Y^T) \begin{bmatrix} O & -A \\ A^T & C \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} = Y_C^T Y \ge 0, \text{ since C is positive}$$

semidefinite.

Hence M is copositive.

Now, suppose that $Z \geq 0$ and $Z^T MZ = 0$, then since $Z^TMZ = Y^T{}_CY = 0$ and C is positive semidefinite, we have $CY = 0$ (by Theorem 6.1) and hence

$$\left(M + M^T\right)Z = \begin{bmatrix} 0 \\ 2CY \end{bmatrix} = 0 \text{, so that M is copositive-plus.}$$

Lemma 24.4: Let A be an m$\dot{\text{x}}$n matrix and C be an nxn symmetric matrix. If C has nonnegative entries, then the matrix $M = \begin{bmatrix} 0 & -A \\ A^T & C \end{bmatrix}$ is copositive. Further, if C has nonnegative elements with positive diagonal elements, then M is copositive-plus.

Proof: Let $Z^T = (x^T, Y^T) \geq 0$, then

$Z^T MZ = Y^TCY$

and since it follows from assumption that $Y^TCY \geq 0$, M is copositive. Further, if C has positive diagnal elements, then $Z \geq 0$, $Z^TMZ = Y^TCY = 0$ imply that $Y = 0$.

Hence $CY = 0$. Thus, $(M+M^T)Z = \begin{bmatrix} 0 \\ 2CY \end{bmatrix} = 0$ and therefore M is copositive-plus.

Theorem 24.5

Consider the quadratic programming problem (24.62). Suppose that the problem is feasible. Further, suppose that Lemke's complementary pivoting algorithm is used to find a solution to the Kuhn–Tucker conditions $W–MZ = q$, $W, Z \geq 0$, $W^TZ = 0$ where M, q, W and Z are given by (24.64). Under the assumption that each almost complemtary basic feasible solution to this system, is nondenerate,

(i) The algorithm stops in a finite number of steps, if C is positive semidefinite. Furthermore, ray termination implies that there is an unbounded solution to the quadratic programming problem (24.62).

(ii) The algorithm stops in a finite number of steps with a solution to the Kuhn–Tucker conditions if any one of the following three coditions is satisfied:

(a) C is positive semidefinite and $p = 0$

(b) C is positive definite

(c) C has nonnegative entries with positive diagonal elements.

Proof: (i) If C is positive semidefinite, then by Lemma 24.3 M is copositive-plus. Then by Theorem 24.4, the algorithm stops in a finite member of steps with either a solution to the Kuhn–Tucker conditions or a ray termination.

Now, suppose ray termination occurs. Then by Theorem 24.4, ray termination is possible only if the system

$$Ax + Y = b$$
$$-Cx - A^TU + V = p$$
$$X, Y, U, V \geq 0$$

has no solution.

By theorem 7.9, we then must have a solution (d, g) of the system

$$A^T q - Cd \geq 0 \tag{24.66}$$

$$Ad \leq 0 \tag{24.67}$$

$$q \geq 0 \tag{24.68}$$

$$d \geq 0 \tag{24.69}$$

$$b^T q = p^T d \leq 0 \tag{24.70}$$

Since $Ad \leq 0$ and $q \geq 0$, multiplying (24.66) by $d^T \geq 0$, we have

$$0 \leq d^T A^T q - d^T Cd \leq 0 - d^T Cd = - d^T Cd \tag{24.71}$$

and C is positive semidefinite, $d^T Cd = 0$ and hence $Cd = 0$, (24.72)

Let $\overline{X}, \overline{Y}$ be a feasible solution to the quadratic programming problem (24.62) so that

$$A\overline{X} + \overline{Y} = b, \overline{X},\ \overline{Y} \geq 0 \tag{24.73}$$

Substituting b from (24.73) in (24.70), we get

$$0 > p^T d + b^T q = p^T d + (A\overline{X} + \overline{Y})^T q$$

$$\geq p^T d + \overline{X}^T A^T q \qquad \because\ \overline{Y} \geq 0, q \geq 0$$

$$\geq p^T d + \overline{X}^T Cd \qquad \text{by (24.66)}$$

Further, from (24.72), we note that cd = 0 and hence $p^T d < 0$.

We note that $Ad \leq 0$, $d \geq 0$, so that $\overline{x} + \lambda d$ is feasible to the problem for all $\lambda \geq 0$ and is a ray emanating from the point $\overline{x}$ in the direction d.

Now, $\qquad f(\overline{x} + \lambda d) = f(\overline{x}) + \lambda (p^T + \overline{x}^T C)d + \frac{1}{2} \lambda^2 d^T Cd$

$\qquad\qquad = f(\overline{x}) + \lambda\, p^T d$, since cd = 0

Since $\qquad p^T d < 0$, $f(\overline{x} + \lambda d) \to -\infty$ as $\lambda \to -\infty$.

and thus there is an unbounded solution to the problem (24.62).

(ii) By Lemma (24.3) and (24.4), M is copositive-plus under conditions (a), (b) or (c) of the theorem. It would therefore, be sufficient to show that row termination is not possible under any of these conditions. Suppose, by contradiction, that ray termination occurs under condition (a), (b) or (c). From (24.71), we note that $d^T Cd \leq 0$. Hence, under condition (b) or (c), d = 0, which contradicts (24.74). Under condition (a), Cd = 0 and moreover, by assumption p = 0. This again contradicts (24.74)

Thus, under conditions (a), (b) or (c), ray termination is not possible and the algorithm finds a solution to the Kuhn–Tucker conditions in a finite number of steps.

24.6. Exercises

1. Solve the following quadratic programs by Wolfe's method:

(i) Minimize $z = x_1 - 2x_3 + 1/2(x^2_1 + x^2_2 + x^2_3)$.

Subject to $x_1 - x_2 + x_3 = 1$

$x_1, x_2, x_3 \geq 0.$

(ii) Minimize $z = x_1 - x_2 + 1/2(x_1 - x_2)^2$

Subject to $2x_1 - x_2 \geq 4$

$2x_1 + x_2 \geq 7.$

$x_1, x_2 \geq 0.$

2. Solve the following problems by the method of Wolfe and by Beale's method.

(i) Maximize $z = 5x_1 + 2x_2 - 2x^2_1 + 2x_1x_2 - x^2_2$

Subject to $-5x_1 + 3x_2 \leq 4$

$3x_1 + 2x_2 \leq 20$

$x_1, x_2 \geq 0.$

(ii) Maximize $z = 2x_1 + x_2 - 1/2\, x^2_1 - 1/2\, x^2_2$

Subject to $4x_1 + x_2 \leq 5$

$3x_1 + 2x_2 \leq 6$

$x_1, x_2 \geq 0.$

3. Solve the following quadratic programming problem for all $\lambda \geq 0$

Maximize $z = \lambda(3x_1 + 2x_2) - 2x^2_1$

Subject to $x_1 + x_2 \leq 2$

$4x_1 + x_2 \leq 4$

$x_1, x_2 \geq 0.$

4. Solve the following quadratic programming problems by Beale's method.

(i) Maximize $z = 25x_1 + 10x_2 - x^2_1 - 4x_1x_2 - 10x^2_2$

Subject to $2x_1 + x_2 \leq 10$

$x_1 + x_2 \leq 9$

$x_1, x_2 \geq 0$

(ii) Maximize $z = 3x_1 + 6x_2 - 4x^2_1 + 4x_1x_2 - x^2_2$

$x_1 + 4x_2 \geq 9$

$x_1 + x_2 \leq 3$

$x_1, x_2 \geq 0.$

5. Show that the convergence of Wolfe's method for solving the quadratic program

Minimize $p^TX + 1/2\, X^TCX$

$AX = b$

$X \geq 0$

requires that either p = 0 or C is positive definite.

6. Discuss the case in Wolfe's method for solving the problem (24.1) when the value of the objective function is unbounded from below.

7. Are the matrices given below copositive-plus?

$$\text{(i)} \quad \begin{bmatrix} 1 & 0 & 2 \\ 3 & 2 & -1 \\ -2 & 1 & 0 \end{bmatrix}, \qquad \text{(ii)} \begin{bmatrix} 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & -2 \\ 1 & -1 & 2 & -2 \\ 1 & 2 & -2 & 4 \end{bmatrix}$$

8. Use the complementary pivoting algorithm to solve the Kuhn-Tucker conditions for the problem.

$$\begin{aligned} \text{Maximize} \quad & 3x_1 + x_2 \\ \text{Subject to} \quad & x_1 - x_2 \le 4 \\ & x_1 + x_2 \le 6 \\ & x_1, x_2 \ge 0 \end{aligned}$$

9. Use the complementary pivoting algorithm to solve the following quadratic programming problem.

$$\begin{aligned} \text{Maximize} \quad & -6x_1 - 2x_2 + 2x_1^2 - 2x_1 + x_2 + x_2^2 \\ \text{Subject to} \quad & 2x_1 - x_2 \le 2 \\ & x_1 + x_2 \le 2 \\ & x_1, x_2 \ge 0 \end{aligned}$$

CHAPTER 25

Methods of Nonlinear Programming

In this chapter we are concerned with the methods of solution for general nonlinear programs.

$$\begin{aligned} &\text{Maximize} && f(X) \\ &\text{Subject to} && g_i(X) \leq 0,\ i = 1, 2...m \\ &&& X \leq 0\ ,\ X \in R^n \end{aligned}$$

Several methods have been developed where either f and all g_i are nonlinear or only f is nonlinear. The methods that we will discuss in this chapter are the following:

(1) Separable Programming
(2) Kelley's Cutting Plane Algorithm
(3) Zoutendijk's Method of Feasible Direction
(4) Rosen's Gradient Projection Method
(5) Wolfe's Reduced Gradient Method
(6) Zangwill's Convex–Simplex Method
(7) Dantzig's Method for Convex Programming

25.1. Separable Programming

Separable programming is a special technique for obtaining solutions of a class of nonlinear programming problems where the functions involved can be expressed as a sum of functions each of a single variable only, that is, the functions f(X) and g_i (X) are of the form.

$$f(X) = \sum_{j=1}^{n} f(x_j) \quad , \quad g_i(x) = \sum_{j=1}^{n} g_{ij}(x_j), \qquad i= 1,2..m$$

The functions are then said to be separable. The linear function

$$f(X) = \sum_{j=1}^{n} c_j x_j$$

and the nonlinear function

$$f(x_1, x_2) = c_1x^2_1 + c_2x^3_2 - c_3x_1 + c_4x_2$$
$$= (c_1x^2_1 - c_3x_1) + (c_2 x^3_2 + c_4x_2)$$
$$= f_1(x_1) + f_2(x_2)$$

are examples of separable function.

There are some functions, which are not directly separable can also be made so by transformation of variables. For example, in a given problem, the product term x_1x_2 can be transformed into a separable function by setting

$$u_1 = 1/2\,(x_1 + x_2)\,;\; u_2 = 1/2\,(x_1 - x_2) \qquad (25.2)$$

so that $\quad x_1x_2 = u^2_1 - u^2_2$

The term $x_1\, x_2$ can then be replaced by the separable function $u^2_1 - u^2_2$, while the separable expressions in (25.2) are included in the set of constraints of the problem.

In general the term $x_1^{a_1}.x_2^{a_2}..x_n^{a_n}$, where the x_j are essentially positive and have a strictly positive lower bound can be replaced by e^w, while the expression

$$w = a_1 \ln x_1 + a_2 \ln x_2 + ... + a_n \ln x_n.$$

is included in the set of constraints.

Consider now, the separable nonlinear program

$$\text{Minimize} \quad f(X) = \sum_{j=1}^{n} f_j(x_j)$$

$$\text{Subject to} \quad \sum_{j=1}^{n} g_{ij}(x_j) \le b_i, \qquad i=1,2...m \qquad (25.3)$$

$$x_j \ge 0\;, \qquad j=1,2...n.$$

In this section, we shall discuss a special method which makes use of a modification of the simplex method to provide good approximate solutions to nonlinear separable programs.

25.1.1. Approximating the Problem

In order to be able to make use of the simplex method for finding a solution of the problem, the separable functions $f_j(x_j)$ and $g_{ij}(x_j)$, i = 1,2..m are first approximated by piecewise linear functions.

To see how this can be done, consider a continuous function h (x) of a single variable x defined over the interval [a, b]. The interval [ab] is then subdivided into smaller intervals by the grid points a_r (r = 1, 2..k) such that

$$a = a_1 < a_2 < ... a_k = b.$$

Now, any point x in the interval $[a_r, a_{r+1}]$ can be uniquely expressed as

$$x = \lambda_r a_r + \lambda_{r+1} a_{r+1} \qquad (25.4)$$

where $\quad \lambda_r + \lambda_{r+1} = 1\;,\; \lambda_r, \lambda_{r+1} \ge 0.$

Then $$\hat{h}(x) = \lambda_r h(a_r) + \lambda_{r+1} h(a_{r+1})$$

gives a linear approximation of the function h in the interval $[a_r, a_{r+1}]$ as shown in Figure 25.1.

It should be noted that the grid points need not be equidistant and that the accuracy of the approximation can be improved by using finer grids. However, with the increase in the number of grid points, the number of variables in the approximating linear program increases and this may result in more iterations in solving the problem. The selection of grid points should therefore be suitably made. It is also noted that the approximation using adjacent grid points yields better approximation.

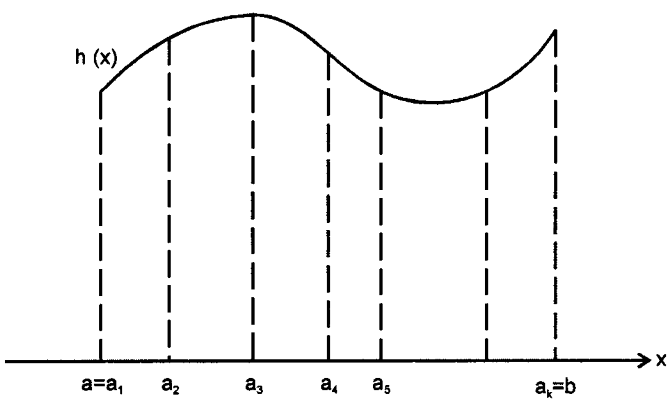


Figure 25.1. Piecewise linear approximation of a function.

The piecewise linear approximation of the function h over the interval [a,b] can then be written as

$$\hat{h}(x)\sum_{r=1}^{k}\lambda_r h(a_r) \tag{25.5}$$

where $$x = \sum_{r=1}^{k}\lambda_r a_r$$

$$\sum_{r=1}^{k}\lambda_r = 1 \ , \quad \lambda_r \geq 0, \qquad r = 1,2...k.$$

and that at most two adjacent λ_r's are positive

Let us now return to the separable programming problem (25.3) and assume that all the f_j and g_{ij} are continuous. Now, it may so happen that for some j, f_j and all g_{ij}, i = 1, 2..m are linear and therefore for convenience we define a set L as

$$L = \{ j : f_j \text{ and } g_{ij}, i = 1, 2..m \text{ are linear}\}. \tag{25.6}$$

Suppose that for each $j \notin L$, the range of the variable x_j is $[a_j, b_j]$, where a_j, $b_j \geq 0$. We then divide the interval by grid points x_{rj}, $r = 1, 2...k_j$ such that

$$a_j = x_{1j} < x_{2j} < ... < x_{kj}\, j = b_j \tag{25.7}$$

where the grid points (not necessarily equidistant) are selected judiciously so that the approximations of the functions can be obtained with sufficient accuracy.

Using the method described above, piecewise linear approximations of $f_j(x_j)$, $g_{ij}(x_j)$, $(i = 1,2...m)$ for each $j \notin L$ are then obtained as

$$\hat{f}(x_j) = \sum_{r=1}^{kj} \lambda_{rj} f(x_{rj}) \ , \text{ for } j \notin L$$

$$\hat{g}_{ij}(x_j) = \sum_{r=1}^{kj} \lambda_{rj} g_{ij}(x_j) \ , \text{ for } i = 1,2...m; j \notin L \tag{25.8}$$

where
$$x_j = \sum_{r=1}^{kj} \lambda_{rj} r_{rj}, j \notin L \tag{25.9}$$

$$\sum_{r=1}^{kj} \lambda_{rj} = 1 \ , \quad \lambda_{rj} \geq 0, \ \text{ for } r = 1,2,...k_j, j \notin L \tag{25.10}$$

and at most two adjacent λ_{rj}'s are positive for $j \notin L$.

Since for $j \in L$, both f_j and g_{ij} for $i = 1, 2...m$ are linear, no grid points need to be introduced in this case and the functions are kept in their present form.

Now, replacing the nonlinear functions $f_j(x_j)$, $g_{ij}(x_j)$, $j \notin L$ in (25.3), by their linear approximations, the separable program is approximated by the following problem

$$\text{Minimize} \quad z = \sum_{j \in L} f_j(x_j) + \sum_{j \notin L} \sum_{r=1}^{kj} \lambda_{rj} f(x_{rj})$$

$$\text{Subject to} \quad \sum_{j \in L} g_{ij}(x_j) + \sum_{j \notin L} \sum_{r=1}^{kj} \lambda_{rj} g_{ij}(x_{rj}) \leq b_i \ , \ i=1,2...m$$

$$\sum_{r=1}^{kj} \lambda_{rj} = 1, \ j \notin L$$

$$\lambda_{rj} \geq 0 \ , r = 1,2...k_j; j \notin L \tag{25.11}$$

$$x_j \geq 0 \ , \ j \in L$$

and the additional restriction that atmost two adjacent λ_{rj}'s are positive for $j \notin L$.

25.1.2. Solution of the Problem

Except for the additional restriction, the approximating problem (25.11) is a linear program. It can therefore, be solved by the simplex method with a little modification. The simplex method is now applied with the 'restricted basis entry rule', which specifies that a nonbasic variable λ_{rj} is introduced into the basis only if it improves the value of the objective function and if for each $j \notin L$, the new basis has no more than two adjacent λ_{rj} that are positive. From the local optimal solution to the approximating problem thus obtained, an approximate local optimal solution $\hat{x}$ to the original problem is obtained by (25.9). Thus if (x^0_j, λ^0_{rj}) is a solution of (25.11), then the jth component of $\hat{x}$ is given by

$$\hat{x}_j = \begin{cases} x_j^0 & , \quad \text{for } j \in L \\ \sum_{r=1}^{kj} \lambda_{rj}^0 x_{rj} & , \quad \text{for } j \notin L \end{cases} \tag{25.12}$$

It should be noted that the solution obtained to the approximating problem may not even be a feasible solution to the original problem. If however, the original constraint set is convex, a feasible solution to the approximating problem will also be a feasible solution to the original problem.

The following theorem shows that if for each $j \notin L$, f_j is strictly convex and g_{ij} is convex for $i = 1, 2,...m$, then we can discard the additional restriction in (25.11) and solve the problem by the usual simplex method.

Theorem 25.1. Consider the separable program (25.3) and let the set L be defined as in (25.6). Suppose that for each $j \notin L$, the function $f_i(x_j)$ is strictly convex and that $g_{ij}(x_j)$ is convex for $i = 1, 2,...m$. Further, suppose that x^0_j ($j \notin L$) and λ^0_{rj} ($r = 1, 2,...k_j$; $j \notin L$) solve the approximating linear program (25.11) without the additional restriction. Then

(i) the vector $\hat{x}$, whose components are given by

$$\hat{x}_j = x_j^0, \qquad \text{for } j \in L$$

$$\hat{x}_j = \sum_{r=1}^{k_j} \lambda_{rj}^0 x_{rj} \quad \text{for } j \notin L$$

is feasible to the original problem

(ii) For each $j \notin L$, atmost two λ^0_{rj}'s are positive and they must be adjacent.

Proof:

(i) Since the functions $g_{ij}(x_j)$ are convex for $j \notin L$ and $i = 1,2,..m$ we have

$$\sum_{j=1}^{n} g_{ij}(\hat{x}_j) = \sum_{j \in L} g_{ij}(\hat{x}_j) + \sum_{j \notin L} g_{ij}(\hat{x}_j)$$

$$= \sum_{j \in L} g_{ij}(x_j^0) + \sum_{j \notin L} g_{ij}\left(\sum_{r=1}^{kj} \lambda_{rj}^0 x_{rj}\right)$$

$$\leq \sum_{j\in L} g_{ij}(x_j^0) + \sum_{j\notin L} \lambda_{rj}^0 g_{ij}(x_{rj})$$

$$\leq b_i, \qquad \text{for} \quad i=1,2,...m.$$

Moreover,

$$\text{for } j\in L, \quad \hat{x}_j = x_j^0 \geq 0$$

$$\text{for } j\notin L, \quad \hat{x}_j = \sum_{r=1}^{kj} \lambda_{rj}^0 x_{rj} \geq 0 \ , \quad \text{since } \lambda_{rj}^0 \geq 0, x_{rj} \geq 0$$

for $r = 1,2...k_j$ and $j\notin L$.

To prove part (ii), it is sufficient to show that for each $j\notin L$, if λ^0_{lj} and λ^0_{qj} are positive, then the grid points x_{rj} and x_{qj} must be adjacent. Let us suppose, on the contrary that λ^0_{lj}, λ^0_{qj} are positive but x_{lj}, x_{qj} are not adjacent. Then there exists a grid point $x_{pj} \in (x_{lj}, x_{qj})$ that can be expressed as

$$x_{pj} = \theta\, x_{lj} + (1-\theta)\, x_{qj}, \text{ for some } \theta,\ 0<\theta<1.$$

Since x^0_j, $(j \in L)$, λ^0_{rj} $(r = 1, 2,,, k_j, j \notin L)$ solve the linear program obtained from (25.11) ignoring the additional restriction, it follows that there exist Lagrange multipliers $u^0_i \geq 0$ for $i = 1, 2...m$ and v^0_j for $j\notin L$ associated with the constraints of the problem such that the following subset of the Kuhn–Tucker necessary conditions are satisfied for each $j\notin L$:

$$f_j(x_{rj}) + \sum_{i=1}^{m} u_i^0 g_{ij}(x_{rj}) + v_j^0 \geq 0$$

$$\left[f_j(x_{rj}) + \Sigma u_i^0 g_{ij}(x_{rj}) + v_j^0\right]\lambda_{rj}^0 = 0, \qquad r = 1,2,...k_j.$$

Since λ^0_{lj} and λ^0_{qj} are positive, we have, from the above relations

$$f_j(x_{rj}) + \sum_{i=1}^{m} u_i^0 g_{ij}(x_{lj}) + v_j^0 = 0 \tag{25.13}$$

$$f_j(x_{qj}) + \sum_{i=1}^{m} u_i^0 g_{ij}(x_{qj}) + v_j^0 = 0 \tag{25.14}$$

$$f_j(x_{pj}) + \sum_{i=1}^{m} u_i^0 g_{ij}(x_{pj}) + v_j^0 \geq 0 \tag{25.15}$$

Now, by strict convexity of f_j and convexity of g_{ij} for $j \notin L$, we get

$$f_j(x_{pj}) + \sum_{i=1}^{m} u_i^0 g_{ij}(x_{pj}) + v_j^0$$

$$= f_j(\theta x_{lj} + (1-\theta)x_{qj}) + \sum_{i=1}^{m} u_i^0 g_{ij}(\theta x_{lj} + (1-\theta)x_{qj}) + v_j^0$$

$$< \theta f_j(x_{lj}) + (1-\theta)f_j(x_{qi}) + \sum_{i=1}^{m} u_i[\theta g_{ij}(x_{lj}) + (1-\theta)g_{ij}(x_{qj})]$$

$$+\theta v_j^0 + (1-\theta)v_j^0 = 0, \qquad (25.16)$$

by (25.13) and (25.14)

This contradicts (25.15) and hence x_{lj} and x_{qj} must be adjacent.

25.1.3. Grid Refinement

It is natural that the accuracy of the above procedure largely depends on the number of grid points for each variable. However, with the increase in the number of grid points, the number of variables in the approximating linear program increases and this may result in more iterations in solving the problem.

One way to select a suitable grid is as follows. [344]

Step 1. Solve the problem initially using a course grid.

Step 2. Refine the grid around the optimal solution obtained with the coarse grid.

Step 3. Resolve the problem using the refined grid.

Step 4. If the improvement in the optimal objective value of the problem using the new grid is significant go to step 2. Otherwise, select the grid used in step 3.

25.1.4. Example

Consider the following separable problem

$$\text{Minimize } f(X) = 2x_1^2 - 4x_1 + x_2^2 - 3x_3$$

$$\text{Subject to} \quad x_1 + x_2 + x_3 \le 4$$

$$x_1^2 - x_2 \le 2$$

$$x_j \ge 0, j = 1, 2, 3.$$

Since the variable x_3 appears linearly both in the objective function and the constraints, the set $L = \{3\}$. We therefore need not take any grid points for x_3. From the constraints it is clear that both x_1 and x_2 lie in the interval [0, 4]. Though it is not necessary for the grid points to be equally spaced, here we use the grid points 0, 2, 4 for both the variables x_1 and x_2, so that

$$x_{11} = 0, x_{21} = 2, x_{31} = 4, \text{ and}$$

$$x_{12} = 0, x_{22} = 2, x_{32} = 4.$$

Then,

$$x_1 = 0\,\lambda_{11} + 2\,\lambda_{21} + 4\,\lambda_{31}.$$

$$x_2 = 0\,\lambda_{12} + 2\,\lambda_{22} + 4\,\lambda_{32}$$

where

$$\lambda_{11} + \lambda_{21} + \lambda_{31} = 1$$

$$\lambda_{12} + \lambda_{22} + \lambda_{32} = 1$$

$$\lambda_{rj} \ge 0, \text{ for } r = 1, 2, 3; j = 1, 2$$

$$\hat{f}(X)=(16\lambda_{31}) + (4\lambda_{22}+16\lambda_{32})-3x_3$$

$$\hat{g}_1(X) = (2\lambda_{21} + 4\lambda_{31}) + (2\lambda_{22} + 4\lambda_{32}) + x_3 \le 4$$

$$\hat{g}_2(X) = (4\lambda_{21} + 16\lambda_{31}) + (-2\lambda_{22} - 4\lambda_{32}) \le 2$$

The approximating program therefore is

$$\begin{aligned}
&\text{Minimize } z = 16\,\lambda_{31} + 4\,\lambda_{22} + 16\,\lambda_{32} - 3x_3 \\
&\text{Subject to} \quad 2\,\lambda_{21} + 4\,\lambda_{31} + 2\,\lambda_{22} + 4\,\lambda_{32} + x_3 \le 4 \\
&\qquad\qquad\quad\ 4\,\lambda_{21} + 16\,\lambda_{31} - 2\,\lambda_{22} - 4\,\lambda_{32} \le 2 \\
&\qquad\qquad\quad\ \lambda_{11} + \lambda_{21} + \lambda_{31} = 1. \\
&\qquad\qquad\quad\ \lambda_{12} + \lambda_{22} + \lambda_{32} \le 0. \\
&\qquad\qquad\quad\ \lambda_{rj} \le 0,\ r = 1, 2, 3; \quad j = 1, 2.
\end{aligned}$$

Introducing the slack variables $x_4 \ge 0$ and $x_5 \ge 0$ in the first two constraints, we solve the problem by the simplex method with the restricted basis entry rule, that is, for each j=1, 2, at the most two of the variables λ_{1j}, λ_{2j}, λ_{3j} are positive and they must be adjacent. The first tableau is given below

Tableau 1

Basic Variables	λ_{11}	λ_{21}	λ_{31}	λ_{12}	λ_{22}	λ_{32}	x_3	x_4	x_5	Constant
x_4	0	2	4	0	2	4	1	1	0	4
x_5	0	4	16	0	–2	–4	0	0	1	2
λ_{11}	1	1	1	0	0	0	0	0	0	1
λ_{12}	0	0	0	1	1	1	0	0	0	1
z	0	0	16	0	4	16	–3	0	0	

It is left to the reader to obtain the subsequent simplex tableaus finally giving an optimal solution to the approximating problem. From this solution an approximate optimal solution to the original problem is obtained by (25.12)

Since the objective function and the constraint functions of this problem satisfy the assumptions of Theorem 25.1, we could have applied the simplex method without the restricted basis entry rule and yet obtained the same optimal solution.

25.1.5. Mixed Integer Programming Formulation

In the method discussed above, we have seen that a nonlinear separable program can be approximated by a linear program with an additional restriction, which can be solved by the simplex method with a minor modification. The procedure however, finds a local optimal solution to the problem (25.11) if the functions are not convex. A mixed integer programming[1] formulation of the problem where the additional restriction is replaced by a set of constraints involving integer valued variables has been discussed by Markowitz and Manne [329] and Dantzig [105]. Such a formulation leads to a global optimal solution to the problem. The simplest such formulation given by Dantzig [105] is as follows:

$$\text{Minimize} \quad \sum_{j\in L} f_j(x_j) + \sum_{j\notin L}\sum_{r=1}^{k_j} \lambda_{rj} f(x_{rj})$$

$$\text{Subject to} \quad \sum_{j\in L} g_{ij}(x_j) + \sum_{j\notin L}\sum_{r=1}^{k_j} \lambda_{rj} g_{ij}(x_{rj}) \le b_i, \quad i = 1,2,..m$$

$$\begin{aligned} &0 \le \lambda_{lj} \le \delta_{lj} \\ &0 \le \lambda_{rj} \le \delta_{r-j} + \delta_{rj}, \quad r = 2, 3,\dots k_j - 1, \\ &0 \le \delta_{kj}, j \le \delta_{kj-1}, {}_j \\ &\sum_{r=1}^{kj-1} \delta_{rj} = 1 \\ &\sum_{r=1}^{kj} \lambda_{rj} = 1 \\ &\delta_{rj} = 0 \text{ or } 1, \quad j \notin L \end{aligned} \tag{25.17}$$

25.2. Kelley's Cutting Plane Method

Consider the nonlinear programming problem

$$\begin{aligned} &\text{Minimize} && f(X) \\ &\text{Subject to} && g_i(X) \le 0, \; i = 1, 2\dots p \end{aligned} \tag{25.18}$$

where f(X) and $g_i(X)$ are continuously differentiable convex functions of $X \in R^n$.

It is easy to see that the problem (25.18) can be expressed in the equivalent form

$$\begin{aligned} &\text{Minimize} && x_{n+1} \\ &\text{Subject to} && f(X) - x_{n+1} \le 0 \\ &&& g_i(X) \le 0, \; i = 1, 2\dots p. \end{aligned} \tag{25.19}$$

Thus by the addition of one constraint and one variable, we have converted the original problem (25.18) into a problem with a linear objective function.

The problem therefore can be expressed in the form

$$\begin{aligned} &\text{Minimize} && C^T X \\ &\text{Subject to} && g_i(X) \le 0, \; i = 1, 2\dots m. \end{aligned} \tag{25.20}$$

where $g_i(X)$ are continuously differentiable convex functions of $X \in R^n$.

It is therefore sufficient to restrict our attention to the convex program (25.20).

1. A mixed integer programming problem is a linear program where some of the variables are restricted to have integer values only.

This is the form Kelley [273] has considered to develop his cutting plane algorithm.

Let G denote the feasible set of the problem (25.20). We assume that G is contained in a compact convex polyhedron S_0, defined by a finite number of closed half spaces and then solve the linear program

$$\begin{aligned} &\text{Minimize} \quad C^T X \\ &\text{Subject to} \quad X \in S_0. \end{aligned} \tag{25.21}$$

Let X_0 denote the optimal solution of (25.21), which is of course an extreme point of S_0. If $X_0 \in G$, then it must be an optimal solution of (25.20). If not, then it produces a linear constraint that cuts off a portion of S_0 to yield a new compact polyhedron S_1 ($S_0 \supset S_1 \supset G$). Thus, a sequence $\{S_k\}$ of nested convex polyhedron is produced such that

$$G \subset S_k,\ S_{k+1} \subset S_k,\ k = 0, 1, 2, \ldots \tag{25.22}$$

and a sequence $\{X_k\}$ of infeasible points is generated which under certain conditions converge to the optimal solution of the original problem. (see Section 25.2.2)

25.2.1. Summary of the Algorithm

We now describe the steps of the algorithm.

Step 1. Find a compact convex polyhedron S_0 containing G defined by a finite number of linear inequalities i.e.

$$S_0 = \{X \mid AX \le b\,,\, X \in R^n\}. \tag{25.23}$$

There are a number of ways to find such an S_0. One possibility is to let S_0 be the set of all points X satisfying the linear constraints $g_i(X) \le 0$ of the problem (25.20), provided the set is compact. Another reasonable procedure for choosing S_0 for many practical problems is to determine the lower and upper bounds L_j and M_j on the value of each variable X_j and define S_0 to be the rectangular hypersolid

$$S_0 = \{X \mid L_j \le x_j \le M_j, j = 1, 2, \ldots n\} \tag{25.24}$$

Now solve the problem

$$\begin{aligned} &\text{Minimize} \quad C^T X \\ &\text{Subject to} \quad X \in S_0 \end{aligned} \tag{25.25}$$

Let X_0 denote an optimal solution to the linear program (25.25) If $X_0 \in G$, X_0 is an optimal solution to the problem (25.20) and the method terminates. Otherwise, go to step 2.

Step 2. If X_0 does not satisfy the constraints of (25.20) then there is at least one constraint in G for which $g_i(X_0) > 0$. Let the most violated constraint be $g_r(X)$, i.e.

$$g_r(X_0) = \text{Max}\,[g_1(X_0), g_2(X_0) \ldots g_m(X_0)] > 0 \tag{25.26}$$

Consider the linear approximating constraint to $g_r(x) \le 0$, given by

$$h_0(X) \equiv g_r(X_0) + \nabla g_r(X_0)^T (X - X_0) \le 0 \tag{25.27}$$

Since all g_i are convex functions on G, we have

$$h_0(X) \equiv g_r(X_0) + \nabla g_r(X_0)^T (X - X_0) \leq g_r(X), \text{ for all } X \in G. \tag{25.28}$$

Obviously, $\nabla g_r(X_0) \neq 0$, otherwise, since $g_r(X_0) > 0$, (25.28) will imply that the problem (25.20) is infeasible.

Hence for all $X \in G$,

$$h_0(X) \leq g_r(X) \leq 0,$$

so that the linear constraint (25.27) is satisfied by all points of G but is not satisfied at the point X_0 since $g_r(X_0) > 0$. Hence the constraint (25.27) cuts off a portion of the convex polyhedron S_0, but not any portion of G. For this reason the hyperplane $h_0(X) = 0$ is called the cutting hyperplane and any algorithm based on the addition of (25.27) to the constraints of (25.25) is called the cutting plane algorithm.

We now solve the problem

$$\begin{aligned} &\text{Minimize} && C^T X \\ &\text{Subject to} && X \in S_1 \end{aligned} \tag{25.29}$$

where $S_1 = \{X \mid X \in S_0, \text{ and } h_0(X) \leq 0\}$

Note that $G \subset S_1 \subset S_0$.

Let X_1 be an optimal solution of (25.29) ($X_1 \neq X_0$). If $X_1 \in G$, then X_1 is an optimal solution of (25.20) and the procedure terminates.

If $X_1 \notin G$, go to step 3.

Step 3. If $X_1 \notin G$, proceeding as before we form a convex polyhedron S_2 by adding to the constraint set of (25.29), the cutting hyperplane

$$h_1(X) \equiv g_{r_1}(X_1) + \nabla g_{r_1}(X_1)^T (X - X_1) \leq 0 \tag{25.30}$$

where $g_{r_1}(X_1) = \text{Max}\,[g_1(X_1), g_2(X_1), \dots g_m(X_1)] > 0$

and solve the problem

$$\begin{aligned} &\text{Maximize} && C^T X \\ &\text{Subject to} && X \in S_2 \end{aligned}$$

where $S_2 = \{X \mid X \in S_1 \text{ and } h_1(X) \leq 0\}$

Continuing in this manner, we generate a sequence of nested convex, polyhedron $\{S_k\}$ so that

$$G \subset S_k, S_{k+1} \subset S_k \text{ for } k = 0, 1, 2\dots$$

and a corresponding sequence of infeasible points $\{X_k\}$ that converges to the optimal solution X* of the original problem (25.20).

25.2.2. Convergence of the Algorithm

Under some mild assumptions on the convex functions, convergence of Kelley's algorithm can be established as follows.

Theorem 25.2. Let the convex functions g_i, i = 1, 2...m be continuously differentiable on a compact convex set S_0 and suppose that there is a positive number M such that $| g_i \leq (X)|$ M, for all $X \in S_0$, i = 1, 2,...m. Further assume that the feasible set G of the problem (25.20) is nonempty and is contained in S_0. Then, any limit point of the sequence generated by Kelley's algorithm is an optimal solution of the problem (25.20)

Proof: Since S_0 is compact, the sequence $\{X_k\}$ in S_0 has convergent subsequence with a limit in S_0.

Let $\{X_k\}$, $k \in K$, where K is a subset of all positive integers, be a subsequence of $\{X_k\}$, converging to $\overline{X}$. Now, if $k, \in K$ $k' \in K$, and $k' > k$, then we must have

$$g_i(X_k) + \nabla g_i(X_k)(X_{k'} - X_k) \leq 0,$$

which implies that

$$g_i(X_k) \leq |\nabla g_i(X_k)|\,|X_{k'} - X_k| \tag{25.31}$$

Since $|\nabla g_i(X_k)|$ is bounded, the right hand side of (25.31) goes to zero as k and k' go to infinity. The left-hand side goes to $g_i(\overline{X})$. Thus $g_i(\overline{X}) \leq 0$ and hence $\overline{X}$ is feasible for the problem (25.20).

Now, if C^TX^* be the optimal value of the problem (25.20) then $C^TX_k \leq C^TX^*$, for each k since X_k is obtained by minimizing C^TX over a set containing G. Thus by continuity $C^T\overline{X} \leq C^TX^*$ and hence $\overline{X}$ is an optimal solution of (25.20)

However, the rate of convergence of the algorithm has not been satisfactorily analysed. In practice therefore, the algorithm is terminated when

$$g_i(X_k) \leq \in, \; i = 1, 2, .. m$$

where $\in > 0$ is a small preselected tolerance limit.

We should also note that in Kelley's algorithm, at each iteration one new constraint is added and this suggests that the dual simplex method can also be used at each stage. Moreover, computational efficiency of the algorithm can be increased by dropping some of the cutting plane constraints which are no longer meaningful. One way to do this is to discard all nonbinding cutting plane constraints at the end of each stage so that the linear programs do not grow too large. Convergence properties are not destroyed by this process since the sequence of objective values will still be monotically increasing.

One may add all constraints of the form of (25.27) at each iteration, but this increases the size of the problem and therefore, the best scheme would probably to add a single linear constraint as discussed above.

For other cutting plane methods see Zangwill [550] and Veinott [486].

25.2.3. Example

Consider the following problem

Minimize $\quad z = x_1 - x_2$

Subject to $\quad g(X) = 3x_1^2 - 2x_1x_2 + x_2^2 - 1 \leq 0.$

Since g(X) is convex, it is a convex programming problem and Kelley's algorithm may therefore be applied. From graphical representation, it can be seen that the optimal solution of the problem is $X^{*T} = (0,1)$ and Min $Z = -1$.

Let $S_0 = \{X \mid -2 \le x_1 \le 2, -2 \le x_2 \le 2\}$ which contains the feasible region G. of the given problem.

We solve the linear program

$$\begin{aligned} \text{Minimize} \quad & z = x_1 - x_2 \\ \text{Subject to} \quad & -2 \le x_1 \le 2 \\ & -2 \le x_2 \le 2 \end{aligned}$$

It can be seen that the optimal solution of the problem is $x_{01} = -2$, $x_{02} = 2$, with $z = -4$

Since $g(X_0) = 23 > 0$, X_0 is not feasible to the original problem. We therefore linearize g(X) at X_0 and get

$$h_0(X) = g(X_0) + \nabla g(X_0)^T(X - X_0) = -16x_1 + 8x_2 - 25 \le 0$$

We then solve the problem

$$\begin{aligned} \text{Minimize} \quad & z = x_1 - x_2 \\ \text{Subject to} \quad & X \in S_1 \end{aligned}$$

where $\quad S_1 = \{X \mid -2 \le x_1 \le 2, -2 \le x_2 \le x_2 \text{ and } -16x_1 + 8x_2 \le 25\}$

and obtain the optimal solution X^T_1 $(-.5625, 2.0)$ with $Z = -2.5625$.

Since $g(X_1) = 6.199 > 0$, X_1 is not feasible to the given problem. We therefore linearize g(X) about X_1 and the next cutting plane constraint is

$$\begin{aligned} h_1(X) &= g(X_1) + \nabla g(X_1)^T (X - X_1) \\ &= -7.375x_1 + 5.125x_2 - 8.199 \le 0 \end{aligned}$$

This gives the new linear program

$$\begin{aligned} \text{Minimize} \quad & z = x_1 - x_2 \\ \text{Subject to} \quad & X \in S_2 \end{aligned}$$

i.e.

$$\begin{aligned} \text{Minimize} \quad & z = x_1 - x_2 \\ \text{Subject to} \quad & -2 \le x_1 \le 2 \\ & -2 \le x_2 \le 2 \\ & -16x_1 + 8x_2 \le 25 \\ & -7.375x_1 + 5.125x_2 \le 8.199. \end{aligned}$$

We find that the optimal solution is

$$X_2 = (.279, 2.000) \text{ with } z = 1.722.$$

The point X_2 still does not satisfy the constraint $g(X) \le 0$ and the process is continued. The results thus obtained are given in the Table 25.1.

Table 25.1

k	X_k	z	$g(X_k)$
0	(–2.00, 2.00)	–4.000	23.000
1.	(–.563, 2.000)	–2.563	6.199
2.	(.279, 2.00)	–1.722	2.120
3.	(–.530, .838)	–1.367	1.431
4.	(–.053, 1.160)	–1.213	.478
5.	(.427, 1.485)	–1.058	.484
6.	(.171, 1.207)	–1.036	.132
7.	(.0183, 1.041)	–1.023	.047
8.	(–.166, .840)	–1.007	.068
9.	(–.073, .930)	–1.003	.017
.	.	.	.
.	.	.	.
.	.	.	.

It is evident from the Table 25.1 that the optimal solution of the problem cannot be obtained in a finite number of iterations. This, in fact, is frequently the case in applications of Kelley's algorithm. We therefore continue the process till $g(X) \leq \epsilon$, where $\epsilon > 0$ is a preassigned tolerance limit.

In this problem, we take $\epsilon = .02$, and obtain an approximate optimal solution of the problem as.

$$X^T = (-.073, .930) \text{ with } z = -1.003.$$

25.3. Zoutendijk's Method of Feasible Directions

Zoutendijk's method of feasible directions [563] is an iterative procedure for solving convex programming problems, which can be considered to be a large step gradient method. Starting with a feasible solution of the problem, the method generates a sequence of improving feasible solutions by moving along the usable directions and finally finds an optimal solution of the problem.

Consider the problem

$$\begin{aligned} \text{Maximize} \quad & f(X) \\ \text{Subject to} \quad & g_i(X) \leq b_i,\ i \in I_C \\ & a^T_i X \leq b_i,\ i \in I_L \\ & 0 \leq X \leq M \end{aligned} \tag{25.32}$$

where $f(X)$ and $g_i(X)$ are concave and convex functions of $X \in R^n$ respectively.

I_C is a subset of a set of integers $I = \{1,2,..m\}$, that is, $I_C \subset I$.

a_i, $i \in I_L$ are n component vectors and $I_L = I - I_C$

b_i are scalars for $i \in I_L + I_C = I$

and M is an n–component vector, some or all components of which may be infinite.

Further let $J = \{1,2...n\}$.

Then the feasible region R defined by the constraints of (25.32) is obviously convex.

It is assumed that the functions f(X) and $g_i(X)$, $i \in I_C$ are differentiable with continuous partial derivatives and further the nonlinear constraints satisfy the regularity condition:

$$\text{there exists an } X \in R, \text{ such that } g_i(X) < b_i \text{ for all } i \in I_C \tag{25.33}$$

Note that the regularity condition implies Kuhn–Tucker constraint qualification

By definition, a direction S at $X \in R$ is called feasible if we do not immediately leave the region R, when making a sufficiently small move in the direction S, that is, S is called a feasible direction at $X \in R$, if there exists a $\delta > 0$, such that $X + \lambda S \in R$, for all $\lambda \in (0, \delta)$. Further, a feasible direction S at $X \in R$ is called usable if $f(X + \lambda S) > f(X)$, for all $\lambda \in (0, \delta)$ or equivalently, if

$$\left(\frac{\delta f(X+\lambda S)}{\delta \lambda}\right)_{\lambda=0} = \nabla f(X)^T S > 0 \tag{25.34}$$

Theorem 25.3. If no feasible direction at $X \in R$ is usable then X maximizes f(X) on R.

Proof: Suppose that no feasible direction at $X \in R$ is usable and that X does not maximize f(X).

Hence there exists $Y \in R$ such that $f(Y) > f(X)$. Since f(X) is concave, we have

$$0 < f(Y) - f(X) \leq \nabla f(X)^T (Y - X).$$

Now, since Y–X is a feasible direction, it follows that it is also usable, which contradicts our assumption.

Hence X maximizes f(X) on R.

Now, for convenience, let the gradient vectors be denoted by

$$\begin{aligned} h(X) &= \nabla f(X) = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots \frac{\partial f}{\partial x_n}\right]^T \\ q_i(X) &= \nabla g_i(X) = \left[\frac{\partial g_i}{\partial x_1}, \frac{\partial g_i}{\partial x_2}, \ldots \frac{\partial fg_i}{\partial g_n}\right]^T, i \in Ic. \end{aligned} \tag{25.35}$$

and let for any $X \in R$,

$$I_C(X) = \{i \mid g_i(X) = b_i\}$$
$$I_L(X) = \{i \mid a^T_i X = b_i\}$$
$$J_1(X) = \{j \mid x_j = 0\}$$

$J_2(X) = \{j \mid x_j = M_j\}$, where $J_1(X) + J_2(X) = J$ and M_j is the j^{th} component of M. (25.36)

and further let

$$S(X) = \left\{ \begin{array}{ll} S \mid q_i(X)_S^T \le 0, & i \in Ic(X) \\ a_i^T S \le 0, & i \in I_L(X). \\ s_j \ge 0, & i \in J_1(X) \\ s_j \le 0, & i \in J_2(X). \end{array} \right\} \qquad (25.37)$$

Theorem 25.4. A feasible direction S at $X \in R$ belongs to S(X). (Exercise.)

Theorem 25.5. If $S \in S(X)$ with $q_i(X)^T S < 0$, $i \in I_C(X)$ then S is a feasible direction at $X \in R$. If further $h(X)^T S > 0$, then S is a usable feasible direction.

Proof: Let $S \in S(X)$ satisfy $q_i(X)^T S < 0$ for $i \in I_C(X)$. By differentiability of g_i, we have

$$g_i(X + \lambda S) = g_i(X) + \lambda q_i(X)^T S + \lambda \|S\| \alpha(X; \lambda S).$$

where $\alpha(X; \lambda S) \to 0$ as $\lambda \to 0$.

and since $q_i(X)^T S < 0$, then

$$g_i(X + \lambda S) < g_i(X) = b_i, \text{ for } \lambda > 0 \text{ sufficiently small.}$$

Hence for $i \in I_C(X)$, $g_i(X + \lambda S) < b_i$.

For $i \notin I_C(X)$, $g_i(X) < b_i$ and since g_i is continuous there exists a $\lambda > 0$, sufficiently small such that $g_i(X + \lambda S) \le b_i$.

Thus there exists a $\lambda > 0$ such that

$$g_i(X + \lambda S) \le b_i, \text{ for all } i \in I_C.$$

It can be easily shown that for some $\lambda > 0$, $(X + \lambda S)$ also satisfies all other constraints of the problem (25.32)

Hence there exists a $\lambda > 0$ such that $X + \lambda S \in R$ and therefore S is a feasible direction at $X \in R$. Now, since f is differentiable and $h(X)^T{}_S > 0$, by a similar argument we get

$f(X + \lambda S) > f(X)$, for $\lambda > 0$ sufficiently small which implies that S is a usable feasible direction.

Theorem 25.6. A feasible point $X \in R$ maximizes f(X) on R if and only if $h(X)^T S \le 0$ for all $S \in S(X)$.

Proof: Let S be a feasible direction at $X \in R$. Then $S \in S(X)$ by theorem 25.4. Now, for all $S \in S(X)$ $h(X)^T S \le 0$ implies that there is no feasible direction S which is usable. Hence by theorem 25.3, X maximizes f(X) on R.

Conversely, let X maximize f(X) on R and suppose that there exists an $S \in S(X)$ satisfying $h(X)^T S > 0$. By our regularity condition, we know that there exists an $\overline{X} \in R$, such that $g_i(\overline{X}) < b_i$, for all $i \in I_C$.

Let $\overline{S} = \overline{X} - X$. Since $g_i(X)$ are convex, we have

$$q_i(X)^T [(X + \alpha\, \bar{S}) - X] \le g_i(X + \alpha\, \bar{S}) - g_i(X)., \text{ for } 0 < \alpha < 1$$

or $\qquad \alpha\, q_i(X)^T\, \bar{S} \le \alpha\, [g_i(\bar{X}) - g_i(X)]$

or $\qquad q_i(X)^T\, \bar{S} < 0$, for $i \in I_C(X)$.

Now, let $S(\lambda) = \lambda\, S + \bar{S}$.

It is easy to see that

$$q_i(X)^T S(\lambda) < 0, \text{ for } i \in I_C(X).$$
$$a_i^T S(\lambda) \le 0, \text{ for } i \in I_L(X).$$
$$[S(\lambda)]_j \ge 0, \text{ for } j \in J_1(X)$$

and
$$[S(\lambda)]_j \le 0, \text{ for } j \in J_2(X)$$

Thus $S(\lambda)$ belongs to $S(X)$ with $q_i(X)^T S < 0$

Hence, by theorem 25.5, $S(\lambda)$ is a feasible direction at $X \in R$, for all $\lambda \ge 0$.

Moreover, $\qquad h(X)^T S(\lambda) = \lambda\, h(X)^T S + h(X)^T\, \bar{S}$

Since $h(X)^T S > 0$, λ can be chosen large enough so that $h(X)^T S(\lambda) > 0$, which implies that $S(\lambda)$ is a usable feasible direction and hence X cannot be a maximum point.

Theorem 25.7 A point $X \in R$ maximizes $f(X)$ on R if and only if the gradient vector in X can be expressed as

$$h(X) = \sum_{i \in I_c(x)} u_i g_i(X) + \sum_{i \in I_L(x)} u_i a_i - \sum_{j \in J_1(x)} v_j^1 e_j + \sum_{j \in J_2(X)} v_j^2 e_j \qquad (25.38)$$

where $u_i \ge 0$, $v_j^1 \ge 0$, $v_j^2 \ge 0$ and e_j is the jth unit vector.

Proof: Suppose $g(X)$ can be expressed as in (25.38). Then for any $S \in S(X)$, we have

$$h(X)^T S = \sum_{i \in I_c(x)} u_i g_i(X)^T S + \sum_{i \in I_L(x)} u_i a_i^T S - \sum_{j \in J_1(x)} v_j^1 e_j S_j + \sum_{j \in J_2(x)} v_j^2 e_j S_j \le 0$$

so that by Theorem 25.6, X maximizes $f(X)$ on R

Conversely, if X maximizes $f(X)$ on R, then by Theorem 25.6 we have $h(X)^T S \le 0$, for all $S \in S(X)$

that is, we have a solutions of

$$q_i(X)^T S \le 0,\ i \in I_C(X).$$
$$a_i^T S \le 0,\ i \in I_L(X).$$
$$s_j \le 0,\ j \in J_1(X)$$
$$s_j \le 0,\ j \in J_2(X)$$

and
$$h(X)^T S \le 0$$

Applying Farkas lemma, we then have.

$$h(X) = \sum_{i \in I_c(x)} u_i g_i(X) + \sum_{i \in I_L(x)} u_i a_i - \sum_{j \in J_1(x)} v_j^1 e_j + \sum_{j \in J_2(X)} v_j^2 e_j.$$

$$u_i \ge 0, v_j^i \ge 0, v_j^2 \ge 0$$

25.3.1. Direction Finding Problem

In order to find a usable feasible direction at a point $X \in R$, we now consider the problem

$$\begin{aligned}
&\text{Maximize} \quad \sigma \\
&\text{Subject to} \quad q_i(X)^T S + \theta_i \sigma \le 0,\ i \in I_C(X) \\
&\qquad a^T_i S \le 0,\ i \in I_L(X) \\
&\qquad s_j \ge 0,\ j \in J_1(X) \\
&\qquad s_j \le 0,\ j \in J_2(X) \\
&\qquad -h(X)^T S + \sigma \le 0
\end{aligned} \tag{25.39}$$

and a normalization requirement, where σ is an extra variable and θ_i are arbitrary positive numbers.

A normalization requirement is introduced in (25.39) to prevent a method of solution producing an infinite solution.

For any (S, σ) satisfying the constraints of (25.39) with $\sigma > 0$, S will be a usable feasible direction since $\sigma > 0$ implies that $q_i(X)^T S < 0$ for $i \in I_c(x)$ and $h(X)^T S > 0$. (Theorem 25.5) Thus if at $X \in R$, $(\hat{S}, \hat{\sigma},)$ is an optimal solution of (25.39), with $\hat{\sigma} > 0$, it will lead to the most suitable usable feasible direction and if $\hat{\sigma} = 0$, it can be shown that X maximizes f(X) on R.

Theorem 25.8 A point $X \in R$ maximizes f(X) on R if and only if the optimal value of 6 in (25.39) is zero.

Proof: The optimal value of σ in (25.39) is zero if and only if the system of inequalities

$$\begin{aligned}
&q_i(X)^T S + \theta_i \sigma \le 0,\ i \in I_C(X) \\
&a^T_i S \le 0,\ i \in I_L(X) \\
&s_j \ge 0,\ j \in J_1(X) \\
&s_j \le 0,\ j \in J_2(X) \\
&-h(X)^T S + \sigma \le 0
\end{aligned} \tag{25.40}$$

and $\qquad \sigma > 0$

has no solution

By Theorem 7.9, then there must exist a solution of

$$\sum_{i \in I_c(x)} q_i(x) u_i + \sum_{i \in I_L(x)} a_i u_i - \sum_{j \in J_1(x)} e_j v_j^1 + \sum_{j \in J_2(x)} e_j v_j^2 = u_0 h(X) \tag{25.41}$$

$$\sum_{i \in I_c(x)} \theta_i u_i + u_0 = 1. \tag{25.42}$$

$u_0 \ge 0, u_i \ge 0,\ \ i \in I_c(X) + I_L(X); v_j^1 \ge 0,\ \ j \in J_1(X)$ and

$v_j^2 \ge 0,\ j \in J_2(X).$

Now, if $u_0 > 0$, it follows from (25.41) and Theorem 25.7 that X maximizes f(X) on R. If $u_0 = 0$, then the equation (25.42) implies that $u_i > 0$ for at least one $i \in I_C(X)$ and.

$$\sum_{i\in I_c(x)} q_i(x)u_i + \sum_{i\in I_L(x)} a_i u_i - \sum_{j\in J_1(x)} e_j v_j^1 + \sum_{j\in J_2(X)} e_j v_j^2 = 0 \tag{25.43}$$

which shows that the gradient vectors of the binding constraints are not linearly independent. Hence, Kuhn–Tucker constraint qualification does not hold and thus our regularity condition is violated.

u_0 must therefore be positive and we arrive at an optimal solution of the problem. Zoutendijk suggested a number of conditions, one of which may be used as a normalization constraint in the direction finding problem (25.39)

N1:		$S^TS \leq 1$
N2:		$-1 \leq s_j \leq 1$, for all j,
N3:		$s_j \leq 1$, if $h_j(X) > 0$
		$s_j \geq -1$, if $h_j(X) < 0$
N4:	(a)	$\sigma \leq 1$, for the problem (25.39)
	(b)	$h(X)^TS \leq 1$ if the problem has no nonlinear constraint
N5:		$q_i(X)^TS + \theta_i\sigma \leq b_i - g_i(X)$, if $i\in I_C$
		$a_i^T S \leq b - a_i^T X$, if $i\in I_L$
		$-x_j \leq s_j \leq M_j - x_j$, $j\in J$.

The normalization N_1: $S^TS \leq 1$, perhaps gives the best possible feasible direction since it makes the smallest angle with the gradient vector and will in general need a fewer steps than the other normalization constraints. On the other hand, the other normalization constraints, which are linear may need more steps but less computation in each step. Therefore, the relative merits of the different normalizations can be judged only after having computational experience with them. Obviously, other constraints can be used as normalization requirement

We now discuss the computational aspect of the direction finding problem with $S^TS \leq 1$ as the normalization constraint.

The direction finding problem now becomes

$$\begin{aligned} &\text{Maximize} && \sigma \\ &\text{Subject to} && q_i(X)^TS + \theta_i\sigma \leq 0,\ i \in I_C(X) \\ &&& a_i^T S \leq 0,\ i \in I_L(X) \\ &&& s_j \geq 0,\ j \in J_1(X) \\ &&& s_j \leq 0,\ j \in J_2(X) \\ &&& -h(X)^TS + \sigma \leq 0 \\ &&& S^TS \leq 1 \end{aligned} \tag{25.44}$$

The problem (25.44) can be expressed in the form

$$\begin{aligned} &\text{Maximize} && \sigma \\ &\text{Subject to} && QS + \theta\,\sigma \le 0 \\ & && -h(X)^T S + \sigma \le 0 \\ & && S^T S \le 1. \end{aligned} \tag{25.45}$$

where Q = Q(X) is the matrix with rows

$q_i(X)^T,\ i \in I_C(X)$; $a^T_i,\ i \in I_L(X)$; $e_j,\ j \in J_1(X)$; $e_j,\ j \in J_2(X)$

and θ is a vector with components $\theta_i > 0$, for $i \in I_C(X)$ and equal to zero otherwise

Theorem 25.9. If (S_1, σ_1) is an optimal solution of the problem (25.45), it is proportional to the optimal solution (S_2, σ_2) of the problem

$$\begin{aligned} &\text{Maximize} && \sigma \\ &\text{Subject to} && QS + \theta\,\sigma \le 0 \\ & && -h(X)^T S + \sigma \le 0 \\ & && S^T S + \sigma^2 \le 1 \end{aligned} \tag{25.46}$$

Proof: It is clear that $S_1/\sqrt{1+\sigma_1^2}$, $\sigma_1/\sqrt{1+\sigma_1^2}$ is feasible to (25.46) and that $S_2/\sqrt{1-\sigma_2^2}$, $\sigma_1/\sqrt{1-\sigma_2^2}$ is feasible to (25.45).

Hence

$$\sigma_1 \ge \frac{\sigma_2}{\sqrt{1-\sigma_2^2}} \tag{25.47}$$

$$\sigma_2 \ge \frac{\sigma_1}{\sqrt{1+\sigma_1^2}} \tag{25.48}$$

From (25.48), we get
$$\sigma_1^2 \le \frac{\sigma_2^2}{1+\sigma_2^2} \tag{25.49}$$

(25.47) and (25.49) then imply that

$$\sigma_1 = \frac{\sigma_2}{\sqrt{1+\sigma_2^2}}\ ,\quad S_1 \frac{S_2}{\sqrt{1-\sigma_2^2}} \tag{25.50}$$

Similarly, we have
$$\sigma_2 = \frac{\sigma_1}{\sqrt{1+\sigma_1^2}}\ ,\quad S_2 \frac{S_1}{\sqrt{1+\sigma_1^2}} \tag{25.51}$$

This proves the theorem.

From Theorem 25.9, we note that if $\sigma_2 = 0$, then $\sigma_1 = 0$ and then $S_1 = 0$, $\sigma_1 = 0$ is an optimal solution of (25.45) and we arrive at an optimal solution of the original problem. If $\sigma_2 > 0$, σ_1, is positive and S_1 is then the usable feasible direction.

Thus through a solution of the problem (25.46), we can obtain a solution of the direction finding problem.

The problem (25.46) is a convex programming problem with a linear objective function and linear and one quadratic constraints. It can however, be shown that this problem can be solved by any method for solving quadratic programming problems.

The problem (25.46) is of the form

$$\begin{aligned} &\text{Maximize} && P^T X. \\ &\text{Subject to} && AX \leq 0 \\ &&& X^T X \leq 1 \end{aligned} \tag{25.52}$$

where $P^T = (0, 1)$; $X^T = (S, \sigma)$ and

$$A = \begin{bmatrix} Q & \theta \\ -h^T & 1 \end{bmatrix}$$

Now, consider the problem

$$\begin{aligned} &\text{Minimize} && X^T X. \\ &\text{Subject to} && AX \leq 0 \\ &&& P^T X = 1. \end{aligned} \tag{25.53}$$

It can be easily seen that if the problem (25.53) is not feasible then the problem (25.52) must have an optimal solution with $P^T X = 0$.

Suppose now that X_1 is an optimal solution of (25.52) with $P^T X_1 > 0$. In that case an optimal solution of the problem (25.52) can be obtained from an optimal solution of (25.53) which is a quadratic programming problem.

Theorem 25.10 If the problem (25.52) is feasible with $P^T X > 0$, then an optimal solution of the problem can be obtained from an optimal solution of the problem (25.53)

Proof: Let X_1 be an optimal solution of (25.52) Then $P^T X_1 > 0$ and for $\beta = 1/P^T X_1$, βX_1 is feasible for (25.53). Suppose X_2 is an optimal solution of (25.53). Then

$$\alpha^2 = X_2^T X_2 \leq \beta^2 (X_1^T X_1) \leq \beta^2 .$$

or
$$\alpha \leq \beta.$$

It is clear that X_2/α is feasible fo: (25.52) and hence

$$\frac{P^T X_2}{\alpha} = \frac{1}{\alpha} \geq \frac{1}{\beta} = \frac{\beta P^T X_1}{\beta} = P^T X_1$$

Since X_1 is optimal for (25.52), α must be equal to β.

Hence

$$\frac{X_2}{\alpha} = \frac{X_2}{(X_2^T X_2)^{1/2}}$$

must be an optimal Solution of (25.52)

We also note that if X_1 is an optimal solution of (25.52) and $P^TX_1 > 0$ then $X_1^TX_1 = 1$ and further X_1 is unique.

25.3.2. Determination of the Length of the Steps

After finding a usable feasible direction S at $X \in R$, we now want to determine λ, the length of the step to be taken in the direction S, so that $X + \lambda S$ is an improved solution.

The step length λ can be obtained by solving the problem

$$\begin{aligned} &\text{Maximize} \quad f(X + \lambda S) \\ &\text{Subject to} \quad 0 \le \lambda \le \lambda_0 \end{aligned} \tag{25.54}$$

where $\lambda_0 = \text{Max}\ \{\lambda \mid X + \lambda S \in R\}$.

Since λ is the only variable involved in the problem (X and S are known vectors), λ_0 can be easily obtained, if all the constraints are linear. In the general case, we find the largest root of $g_i(X + \lambda S) = b_i$, for each $i \in I_C$ and then take the smallest of the figures obtained. This can be done by using some numerical technique such as Newton's method. Let this value of λ be denoted by λ_c'.

For the linear constraints, we define

$$\lambda_L' = \text{Min}\ (\lambda_1', \lambda_2', \lambda_3') \tag{25.55}$$

where $$\lambda_L' = \underset{i}{\text{Min}}\left\{ \frac{b_i - a_i^T X}{a_i^T S} \,\middle|\, a_i^T S > 0 \ \text{ for } i \in I_L - I_L(X) \right\}.$$

$$\lambda_2' = \underset{j}{\text{Min}}\left\{ \frac{x_j}{-s_j} \,\middle|\, s_j < 0, \ \ j \in J - J_1(X) \right\}$$

$$\lambda_3' = \underset{j}{\text{Min}}\left\{ \frac{M_j - x_j}{s_j} \,\middle|\, s_j > 0, \ \ j \in J - J_2(X) \right\}$$

$$\lambda_0 \text{ is then obtained as } \lambda_0 = \text{Min}\ (\lambda_c', \lambda_L') \tag{25.56}$$

Note that the value of λ_0 may be infinite if some of the M_j are infinite and if no finite value of λ is optimal for (25.54), the original problem has an unbounded solution.

It now remains to find a starting feasible solution to the problem to initiate the algorithm.

25.3.3. Finding an Initial Feasible Solution

Frequently, a feasible solution to the problem is immediately available. If however, an initial feasible solution is not readily available, we can obtain it as follows.

The solution is obtained in two phases.

Phase 1: Take any point $\overline{X}$ satisfying $0 \le \overline{X} \le M$ and solve the problem

$$\begin{aligned} &\text{Maximize} && -\xi_1 \\ &\text{Subject to} && a_i^T X - \rho_i \xi_1 \le b_i, \ i \in I_L \\ &&& 0 \le X \le M, \ \xi_1 \ge 0 \end{aligned} \tag{25.57}$$

where ξ_1 is an extra variable and ρ_i, $i \in I_L$ are nonnegative numbers with $\rho_i = 0$ if $a_i^T X \le b_i$ and $\rho_i > 0$ if $a_i^T X > b_i$.

This is a linear programming problem and can be solved by the simplex method. An initial feasible solution to this problem is given by

$$X = \overline{X}, \quad \xi_1 = \text{Max}\left[\frac{a_i^T \overline{X} - b_i}{\rho_i} \middle| \rho_i > 0\right]. \tag{25.58}$$

If the constraint set R is not empty we will obtain an optimal solution ($\overline{X}^1$, $\xi_1 = 0$) and $\overline{X}^1$ will then satisfy the linear constraints of the original problem. If now $g_i(\overline{X}^1) \le b_i$, for all $i \in I_C$, $\overline{X}^1 \in R$ and an initial feasible solution to the problem is obtained. If $g_i(X^1) > b_i$, for same $i \in I_C$, we proceed to phase 2:

Phase 2: We now solve the problem

$$\begin{aligned} &\text{Maximize} && -\xi_2 \\ &\text{Subject to} && g_i(X) - \mu_i \xi_2 \le b_i, \ i \in I_C \\ &&& a_i^T X \le b_i, \ i \in I_L \\ &&& 0 \le X \le M, \ \xi_2 \ge 0. \end{aligned} \tag{25.59}$$

where ξ_2 is an extra variable and μ_i $i \in I_C$ are nonnegative numbers with $\mu_i = 0$, if $g_i(\overline{X}^1) \le b_i$ and $\mu_i > 0$, if $g_i(\overline{X}^1) > b_i$.

An initial feasible solution to the problem (25.59) is

$$X^1 = \overline{X}^1$$

$$\xi_2 = \text{Max}\left[\frac{f_i(\overline{X}^1) - b_i}{\mu_i} \middle| \mu_i > 0\right] \tag{25.60}$$

Since the problem (25.59) is a convex programming problem of the original type, the method of feasible directions can be applied to obtain an optimal solution to this problem with $\xi_2 = 0$, which can be taken as an initial feasible solution to the original problem.

25.3.4. Summary of the Algorithm

Step 1. Start with a feasible point X^0 and suppose that the algorithm has produced a sequence of feasible points $X^0, X^1, .. X^k$.

Step 2. Solve the kth direction finding problem, Let $\hat{S}^k$, $\hat{\sigma}^k$ be the optimal solution.

Step 3. (a). If $\hat{\sigma}^k = 0$, X^k is an optimal solution to the original problem and the process terminates

(b). If $\hat{\sigma}^k > 0$, $S^k = \hat{S}^k$ is a usable feasible direction at X^k.

Step 4. Determine the step length λ^k.

(a) If $\lambda^k = \infty$, the original problem is unbounded.

(b) If $\lambda^k < \infty$, determine $X^{k+1} = X^k + \lambda^k S^k$ and proceed to the next iteration.

Step 5. Repeat the process from step 2.

25.3.5. Convergence of the Procedure

In Zoutendijk's method of feasible directions discussed above, convergence is not generally guaranteed. Recall that in the direction finding problem only the binding constraints are considered. The usable feasible directions S^k at X^k are thus chosen by examining only the binding constraints at X^k, without any regard to other constraints. But if the point X^k is close to any boundaries of the constraints not considered, the directions, S^k thus chosen might point toward these close boundaries. The algorithm then generates a sequence of points $\{X^k\}$ such that the entire sequence converges to the corner point formed by these close boundaries thus causing jamming (zigzagging) at a nonoptimal point.

It is therefore necessary to apply some device to avoid jamming in the direction finding problems to ensure convergence of f(X) to an optimal solution. One such device is the '$\in$-perturbation' method where feasible directions are determined by considering all constraints which are binding or almost binding at X^k.

The $\in$–Perturbation Method

Let $\in > 0$ be an arbitrary number.

For $X \in R$, we now define.

$$\begin{aligned}
I_C(X,\in) &= \{\, i \in I_C \mid b_i - \in \le g_i(X) \le b_i \} \\
I_L(X,\in) &= \{\, i \in I_L \mid b_i - \in \le a^T_i X \le b_i \} \\
J_1(X,\in) &= \{\, j \in J \mid 0 \le x_j \le \in \}. \\
J_2(X,\in) &= \{\, j \in J \mid M_j - \in \le x_j \le M_j \}.
\end{aligned} \tag{25.61}$$

The direction finding problem (25.39) now becomes,

$$\begin{aligned}
\text{Maximize} \quad & \sigma \\
\text{Subject to} \quad & q_i(X)^T S + \theta_i \sigma \le 0 , \ i \in I_C(X,\in) \\
& a^T_i S \le 0 , \ i \in I_L(X,\in) \\
& s_j \ge 0 , \ j \in J_1(X,\in)
\end{aligned}$$

$$s_j \leq 0\,, j \in J_2\,(X,\epsilon)$$
$$-h(X)^T S + \sigma \leq 0 \tag{25.62}$$
$$S^T S \leq 1$$

If $(\hat{S},\hat{\sigma})$ be the optimal solution of (25.62) with $\hat{\sigma} < \epsilon$, we replace ϵ by $\epsilon/2$ and solve the direction finding problem again. If $\hat{\sigma} \leq \delta$ a pre-specified small positive number called the termination constant, then the process terminates and X is an optimal solution of the original problem. Otherwise, $\hat{S}$ will be the usable feasible direction with no danger of jamming.

For proofs and detail discussion on convergence, refer to Zoutendijk [563] and Zangwill [550]. Also, see Topkis and Veinott [471]

25.3.6. Summary of the Algorithm using Perturbation Method

1. Start with a feasible point X^0 and suppose that the algorithm has produced a sequence of feasible points X^0, X^1,.. X^k and perturbation constants ϵ_1, ϵ_2,.. ϵ_k.
2. Solve the kth direction finding problem (25.62) using $\epsilon = \epsilon_k$. Let $\hat{S}^k$, $\hat{\sigma}^k$ be the optimal solution.
3. a) If $\hat{\sigma}^k < \epsilon_k$, then replace ϵ_k by $\epsilon_k/2$ and solve the direction finding problem again.
   b) If $\hat{\sigma}^k \leq \delta$ (the termination constant), the process terminates and X^k is an optimal solution of the original problem; otherwise set $\epsilon_{k+1} = \epsilon_k/2$ and proceed to next step.
   c) If $\hat{\sigma}^k \geq \epsilon_k$, then $S^k = \hat{S}^k$ is a usable feasible direction at X^k.
4. Determine the step length λ^k
   If $\lambda^k = \infty$, the original problem is unbounded and the process terminates; otherwise $X^{k+1} = X^k + \lambda^k S^k$ and proceed to the next iteration.
5. Repeat the process from step 2.

25.3.7. The Case of Linear Constraints

Consider the problem

$$\begin{aligned} \text{Maximize}\quad & f(X) \\ \text{Subject to}\quad & AX \leq b \\ & EX = d \\ & 0 \leq X \leq M \end{aligned} \tag{25.63}$$

where f(X) is a concave differentiable function of $X \in R^n$, A is an $m \times n$ matrix, E is an $l \times n$ matrix, b is an m-vector, d is an l-vector and M is a vector of n components, some or all of which may be infinite. Let R denote the feasible set of the problem.

It can be seen from the previous result that in this case the following steps are to be followed to obtain a solution of the problem.

Step 1. Find an initial feasible solution X^0 to the problem and suppose that the algorithm has produced a sequence of feasible points $X^0, X^1, \ldots X^k$.

Step 2. At X^k, solve the direction finding problem. The direction finding problem, in this case reduces to

$$\begin{aligned}
&\text{Maximize} && h(X^k)^T S \\
&\text{Subject to} && A_i S \leq 0\,,\ i \in I_L(X). \\
& && ES = 0 \\
& && s_j \geq 0\,,\ j \in J_1(X) \\
& && s_j \leq 0,\ j \in J_2(X).
\end{aligned} \tag{25.64}$$

and one of the normalization constraints N_1, N_2, N_3, N_4 (b).

Let S^k be an optimal solution of (25.64)

(a) If $h(X^k)^T S^k = 0$, then X^k is an optimal solution of (25.63) and the algorithm terminates.

(b) If $h(X^k)^T S^k > 0$, S^k is a usable feasible direction and go to step 3.

Step 3. Determine the step length λ^k by solving

$$\begin{aligned}
&\text{Maximize} && f(X^k + \lambda\, S^k) \\
&\text{Subject to} && 0 \leq \lambda \leq \lambda_0
\end{aligned}$$

where $\lambda_0 = \max\, \{\lambda \mid X^k + \lambda\, S^k \in R\}$.

(a) If $\lambda^k = \infty$, the problem (25.63) is unbounded

(b) If $\lambda^k < \infty$, determine $X^{k+1} = X^k + \lambda^k S^k$ and proceed to the next step.

Step 4. Repeat the process from step 2.

25.3.8. Example

Consider the problem

$$\begin{aligned}
&\text{Maximize} && f(X) = 3x_1 + 2x_2 + x_1x_2 - x^2{}_1 - x^2{}_2 \\
&\text{Subject to} && 5x_1 + x_2 \leq 5 \\
& && x_1 + x_2 \leq 2 \\
& && x_1 \geq 0 \\
& && x_2 \geq 0
\end{aligned} \tag{25.65}$$

Note that the constraints are linear only.

$$\nabla f(X) = h(X) = \begin{bmatrix} 3 - 2x_1 + x_2 \\ 2 + x_1 - 2x_2 \end{bmatrix}$$

Let the initial point be

$$X^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Direction Finding Problem: At the point X^0, the third and the fourth constraints are binding and the direction finding problem is then

$$\begin{aligned} \text{Maximize} \quad & h(X^0)^T S = 3s_1 + 2s_2 \\ \text{Subject to} \quad & -s_1 \le 0. \\ & -s_2 \le 0. \\ & -1 \le s_1 \le 1. \\ & -1 \le S_2 \le 1. \end{aligned} \tag{25.66}$$

where we use $-1 \le S \le 1$ as our normalization constraint.

The optimal solution is $s_1^0 = 1$, $s_2^0 = 1$, and $h(X^0)^T S^0 > 0$. The direction S^0 therefore, is a usable feasible direction.

Finding Step Length

$$\lambda_0 = \operatorname*{Min}_i \left[\frac{b_i - a_i^T X^0}{a_i^T S^0} \middle| \; a_i^T S^0 > 0 \right]. \text{ for nonbinding constraints}$$

$$= \text{Min}\left[\frac{5-0}{6}, \frac{2-0}{2}\right] = 5/6.$$

The value of λ^0 is then obtained by solving

$$\begin{aligned} \text{Maximize} \quad & f(X^0 + \lambda S^0) = 5\lambda - \lambda^2 \\ \text{Subject to} \quad & 0 \le \lambda \le 5/6. \end{aligned}$$

Now, $df(X^0 + \lambda S^0)/d\lambda = 0 \Rightarrow 5 - 2\lambda = 0$ so that $\lambda_{opt} = 5/2$.

Hence the desired steps length $\lambda^0 = 5/6$ and

$$X^1 = X^0 + \lambda^0 S^0 \begin{pmatrix} 5/6 \\ 5/6 \end{pmatrix}.$$

$$\text{At } X^1 = \begin{pmatrix} 5/6 \\ 5/6 \end{pmatrix}, h(X^1) \begin{pmatrix} 13/6 \\ 7/6 \end{pmatrix} \text{ and}$$

the first constraint is binding.

The direction finding problem is then

$$\begin{aligned} \text{Maximize} \quad & h(X^1)^T S = 13/6\, s_1 + 7/6\, s_2 \\ \text{Subject to} \quad & 5s_1 + s_2 \le 0. \\ & -1 \le s_1 \le 1 \\ & -1 \le s_2 \le 1. \end{aligned} \tag{25.67}$$

It is clear from the Figure 25.1 that the optimal solution to the above problem is $s_1 = -1/5$, $s_2 = 1$ and $h(X^1)^T S^1 > 0$, so that the direction S^1 is usable.

Finding Step Length

$$\lambda_0 = \text{Min}\left[\frac{2-2X5/6}{4/5}, \frac{5/6}{1/5}\right]$$

$$= \text{Min}\left[\frac{5}{12}, \frac{25}{6}\right] = \frac{5}{12}$$

The value of λ^1 is obtained by solving.

Maximize $f(X^1 + \lambda S^1) = 22/30\ \lambda - 31/25\ \lambda^2$

Subject to $0 \le \lambda \le 5/12$

$$\frac{df}{d\lambda} = 0 \Rightarrow \lambda = \frac{55}{186}$$

Thus $\lambda^1 = 55/186$ and

$$X^2 = X^1 + \lambda^1 S^1 \begin{pmatrix} 24/31 \\ 35/31 \end{pmatrix}$$

Direction Finding Problem

At X^2, the first constraint is binding and $h(X^2)^T S = 80/31\ s_1 + 16/31\ s_2$

Hence we

Maximize $h(X^2)^T S = 80/31\ s_1 + 16/31\ s_2$

Subject to
$$5s_1 + s_2 \le 0. \tag{25.68}$$
$$-1 \le s_1 \le 1$$
$$-1 \le s_2 \le 1.$$

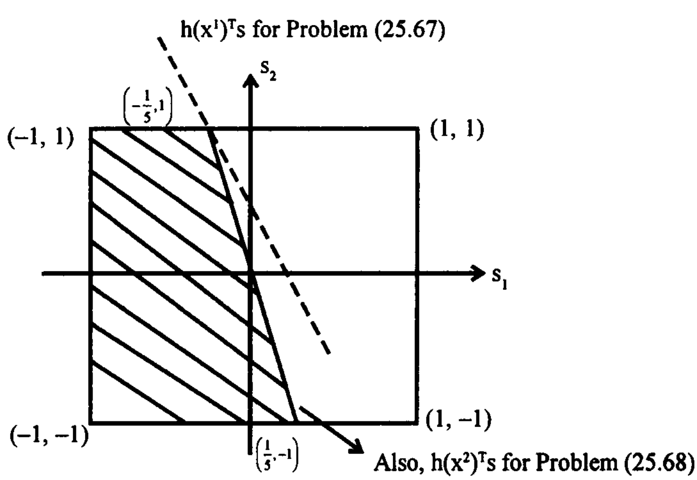


Figure 25.2

From the Figure 25.2, we note that $s^2 = (-1/5, 1)^T$ and $s^2 = (1/5, -1)^T$ and every convex combination of these optimizes the above problem. Hence there exist multiple optimal solutions of the problem and the optimal value of $h(X^2)^T S = 0$. Hence $X^2 = (24/31, 35/31)$ is an optimal solution to the original problem.

25.4. Rosen's Gradient Projection Method

In this section we shall discuss another feasible direction method for maximizing a concave function under linear constraints—the gradient projection method developed by J. B. Rosen [382].

For an unconstrained maximization problem, one can improve the value of the function by moving in the direction of its gradient. However, in the presence of constraints, this method may lead to infeasible points. Zoutendijk therefore considered a vector which makes an acute angle with the gradient of the objective function as the usable feasible direction. Rosen's method is to project the gradient onto the boundary of the feasible domain and then proceed in the direction of this projection.

Consider the problem,

$$\begin{aligned} &\text{Maximize} && f(X) \\ &\text{Subject to} && A^T X \leq b \end{aligned} \tag{25.69}$$

where $f(X)$ is a differentiable concave function of $X \in R^n$, $A = (a_1, a_2, .. a_m)$, $a_i \in R^n$ and b is an m-component vector.

Then the feasible region defined by the constraints of (25.69) is convex.

Suppose that starting with a feasible point X^0, after k iterations we obtain the feasible solution X^k, which is a boundary point and lies on exactly q hyperplanes. Without loss of generality, we assume that these are the first q hyperplanes, so that

$$a^T_i X^k - b_i = 0 \text{ , } i = 1,2..q \tag{25.70}$$

$$a^T_i X^k - b_i < 0 \text{ , } i = q+1..m \tag{25.71}$$

Now, (25.70) can be conveniently expressed as

$$A^T_q X^k - b^q = 0. \tag{25.72}$$

where $A_q = (a_1, a_2..a_q)$ and $b^q = (b_1, b_2, b_q)^T$.

Assume that $a_1, a_2, . a_q$ are linearly independent so that the binding constraints at X^k are linearly independent.

Let V denote the $(n - q)$ dimensional linear subspace defined by the intersection of the q hyperplanes (25.72) and since the vectors $a_1, a_2, .. a_q$ are normal to V, the q-dimensional linear subspace spanned by the columns of A_q is orthogonal to V. Let this be denoted by . $\hat{V}$. $\hat{V}$ is thus the set of points.

$$X = \sum_{i=1}^{q} a_i u_i = A_q U \tag{24.73}$$

where $U = (u_1, u_2, ... u_q)^T$

V and $\hat{V}$ are orthogonal to each other and together they span the whole space R^n. Any vector $X \in R^n$ can then be expressed uniquely as

$$X = X_V + X_{\hat{V}} \tag{25.74}$$

where $X_{\hat{V}}$ lies in $\hat{V}$ and X_V is in the linear subspace parallel to V and by the orthoganility of V and $\hat{V}$

$$X_V^T X_{\hat{V}} = 0$$

Before we proceed to discuss Rosen's algorithm, we first consider the definition of a projection matrix.

25.4.1. The Projection Matrix

An $n \times n$ matrix P is called a projection matrix if $P = P^T$ and $PP = P$.

It can be easily seen that

(i) a projection matrix is positive semidefinite and

(ii) P is a projection matrix if and only if $Q = I - P$ is a projection matrix.

Define the matrix

$$\hat{P} = A_q (A_q^T A_q)^{-1} A_q^T$$

Since the columns of A_q are linearly independent, the $q \times q$ symmetric matrix $Aq^T A_q$ is nonsingular and therefore its inverse $(A^T_q A_q)^{-1}$ exists.

Lemma 25.2 The matrix $\hat{P}_q$ is a projection matrix which projects any vector $X \in R^n$ into $\hat{V}$.

Proof: $\hat{P}_q$ is symmetric and $\hat{P}_q \; \hat{P}_q = \hat{P}_q$ and hence $\hat{P}_q$ is a projection matrix.

Let $X \in R^n$, then

$$\hat{P}_q X = \left[A_q (A_q^T A_q)^{-1} A_q^T\right](X_v + X_{\hat{v}}).$$

Since X_V is orthogonal to all vectors in $\hat{V}$, $A^T_q X_V = 0$ and we have

$$\begin{aligned} \hat{P}_q X &= A_q (A_q^T A_q)^{-1} A_q^T X_{\hat{v}} \\ &= A_q (A_q^T A_q)^{-1} A_q^T . A_q U \qquad (25.76) \\ &= A_q U = X_{\hat{v}}. \end{aligned}$$

Theorem 25.11. $P_q = I - \hat{P}_q$ is a projection matrix which projects any vector $X \in R^n$ into the orthogonal complement of $\hat{V}$.

Proof: From (25.74) we have

$$X_V = X - X_{\hat{V}}$$
$$= X - \hat{P}_q X$$
$$= (I - \hat{P}_q)X = P_q X$$

Thus $$P_q X = X_V \tag{25.77}$$
where $P_q = I - A_q(A^T_q A_q)^{-1} A^T_q$

and X_V is the projection of X into the orthogonal complement of the subspace $\hat{V}$, generated by the columns of A_q. Also, we note that $P_q X_{\hat{V}} = 0$ and $P_q X_V = X_V$.

Corollary 25.4.1

A necessary and sufficient condition that a nonzero vector X be linearly independent of $a_1, a_2, \ldots a_q$ is that $P_q X \neq 0$.

25.4.2. Rosen's Algorithm

We now describe the optimization procedure for the problem (25.69) as developed by Rosen.

It has already been pointed out that in the presence of constraints, moving along the direction of the gradient of the objective function may lead to infeasible points. Rosen therefore considered the projection of the gradient onto the boundary of the feasible domain as the feasible direction, moving in which it is possible to improve the value of the objective function.

Let for convenience, the gradient of the objective function be denoted by h(X), so that

$$\nabla f(X) = h(X)\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots \frac{\partial f}{\partial x_n}\right)^T. \tag{25.78}$$

Theorem 25.12 Let X^k be a boundary point of the feasible domain R which lies on exactly q $(1 \leq q \leq n)$ hyperplanes which are assumed to be linearly independent. Then X^k is an optimal solution to the problem (25.69) if and only if

$$P_q h(X^k) = 0 \tag{25.79}$$
and $$(A^T_q A_q)^{-1} A^T_q h(X^k) \geq 0 \tag{25.80}$$

Proof: The condition (25.79) implies that h (X^k) is orthogonal to the subspace V and hence lies in $\hat{V}$. $h(X^k)$ can therefore be expressed uniquely as

$$h(X^k) = A_q U \tag{25.81}$$
where $U = (u_1, u_2, \ldots u_q)^T$

Substituting (25.81) in (25.80) we find that the condition (25.80) implies that

$$U \geq 0 \tag{25.82}$$

Thus the conditions (25.79), (25.80) reduce to

$$h(X^k) - A_q U = 0 \tag{25.83}$$

$$U \geq 0. \tag{25.84}$$

Now taking u_l corresponding to nonbinding constraints (25.71) to be zero, we find that the relations (25.83), (25.84) along with (25.70), (25.71) satisfy the K–T necessary and sufficient conditions for optimality and hence X^k is a global optimal solution to the problem.

It should be noted that if q = 0, that is if X^k is an interior point of the feasible domain, the optimality criterion reduces to

$$h(X^k) = \nabla f(X^k) = 0. \tag{25.85}$$

25.4.3. Determination of Usable Feasible Directions

Suppose now that the conditions for optimality are not satisfied. Two cases may arise.

(i) $P_q h(X^k) \neq 0$

(ii) $P_q h(X^k) = 0$, but $u_i < 0$ for at least one i.

It will be shown below that in such cases a feasible direction can be found, moving in which a feasible point can be obtained with an improved value of the objective function.

Case (i):

Theorem 25.13: Let X^k be a feasible solution to the problem (25.69), at which the first q constraints are binding and suppose that the vectors $a_1, a_2, \dots a_q$ are linearly independent. Let $P_q h(X^k) \neq 0$. Then $P_q h(X^k)$ is a usable feasible direction at X^k.

Proof: Let $P_q h(X^k) = S^k$, $S^k \neq 0$.

and consider the point

$$X^{k+1} = X^k + \lambda S^k, \text{ where } \lambda \text{ is a scalar.}$$

For i = 1,2,...q,

$$\begin{aligned} a^T_i (X^k + \lambda S^k) &= a^T_i X^k + \lambda a^T_i S^k \\ &= a^T_i X^k + \lambda a^T_i P_q h(X^k) \\ &= a^T_i X^k, \text{ since } A^T_q P_q = 0 \\ &= b_i, \text{ for all } \lambda. \end{aligned} \tag{25.86}$$

For i = q + 1,...m

$$a^T_i (X^k + \lambda S^k) = a^T_i X^k + \lambda a^T_i S^k$$

Since $a_i^T X^k < b_i$, there exists a $\lambda > 0$, such that

$$a^T_i (X^k + \lambda S^k) \leq b_i \tag{25.87}$$

Thus S^k is a feasible direction.

Now, $\quad \nabla f(X^k)^T S^k = h(X^k)^T P_q h(X^k)$

$$= h(X^k)^T P_q^T P_q h(X^k).$$

$$= (S^k)^T S^k > 0. \tag{25.88}$$

Thus (25.88), [see section 25.3] implies that $P_q h(X^k) = S^k$ is a usable feasible direction.

25.4.4. Determination of the Length of the Steps

Having found the usable feasible direction S^k at X^k, we now determine the step length λ, that will yield the greatest increase in the value of the objective function.

It is easily seen that the largest value of λ for which $X^k + \lambda S^k$ is a feasible point is given by

$$\lambda' = \underset{i}{\text{Min}} \left\{ \frac{b_i - a_i^T X^k}{a_i^T S^k} \middle| \; a_i^T S^k > 0, \; i = q+1, \ldots m \right\} > 0 \tag{25.89}$$

If however, $a_i^T S^k \leq 0$, for all nonbinding constraints, $\lambda' = \infty$

We then determine a λ in the interval $0 \leq \lambda \leq \lambda'$, so that at $(X^k + \lambda S^k)$, the objective function has the maximum possible value.

To find the desired step length λ^k, we solve the one-dimensional problem

$$\begin{array}{ll} \text{Maximize} & f(X^k + \lambda^k) \\ \text{Subject to} & 0 \leq \lambda \leq \lambda' \end{array} \tag{25.90}$$

If $\lambda' = \infty$, and no finite value of λ is optimal for the problem (25.90), the original problem has an unbounded solution. Otherwise, we take the next feasible point as $X^{k+1} = X^k + \lambda^k S^k$.

Case (ii): We now consider the case when $P_q h(X^k) = 0$ but $u_i < 0$, for at least one i (i = 1,2,..q). The current solution X^k is then not an optimal solution and we therefore proceed to find a usable feasible direction at X^k. We select one of the indices i for which $u_i < 0$, say i = q, so that $u_q < 0$ (in practice however, the index i is selected by some rule such as selecting i for which $|a_i|\, u_i$ is most negative) and then disregard the q th hyperplane as if X^k lies on the first (q – 1) hyperplanes only. We now show that, we can then find a usable feasible direction at X^k, so that moving along this direction a better solution can be obtained.

Theorem 25.14. Let X^k be a feasible solution to the problem (25.69) at which the first q constraints are binding and suppose that the vectors $a_1, a_2, \ldots a_q$ are linearly independent. Let $P_q h(X^k) = 0$ and $u_q < 0$. If the qth hyperplane is relaxed and it is considered as if at X^k only the first (q – 1) constraints are binding, then $P_{q-1} h(X^k)$ is a usable feasible direction where P_{q-1} is the projection matrix associated with A_{q-1}, that is $P_{q-1} = I - A_{q-1} (A^T_{q-1} A_{q-1})^{-1} A^T_{q-1}$

Proof: Since $P_{q-1}A_{q-1} = 0$ and a_q is independent of $a_1, a_2, \ldots a_{q-1}$, we have $P_{q-1}a_q \neq 0$ and thus

$$P_{q-1} h(X^k) = P_{q-1} A_q U$$

$$= U_q P_{q-1} a_q \neq 0 \tag{25.91}$$

Let $P_{q-1}h(X^k) = S^k$, $S^k \neq 0$, and consider the point $X^{k+1} = X^k + \lambda S^k$, where λ is a scalar

Then, for i = 1,2,..q–1

$$a^T_i (X^k + \lambda S^k) = a^T_i X^k + \lambda a^T_i P_{q-1} h(X^k)$$
$$= a^T_i X^k \text{, since } A^T_{q-1} P_{q-1} = 0$$
$$= b_i \text{, for all } \lambda.$$

and for i = q+1,...m

$$a^T_i (X^k + \lambda S^k) = a^T_i X^k + \lambda a^T_i S^k$$

Since $a_i^T X^k < b_i$, there exists a $\lambda > 0$, such that $a_i^T(X^k + \lambda S^k) \leq b_i$.

Moreover, for i = q,

$$a^T_q (X^k + \lambda S^k) = a^T_q X^k + \lambda a^T_q S^k$$
$$= a^T_q X^k + \lambda a^T_q P_{q-1} h(X^k)$$
$$= a^T_q X^k + \lambda u_q a^T_q P_{q-1} a_q \text{, by (25.91)}$$
$$= a^T_q X^k + \lambda u_q a^T_q P^T_{q-1} P_{q-1} a_q .$$
$$= a^T_q X^k + \lambda u_q \| P_{q-1} a_q \|^2 .$$
$$\leq b_i \text{, for } \lambda \geq 0$$

Hence S^k is a feasible direction.

As in case (i), it can now be shown that $\nabla f(X)^T S^k > 0$, so that S^k is a usable feasible direction at X^k and we proceed to find the optimal step length λ^k, so that $(X^k + \lambda^k S^k)$ yields the maximum possible value of the objective function.

25.4.5. Summary of the Algorithm

The various steps of the gradient projection method can now be stated as follows.

Step 1: Find a feasible point X^0 for the problem. After k-iterations, suppose that the feasible point X^k is obtained.

Step 2: Identify the q independent constraints that are binding at X^k.

Step 3: Compute P_q, $P_q h(X^k)$ and U

Step 4 (a): If $P_q h(X^k) = 0$ and $U \geq 0$, then X^k is an optimal solution and the process terminates.

Step 4 (b): If $P_q h(X^k) \neq 0$, $S^k = P_q h(X^k)$ is a usable feasible direction and go to step 5.

Step 4 (c): If $P_q h(X^k) = 0$ and $u_i < 0$ for at least one i, select any one of them say $u_q < 0$ and disregard the qth binding constraint. Then $S^k = P_{q-1} h(X^k)$ is a usable feasible direction and go to step 5.

Step 5: Determine the largest permitted step length λ' by (25.89)

Step 6: Find the desired step length λ^k by solving

$$\text{Maximize} \quad f(X^k + \lambda S^k)$$

Subject to $0 \le \lambda \le \lambda'$.

If $\lambda' = \infty$ and no finite value of λ is optimal, the original problem has an unbounded solution. Otherwise set $X^{k+1} = X^k + \lambda^k S^k$ and go to step 2 and repeat the process.

To this day, no convergence proof for Rosen's method is available, neither is there a counter example. Indeed, the method has been used successfully for numerous nonlinear programming problems and jamming has never been detected. Readers interested in detail discussions convergence theory should consult Zangwill [550] and Luemberger [306].

We have discussed the gradient projection method for maximizing a concave objective function under linear constraints. It should be noted that if the objective function is a general nonlinear differentiable function, the conditions (25.83), (25.84), (25.70), (25.71) are the K–T necessary conditions for optimality and thus any feasible point for the problem satisfying them is a local optimal point.

The gradient projection method can also be applied to a more general case, where nonlinearity occurs in both the objective function and the constraints. For a discussion on such a case, the reader is referred to Rosen [382]. Also, see Goldfarb [201] and Davies [123].

25.4.6. Example

Consider the problem,

$$\begin{aligned} \text{Maximize} \quad & f(X) = 3x_1 + 2x_2 - x_1^2 - x_2^2 + x_1x_2 \\ \text{Subject to} \quad & 5x_1 + x_2 \le 5 \\ & x_1 + x_2 \le 2 \\ & x_1 \ge 0 \\ & x_2 \ge 0 \end{aligned}$$

The problem can be expressed as

$$\begin{aligned} \text{Maximize} \quad & f(X) = P^T X - X^T C X. \\ & A^T X \le b. \\ \text{where} \quad & P^T = (3, 2) \end{aligned}$$

$$C = \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix}, \quad A^T = \begin{pmatrix} 5 & 1 \\ 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}$$

and

$$b = \begin{pmatrix} 5 \\ 2 \\ 0 \\ 0 \end{pmatrix}$$

Iteration 1. The initial point is selected which satisfy the first two constraints as equalities.

Thus $\quad X^0 = (A^T_q)^{1}\, b_q$

where $\quad q = 2 = n.$

$$A_q^T = \begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix},\ (A_q^T)^{-1} = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{5}{4} \end{pmatrix},\quad b_q = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$$

Hence $\quad X^0 = \begin{pmatrix} 3/4 \\ 5/4 \end{pmatrix}$

$X^0 \geq 0$ and therefore X^0 is a feasible point.

At $X^0 = \left(\frac{3}{4}\ \frac{5}{4}\right)^T$, we have

$$\nabla f(X^0) = h(X^0) = P - 2 \subset X^0.$$

$$= \begin{bmatrix} 3 \\ 2 \end{bmatrix} - 2\begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix} \begin{bmatrix} 3/4 \\ 5/4 \end{bmatrix} = \begin{bmatrix} 11/4 \\ 1/4 \end{bmatrix}$$

Since $q = n$ $\ P_q\, h(X^0) = 0$ and

$$(A_q^T A_q)^{-1} A_q^T = \begin{pmatrix} 2/16 & 6/16 \\ -6/16 & 26/16 \end{pmatrix}\begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1/4 & -1/4 \\ -1/4 & 5/4 \end{pmatrix}.$$

Then, $\quad U = (A_q^T A_q)^{-1} A_q^T h(X^0)$

$$= \begin{pmatrix} 1/4 & -1/4 \\ -1/4 & 5/4 \end{pmatrix}\begin{pmatrix} 11/4 \\ 1/4 \end{pmatrix} = \begin{pmatrix} 10/16 \\ -6/16 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

since $u_2 < 0$, the second constraint is discarded.

$$\text{Now, } A_q^T = (5 \quad 1), (A_q^T A_q) = 26,\ (A_q^T q_q)^{-1} = \frac{1}{26}$$

$$P_q = I - A_q (A_q^T A_q)^{-1} A_q^T$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{26}\begin{pmatrix} 25 & 5 \\ 5 & .1 \end{pmatrix} = \begin{pmatrix} 1/26 & -5/26 \\ -5/26 & 25/26 \end{pmatrix}$$

$$S^0 = P_q h(X^0) = \begin{pmatrix} 6/4 \times 26 \\ -30/4 \times 26 \end{pmatrix}$$

since $A_3^T S^0 < 0$, we compute

$$\lambda' = \frac{0-(0,-1)\begin{pmatrix}3/4\\5/4\end{pmatrix}}{(0-1)\begin{pmatrix}6/4\times 26\\-30/4\times 26\end{pmatrix}} = \frac{26}{6} = \frac{13}{3}$$

Hence $\lambda' = \dfrac{13}{3}$.

Therefore, the optimal value of $\lambda = \lambda^0$ for which $X^0 + \lambda^0 S^0$ is feasible is obtained from the solution of

Maximize $f(X^0 + \lambda S^0)$

Subject to $0 \le \lambda \le 13/3$

which is reduced to

Maximize $26\lambda - 31\lambda^2$

Subject to $0 \le \lambda \le 13/3$.

The optimal solution is $\lambda^0 = 13/31$ so that

$$X^1 = X^0 + \lambda^0 S^0 = \begin{pmatrix}3/4\\5/4\end{pmatrix} + \frac{13}{31}\begin{pmatrix}6/4\times 26\\-30/4\times 26\end{pmatrix} = \begin{pmatrix}24/31\\35/31\end{pmatrix}$$

Iteration 2. At the point $X^1 = \begin{pmatrix}24/31\\35/31\end{pmatrix}$,

the first constraint is binding and.

$$\nabla f(X^1) = h(X^1) = \begin{pmatrix}80/31\\16/31\end{pmatrix}$$

$$A_q^T = (5,1).,\quad (A_q^T A_q)^{-1} = \frac{1}{26}$$

Thus again

$$P_q = I - A_q(A_q^T A_q)^{-1} A_q^T = \begin{pmatrix}1/26 & -5/26\\-5/26 & 25/26\end{pmatrix}$$

We then have

$$S^0 = P_q\, h(X) = \begin{pmatrix}0\\0\end{pmatrix}$$

$$u_1 = (A_q^T A_q)^{-1} A_q^T h(X^1).$$

$$= \frac{1}{26}(5 \quad 1)\begin{pmatrix} 80/31 \\ 16/31 \end{pmatrix} = \frac{16}{31} > 0$$

Hence the point $X^1 = \left(\frac{24}{31}, \frac{35}{31}\right)^T$ is optimal for the problem.

25.5. Wolfe's Reduced Gradient Method

In this section we describe the reduced gradient method to solve a nonlinear programming problem with linear constraints, first developed by Wolfe [527]. From a computional viewpoint, the reduced gradient method is closely related to the simplex method for linear programming. It consists in partitioning the feasible solution of the problem into sets of dependent (basic) and independent (nonbasic) variables and to consider the problem only in terms of the independent variables, thus reducing the dimensionality of the problem. Like the gradient projection method, the method then proceeds to generate usable feasible directions and finally finds a feasible solution which is a Kuhn–Tueker point, i.e. it satisfies the Kuhn–Tucker conditions.

Consider the problem

$$\begin{aligned} \text{Maximize} \quad & f(X). \\ \text{Subject to} \quad & AX = b \\ & X \geq 0 \end{aligned} \tag{25.92}$$

where $X \in R^n$, A is an m × n matrix of rank m, b is an m-vector and f is a continuously differentiable function on R^n.

We make the nondegeneracy assumption that every set of m columns of A are linearly independent and every basic solution to the constraints has m strictly positive variables. With this assumption, any feasible solution will have atmost n – m zero variables.

Let X be a feasible solution. We partition A into [B, N] and X^T into $\left[X_B^T, X_N^T\right]$, where B is an m × m nonsingular matrix and $X_B > 0$, X_N are basic and nonbasic vectors respectively. The components of X_N may either be positive or zero.

We then have

$$X_B = B^{-1} b - B^{-1} N X_N.$$

The objective function f can then be expressed as a function of the independent variables only and the gradient with respect to the independent variables (the reduced gradient) is found by evaluating the gradient of $f(B^{-1} b - B^{-1} N X_N, X_N)$.

It is given by

$$r^T_N = \nabla_N f(X)^T - \nabla_B f(X)^T B^{-1} N \tag{25.93}$$

where $\nabla_B f(X)$ and $\nabla_N f(X)$ are the gradients of f with respect to X_B and X_N respectively.

Thus $\quad f(X)^T = [0^T, r^T_N]$ (25.94)

Now, we define a direction vector S so that it is a usable feasible direction of f at X, i.e. S is such that AS = 0, and $s_j \geq 0$ if $x_j = 0$ and $\nabla f(X)^T S > 0$.

Let the direction vector S be decomposed into $S^T = \left[S_B^T, S_N^T\right]$ and is defined as follows.

$$s_{Nj} = \begin{cases} 0 & \text{if } r_{Nj} \leq 0 \quad \text{and } x_{Nj} = 0 \\ r_{Nj}, & \text{otherwise} \end{cases} \tag{25.95}$$

$$\text{and} \quad S_B = -B^{-1} N S_N \tag{25.96}$$

The following theorem shows that the direction vector S defined above is a usable feasible direction if $S \neq 0$.

Theorem 25.15

Consider the problem (25.92) where the m × n matrix A is partitioned into A = [B, N] and B is an m × m nonsingular matrix. Let X be a feasible solution to the problem such that $X = [X_B, X_N]$ and $X_B > 0$. Let $S^T = [S_B{}^T, S_N{}^T]$ be the direction vector given by (25.95), (25.96). Then S is a usable feasible direction at X if $S \neq 0$. Further, S = 0 if and only if X is a Kuhn–Tucker point.

Proof: Note that S is a feasible direction if and only if AS = 0 and $S_j \geq 0$ if $x_j = 0$, for j = 1,2,..n.

Now, by definition of S,

$$AS = BS_B + NS_N = B(-B^{-1}NS_N) + NS_N = 0 \tag{25.97}$$

If x_j is basic, then by assumptions $x_j > 0$ and if x_j is nonbasic then by (25.95), s_j could be negative only if and if the nonbasic $x_j = 0$, $s_j \geq 0$. Hence S is a feasible direction.

Further,

$$\nabla f(X)^T S = [\nabla_B f(X)^T, \nabla_N f(X)^T] [S^T_B, S^T_N] = r^T_N S_N \tag{25.98}$$

It is obvious from (25.95) that $\nabla f(X)^T \geq 0$ and equality holds if and only if $S_N = 0$. In that case $S_B = 0$ and therefore S = 0. Hence, if $S \neq 0$, $\nabla f(X)^T > 0$ and then S is a usable feasible direction.

Now, the feasible point X is a Kuhn–Tucker point if and only if there exists a vector $V \in R^m$ such that

$$\nabla_B f(X) + B^T V \leq 0, \tag{25.99}$$

$$\nabla_N f(X) + N^T V \leq 0, \tag{25.100}$$

$$[\nabla_B f(X) + B^T V]^T X_B = 0, \tag{25.101}$$

$$[\nabla_N f(X) + N^T V]^T X_N = 0, \tag{25.102}$$

Since $X_B > 0$, from (25.101) we have, $\nabla_B f(X) + B^T V = 0$ and hence

$$V = -(B^{-1})^T \nabla_B f(X).$$

Substituting the value of V in (25.100), we get

$$\nabla_N f(X) - (B^{-1} N)^T \nabla_B f(X) \leq 0$$

or, $$r_N \leq 0 \qquad (25.103)$$

and from (25.102), we have

$$r^T_N X_N = 0 \qquad (25.104)$$

By definition of S, we note that the above two conditions (25.103) and (25.104) hold if and only if $S_N = 0$ and hence by (25.96), $S_B = 0$. Thus, X is a Kuhn–Tucker point if and only if S = 0.

Note that if the objective function f is concave, the solution X is global optimal.

25.5.1. Determination of the Step Length

Having found a usable feasible direction S at X we want to improve the solution X by continuing moving in the direction S until either the objective function stops improving or a boundary of the feasible region is encountered (i.e. a variable is driven to zero). To determine the maximum possible step length, λ_{max}, we compute

$$\lambda_1 = \min\left(-\frac{x_{Nj}}{s_{Nj}} : \; s_{Nj} < 0 \right)$$

$$\lambda_2 = \min\left(-\frac{x_{Bj}}{s_{Bj}} : \; s_{Bj} < 0 \right)$$

Then and

$$\lambda_{max} = \begin{cases} \min(\lambda_1, \lambda_2) \text{ if } S \ngeq 0 \\ \infty, \text{ if } S \geq 0 \end{cases} \qquad (25.105)$$

The desired step length λ must therefore be less than or equal to λ_{max}. This is obtained by solving the problem

$$\begin{aligned} &\text{Maximize} && f(X + \lambda S) \\ &\text{Subject to} && 0 \leq \lambda \leq \lambda_{max} \end{aligned} \qquad (25.106)$$

This is an one-dimensional problem and can easily be solved. If however, $\lambda_{max} = \infty$ and no finite value of λ is optimal for (25.106), the solution to the original problem is unbounded.

25.5.2. Summary of the Algorithm

We now summarize Wolfe's reduced gradient algorithm for solving the problem (25.92)

Step 1. Find a feasible point X^1. If such a point is not immediately available, the Phase 1 approach of the simplex algorithm may be used.

Suppose that the feasible points $X^1, X^2, \ldots X^k$ have already been produced by the algorithm.

Step 2. At the feasible point X^k with the basis matrix B^k, find S^k from (25.95) and (25.96). If $S^k = 0$, the process terminates; X^k is a Kuhn–Tucker point. Otherwise, proceed to step 3.

Step 3. Obtain the step length λ^k by solving the problem (25.106)

(a) If $\lambda^k = \infty$, the solution of the original problem is unbounded and the process terminates.

(b) If λ^k is finite, $X^{k+1} = X^k + \lambda^k S^k$ is the next feasible point. If $\lambda^k < \lambda_2^k$ [see (25.105)], we do not change the basis. Otherwise, for some index j, $x_{Bj}^k + \lambda^k s_j^k = 0$ and x_{Bj}^k is dropped from the set of basic variables in exchange of the largest positive nonbasic variable. Thus form the new basis matrix B^{k+1} and repeat the step 2.

25.5.3. Example

Consider the problem

$$\begin{aligned} \text{Maximize} \quad & f(X) = 18x_2 - x_1^2 + x_1x_2 - x_2^2 \\ \text{Subject to} \quad & x_1 + x_2 + x_3 = 12 \\ & x_1 - x_2 + x_4 = 6 \\ & x_j \geq 0, j = 1, 2, 3, 4. \end{aligned}$$

Iteration 1

Let us take

$$X_B = (x_3, x_4)^T, X_N = (x_1, x_2)^T \text{ so that}$$

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = B^{-1}, \quad N = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

and the initial feasible point

$$X^1 = (0, 0, 12, 6)^T$$

Note that

$$\nabla f(X) = (-2x_1 + x_2, 18 + x_1 - 2x_2, 0, 0)^T$$

and thus
$$\nabla f(X^1) = (0, 18, 0, 0)^T.$$

From (25.93), the reduced gradient is obtained as

$$r_N^1 = \begin{pmatrix} 0 \\ 18 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 18 \end{pmatrix}$$

The above information is recorded in the tableau below

x_B	x_1	x_2	x_3	x_4
x_3	1	1	1	0
x_4	1	−1	0	1
Solution X^1	0	0	12	6
$\nabla f(X^1)$	0	18	0	0
r_N^1	0	18	–	–

By (25.95), (25.96), we then have

$$S_N^1 = \begin{pmatrix} 0 \\ 18 \end{pmatrix} \quad \text{and}$$

$$S_B^1 = \begin{pmatrix} -18 \\ 18 \end{pmatrix}$$

Hence $\quad S^1 = (0, 18, -18, 18)^T$

Note that x_3, x_4 are the basic variables and x_1, x_2 are nonbasic.

We now compute the step length λ^1 along the direction S^1. By (25.105) we have $\lambda_{max} = 2/3 = \lambda_2^1$ and λ^1 is determined from the optimal solution of the problem

Maximize $\quad f(X^1 + \lambda S^1) = (18)^2 \lambda - (18)^2 \lambda^2$

Subject to $\quad 0 \le \lambda\ 2/3$

$\lambda_{opt} = 1/2.$

Hence $\quad \lambda^1 = 1/2$ and

$x^2 = (X^1 + \lambda^1 S^1) = (0, 9, 3, 15)^T.$

Iteration 2

Since $\lambda^1 < \lambda_2^1 = \dfrac{2}{3}$, we do not change the basic and

at $\quad X^2,\ B = (a_3, a_4) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ N = (a_1, a_2) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

and $\quad \nabla f(X^2) = (9, 0, 0, 0)^T$

We then have

$$r_N^2 = \begin{pmatrix} 9 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 9 \\ 0 \end{pmatrix}$$

The new tableau is now as follows:

x_B	x_1	x_2	x_3	x_4
x_3	1	1	1	0
x_4	1	−1	0	1
Solution X^2	0	9	3	15
$\nabla f(X^2)$	9	0	0	0
r_N^2	9	0	–	–

By (25.95) we have

$$S_N^2 = r_N = \begin{pmatrix} 9 \\ 0 \end{pmatrix}$$

Then $$S_N^2 = -B^{-1}NS_N = -\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 9 \\ 0 \end{pmatrix} = \begin{pmatrix} -9 \\ -9 \end{pmatrix}$$

Hence $S^2 = (9, 0, -9, -9)^T$

Now, to compute step length λ^2, we obtain $\lambda_{max} = 1/3$ and then solve the problem

Maximize $f(X^2 + \lambda S^2) = 81 + 81\lambda - 81\lambda^2$

Subject to $0 \le \lambda \le 1/3$.

We find $\lambda_{opt} = 1/3$.

Hence $\lambda^2 = 1/3$ and $X^3 = 3, 9, 0, 12)^T$

Iteration 3. Since $\lambda^2 = \lambda_2^2 = \frac{1}{3}$, the variable x_2 enters the basis replacing x_3.

Atx^3 = (3, 9, 0, 12)^r, we then have

$B = (a_2, a_4)$, $N = (a_1, a_3)$ and $\nabla f(x^3) = (3, 3, 0, 0)^T$

The new tableau obtained by pivot operation is

x_B	x_1	x_2	x_3	x_4
x_2	1	1	1	0
x_4	2	0	1	1
Solution X^3	3	9	0	12
$\nabla f(X^3)$	3	3	0	0
r_N^3	0	–	−3	–

where r_N^3 is computed by (25.93),

$$r_N^3 = \begin{pmatrix} 3 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -3 \end{pmatrix}$$

By (25.95) we have $S_N^3 = 0$

and hence $S_B^3 = -B^{-1}NS_N = 0$

Thus $S^3 = (S_B\ S_N)^T = (0, 0)^T$. Hence the algorithm terminates and the solution $X^3 = (3, 9, 0, 12)^T$ is optimal to the problem.

We have discussed above the reduced gradient method for optimizing a nonlinear function under linear constraints as given by Wolfe in 1963 [527]. The method, however, does not necessarily converge to a Kuhn–Tucker point. In 1969, McCormick [337] modified Wolfe's method (see Exercise 14), which does converge to a Kuhn–Tucker point. A generalization of the method to nonlinear constraints was given by Abadie and Carpentier [4].

25.6. Zangwill's Convex–Simplex Method

The convex–simplex method of Zangwill is quite similar to the reduced gradient method of Wolfe. The major difference between this method and the reduced gradient method is that instead of permitting all of the nonbasic variables to change at each iteration, only one nonbasic variable is changed while all other nonbasic variables are fixed at their current levels. The selection of the one-nonbasic variable to change is made much as in the ordinary simplex method. Thus, the method behaves very much like the simplex method for linear programs. The method was originally posed by Zangwill [545] for the problem of minimizing a convex function subject to linear inequality constraints. Because of this and its simplex method nature, it was termed the convex–simplex method.

Consider the problem

$$\begin{aligned} &\text{Maximize} && f(X) \\ &\text{Subject to} && AX = b \\ &&& x \geq 0 \end{aligned} \qquad (2.5.107)$$

Where $X \in R^n$, A is an m × n matrix of rank m, b is an m vector and f is a continuously differentiate function on R^n.

It is assumed that every set of m cohimns of A are linearly independent and every basic solution to the constraints has m strictly positive variables.

Suppose that the feasible point X^T is partitioned into $X^T = (X_B^T, X_N^T)$ and A into [B, N], where B is an m × m nonsingular matrix and $X_B > 0$, X_N are basic and non basic vectors respectively. The components of X_N may either be positive or zero.

By using the reduced gradient of f, the convex–simplex method develops the criteria to determine if the current feasible point is optimal and if not to determine the next feasible point.

25.6.1. Summary of the Convex–Simplex Method

We now give the summary of the convex–simplex method for solving the problem. (25.107)

Step 1. Find a basic feasible solution X^1, for which Phase 1 procedure of the simplex algorithm may be used.

Suppose that the feasible points $X^1, X^2, \ldots X^k$ have already been produced by the algorithm.

Step 2. A t x^k, identify $B^k = (a_j, j \in L^k)$,

$N^k = (a_j, j \notin L^k)$, where L^k = index set of m largest components of X^k and compute the reduced gradient of f.

$$r^{kT} = \nabla_N f(X^k)^T - \nabla_B f(X^k)^T (B^k)^{-1} N \tag{25.108}$$

Let

$$r_p^k = \max\left\{r_j^k j \in J\right\},$$

$$r_q^k x_q^k = \max\left\{r_j^k x_j^k j \in J\right\}$$

where J is the set of indices of the nonbasic variables.

If $r_p^k = r_q^k x_q^k = 0$, terminate x^k is a Kuhn–Tucker point.

Step 3. Determine the nonbasic variable to change:

If $r_p^k \geq \left|r_q^k x_q^k\right|$, increase x_p, and

$$\text{If } r_p^k \leq \left|r_q^k x_q^k\right|, \text{decrease } x_q \tag{25.109}$$

Step 4. Calculate the next feasible point x^{K+1}:

There are three cases to consider.

Case (a) x_p is to be increased and at least one component of $Y_p^k = (B^k)^{-1} a_p$ is positive, where a_p is the pth column of A.

Compute improving feasible directions S^K, whose components are given by

$$S_p^k = 1$$

$$S_j^k = 0, \text{ if } j \in J, j \neq p.$$

$$S_j^k = y_{ip}^k, \text{, if } j \notin J, i \in p. \tag{25.110}$$

where I represent the set of indices of basic variables and J is the set of indices of nonbasic variables.

Increasing x_p will drive a basic variable to zero and to maintain feasibility x_p cannot be increased by more than λ_{max} where

$$\lambda_{max} = \frac{x_{Bl}^k}{y_{lp}^k} = \min_i \left\{ \frac{x_{Bi}^k}{y_{ik}^k}, y_{ip}^k > 0 \right\} \tag{25.111}$$

The desired step length $\lambda = \lambda^k$ must therefore be less than or equal to λ_{max}. This is obtained be solving the one-dimensional problem

$$\begin{aligned} &\text{Maximize} && f\left(x^k + \lambda S^k\right) && (25.112)\\ &\text{Subject to} && 0 \le \lambda \le \lambda_{max} \end{aligned}$$

Obtain the new feasible point X^{k+1} by

$$X^{k+1} = X^k + \lambda^k S^k$$

Case (b) x_p is to be increased and all components of y_p are nonpositive. In this case x_p may be increased indefinitely without driving a basic variable to zero.

Determine the step length $\lambda = \lambda^k$ by solving the problem

$$\begin{aligned} &\text{Maximize} && f(X^k + \lambda S^k)\\ &\text{Subject to} && \lambda \ge 0 && (25.113) \end{aligned}$$

where the feasible direction S^k is as given in (25.110) If $\lambda_{opt} = \lambda^k$ is finite, the next feasible point is obtained and proceed to the next iteration. If no finite value of λ optimizes the problem (25,113), the method terminates and the solution is unbounded (However, the optimal value of the objective function is not necessarily infinite)

Case (c) x_q is to be decreased. This implies that $x_q > 0$ in the current solution.

The maximum permitted decrease in x_q is computed as follows.

$$\lambda_1^k = \frac{x_{Bl}^k}{y_{lq}^k} = \max_i \left\{ \frac{x_{Bi}^k}{y_{iq}^k} \middle| y_{iq}^k < 0 \right\} \tag{25.114}$$

$$\lambda_2^k = x_q^k$$

and $\quad \lambda_{max} = \min\left(-\lambda_1^k, \lambda_2^k\right)$

If $y_{iq}^k \ge 0$, for all $i \in I$, take $\lambda_1^k = -\infty$. Then either a basic variable becomes zero or x_q itself becomes zero before a basic variable reaches zero.

The desired amount of decrease in x_q is given by the optimal solution $\lambda_{opt} = \lambda^k$ of the problem.

$$\begin{aligned} &\text{Maximize} && f(x^k - \lambda s^k)\\ &\text{Subject to} && 0 \le \lambda \le \lambda_{max} && (25.115) \end{aligned}$$

Where s^k is as given in (25.110)

Then the new feasible point is $X^{k+1} = X^k - \lambda^k S^k$.

Step 5. After calculating X^{K+1}, go to the next iteration and repeat from step 2 replacing k by k + 1, until the algorithm converges to a Kuhn–Tucker point.

For the proof of convergence of the method and other details see Zangwill [545, 550].

25.6.2. Example

Consider the problem

$$\begin{aligned}\text{Maximize} \quad & f(X) = 18x_2 - x_1^2 + x_1x_2 - x_2^2 \\ \text{Subject to} \quad & x_1 + x_2 + x_3 = 12 \\ & x_1 - x_2 + x_4 = 6 \\ & x_j \geq 0, \ j = 1, 2, 3, 4.\end{aligned}$$

Interation 1. Let us take $X_B = (x_3, x_4)^T$, $X_N = (x_1, x_2)^T$. We then have

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad N = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and the initial feasible point $X^1 = (0, 0, 12, 6)^T$

Note that $\nabla f(X) = (-2x_1 + x_2, 18 + x_1 - 2x_2, 0, 0)^T$

and thus $\nabla f(X^1) = (0, 18, 0, 0)^T$

By (25.108), the reduced gradient is obtained as

$$r^1 = \begin{pmatrix} 0 \\ 18 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 18 \end{pmatrix}$$

The above information is recorded in the following tableau.

X_B	x_1	x_2	x_3	x_4
x_3	1	1	1	0
x_4	1	–1	0	1
Solution X^1	0	0	12	6
$\nabla f(X^1)$	0	18	0	0
r^1	0	18	–	–

By (25.108), we get

$$\max\{r_j^1, j \in J\} = r_2^1 = 18$$

$$\min\{r_j^1 x_j^1, j \in J\} = 0$$

Since $r_2^1 = 18 > 0$, we note from (25.109) that the nonbasic variable x_2 is to be increased and since $Y_2 = (1, -1)^T$ has a positive component, we are in case (a).

By (25.110) then, we have the usable variable direction $S^1 = (0, 1, -1, 1)^T$. The maximum value of λ such that $X^1 + \lambda S^1$ is feasible is computed by (25.111) and we get

$$\lambda_{max} = \min_i \left(\frac{x_{Bi}}{y_{i2}}, y_{i2} > 0 \right) = \frac{x_{B1}}{y_{12}} = \frac{12}{1} = 12$$

The desired step length λ^1 at x^1 is then obtained by solving the problem.

Maximize $\quad f(x^1 + \lambda s^1) = 18\lambda - \lambda^2$

Subject to $\quad 0 \le \lambda \le 12$

Now, $\frac{df}{d\lambda} = 0 \Rightarrow \lambda = 9$ which is less than 12 and therefore, the optional solution $\lambda_{opt} = 9$

Thus $\quad \lambda^1 = 9$

and $\quad X^2 = X^1 + \lambda^1 S^1 = (0, 0, 12, 6)^T + 9\,(0, 1, -1, 1)^T$

$\quad = (0, 9, 3, 15)^T$

Iteration 2. At $X^2 = (0, 9, 3, 15)^T$, $L^2 = \{2, 4\}$ so that $B = (a_2, a_4)$ and $N = (a_1, a_3)$. The new tableau obtained by pivot operation is given below.

X_B	x_1	x_2	x_3	x_4
x_2	1	1	1	0
x_4	2	0	1	1
Solution X^2	0	9	3	15
$\nabla f(X^2)$	9	0	0	0
r^2	9	–	0	–

where r^2 is computed by (25.108)

$$r^2 = \begin{pmatrix} 9 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 9 \\ 0 \end{pmatrix}$$

Since $\quad \max\left(r_1^2, r_3^2\right) = r_1^2 = 9$

$$\min\left(r_1^2 x_1^2, r_3^2 x_3^2\right) = 0$$

we note from (25.109) that x_1 has to be increased.

From (25.110) we get $s^2 = (1, 0, -1, -1)^T$. The maximum value of λ so that $X^2 + \lambda S^2$ is feasible is then computed by (25.111).

$$\lambda_{max} = \min\left(\frac{3}{1}, \frac{15}{2}\right) = 3$$

The desired step length λ^2 at X^2 is then obtained by solving the problem.

$$\begin{array}{ll} \text{Maximize} & f(X^2 + \lambda S^2) = 81 + 9\lambda - \lambda^2 \\ \text{Subject to} & 0 \le \lambda \le 3. \end{array}$$

The optimal solution $\lambda_{opt} = 3$.

Thus, $\lambda^2 = 3$

and $X^3 = (X^2 + \lambda S^2) = (3, 9, 0, 12)^T$

Iteration 3. At $X^3 = (3, 9, 0, 12)^T$, $L^3 = \{2, 4\}$ so that $B = (a_2, a_4)$ and $N = (a_1, a_3)$. Thus the basis remains the same and we have

x_B	x_1	x_2	x_3	x_4
x_2	1	1	1	0
x_4	2	0	1	1
Solution X^3	3	9	0	12
$\nabla f(X^3)$	3	3	0	0
r^3	0	–	–3	–

where r^3 is computed by (25.108)

$$r^3 = \begin{pmatrix} 3 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -3 \end{pmatrix}$$

We then have

$$\max\left(r_1^3, r_3^3\right) = \max(0, -3) = 0$$

$$\min\left(r_1^3 x_1^3, r_3^3 x_3^3\right) = \min(0, 0) = 0$$

Hence $X^3 = (3, 9, 0, 12)^T$ is optimal.

25.7. Dantzig's Method for Convex Programs

The methods that we have discussed so far for solving convex programming problems assumed that the functions are differentiable. In this section we shall discuss a method developed by Dantzig [109] for solving a general convex programming problem even if the functions are not differentiable provided some mild regularity conditions are satisfied. The burden of the work in this iterative procedure shifts to a subproblem, which must be solved afresh at each iteration. This itself again is convex programming problem, which may or may not be easy to solve for general convex functions.

Consider a general convex programming problem

$$\begin{aligned} &\text{Minimize} \quad z = \Phi_0(X) \\ &\text{Subject to} \quad \Phi_i(X) = 0,\ i = 1,2,\ldots\ldots r \\ &\qquad\qquad\quad \Phi_i(X) \le 0,\ i = r+1,\ldots\ldots m \end{aligned} \tag{25.116}$$

Where $\Phi_i(X)$, $i = 1,2,\ldots\ldots\ldots r$ are linear and

$\Phi_i(X)$, $i = 0, r+1,\ldots m$ are general

convex function with the assumptions that

(a) the domain of variation is restricted to a closed bounded convex set R.
(b) the convex functions are continuous in R. (i.e. continuity extends to the boundary).
(c) there exists a nondegenerate basic feasible solution X^0 of $\Phi_i(X) = 0$, $i = 1,2,\ldots\ldots r$ such that $\Phi_i(X^0) < 0$, $i = r+1,\ldots\ldots\ldots m$

The problem (25.116) may be rewritten in the form of a generalized linear program [109, 120]

$$\begin{aligned} \lambda_0 &= 1 \\ y_1\lambda_0 &= 0 \\ &\vdots \\ y_r\lambda_0 &= 0 \\ y_{r+1}\lambda_0 + \mu_{r+1} &= 0 \\ y_m\lambda_0 + \mu_m &= 0 \\ y_0\lambda_0 &= z\ (\text{Min}),\ \mu_i \ge 0,\ i = r+1,\ldots m. \end{aligned} \tag{25.117}$$

where y_i are variable coefficients which may be freely chosen subject to the conditions that

$$y_i = \Phi_i(X),\ i = 1, 2, \ldots\ldots r \text{ and } y_i \ge \Phi_i(X),\ i = r+1\ldots m$$

for some $X \in R$

we then consider the following problem called a "restricted master problem"

$$\begin{aligned} &\lambda_1 P_1 + \lambda_2 P_2 + \ldots + \lambda_p P_p + \mu_{r+1} u_{r+1} + \ldots \mu_m u_m + (-z)\, U_{m+1} = U_0 \\ &\lambda_j \ge 0,\ j = 1,2,\ldots p;\ \mu_i \ge 0,\ i = r+1,\ldots m \end{aligned} \tag{25.118}$$

where Ui ($i = 0, r+1,\ldots m+1$) denotes the unit vector with ith component being unity and vectors

$$P_j^T = (1, y^j_1, y^j_2 \ldots y^j_m, y^j o),\ j = 1,2,\ldots p \tag{25.119}$$

are a set of admissible

$$P^T = (1, y_1, y_2 \ldots y_m, yo) \tag{25.120}$$

Our assumption (c) implies that there exists a nondegenerate basic feasible solution to a restricted master problem which can be used to initiate the algorithm. We assume that this has already been done. The restricted master problem for iteration k is shown explicitly in detached coefficient form in (25.121)

Restricted Master Problem for Iteration k							
$\lambda_1 \geq 0$	$\lambda_2 \geq 0 \ldots$	$\lambda_{p+k} \geq 0$	$\mu_{r+1} \geq 0 \ldots$	$\mu_m \geq 0$	$-z$	Constants	Multipliers
1	1 ...	1	0	0	0	1	Π_0
y_1^1	$y_1^2 \ldots$	y_1^{p+k}	0	0	0	0	Π_1
y_2^1	$y_2^2 \ldots$	y_2^{p+k}	0	0	0	0	Π_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_r^1	$y_r^2 \ldots$	y_r^{p+k}	0	0	0	0	$\vdots$
y_{r+1}^1	$y_{r+1}^2 \ldots$	y_{r+1}^{p+k}	1	0	0	0	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_m^1	$y_m^2 \ldots$	y_m^{p+k}	0	1	0	0	Π_m
y_0^1	$y_0^2 \ldots$	y_0^{p+k}	0	0	1	0	1

(25.121)

Iterative Procedure: For iteration k, the restricted master problem (25.118) is optimized obtaining a new basic feasible solution.

$$\lambda^k = \left(\lambda_1^k, \lambda_2^k, \ldots \lambda_{p+k}^k, \mu_{r+1}^k, \ldots \mu_m^k\right) \geq 0 \text{ with } \quad z^k = \sum_{j=1}^{p+k} \lambda^k y_0^j \tag{25.122}$$

and a new set of simplex multipliers

$$\Pi^k = \left(\Pi_0^k, \Pi_1^k, \ldots \Pi_m^k 1\right) \tag{25.123}$$

To express conveniently, the kth approximate solution to the problem (25.116) we now assume that the vector P_j of the master problem is defined by

$$P_j^T = (1, \Phi_1(X^j), \Phi_2(X^j), \ldots \Phi_m(X^j), \Phi_0(X^j)), j = 1,2,\ldots p+k$$

where $y_i^j = \Phi_i(X^j)$, $i = 0, 1, 2, \ldots m$ for some $X = X^j$, chosen from the closed bounded convex set (25.124)

The kth approximate solution to (25.116) is then given by,

$$\hat{X}^k = \sum_{j=1}^{p+k} \lambda_j^k x^j, \ \hat{z}^k = \phi_0\left(\hat{X}^k\right) \tag{25.125}$$

It should be noted that the simplex multipliers

π^k satisfy the following conditions:

$\pi^k P_j = 0$; $\pi^k U_i = 0$, if λ_j or μ_i is a basic variable

$\pi^k P_j \geq 0$; $\pi^k U_i \geq 0$, if π_j or μ_i is a nonbasic variable

and $\pi^k U_{m+1} = 1$ (25.126)

It then follows thth, $\pi_i^k \geq 0$, for I = r +1, ...m (25.127)

Further note from (25.121) and (25.126) that $\pi_0^k = -z^k$ (25.128)

To test whether or not the kth approximate solution (25.125) is optimal, the function

$$\Delta(X \mid \pi^k) = \phi_0(X) + \sum_{i=1}^{m} \pi_i^k \phi_i(X) + \pi_0^k \tag{25.129}$$

is minimized over all X ε R

Theorem 25.16

$$\text{If } \underset{X \in R}{\text{Min}} \left\{ \Delta(X \mid \pi^k) = \phi_0(X) + \sum_{i=1}^{m} \pi_i^k \phi_i(X) + \pi_0^k \right\} \geq 0$$

then the kth approximate solution $\hat{X}^k$ is optimal for the problem (25.116)

Proof: Let $X \in R$ be feasible for (25.116) and suppose

that $\underset{X \in R}{\text{Min}}\, \Delta(X \mid \pi^k) \geq 0$

Since $\pi_i^k \geq 0$, for $i = r+1, \ldots m$, $\sum_{i=r+1}^{m} \pi_i^k \phi_i(X) \leq 0$

and hence

$$\Delta(X|\pi^k) \leq \phi_0(X) + \pi_0^k \tag{25.130}$$

and since $\pi_0^k = -z^k$, we have

$$\Delta(X|\pi_0^k) + z^k \leq \phi_0(X) \tag{25.131}$$

Now, from the convexity of $\phi_0(X)$, we have

$$\phi_0(\hat{X}^k) = \phi_0\left(\sum_{j=1}^{p+k} \lambda_j^k X^j\right) \leq \sum_{j=1}^{p+k} \lambda_j^k \phi_0(X^j)$$

$$= \sum_{j=1}^{p+k} \lambda_j^k y_0^j = Z^k \tag{25.132}$$

From (25.130) we get

$$\text{Min } \Delta(X|\pi^k) \leq \Delta(\hat{X}|\pi^k) \leq \phi_0(\hat{X}^k) + \pi_0^k$$

$$= \phi_0(\hat{X}^k) - z^k$$

$$\leq 0, \text{ by } (25.132)$$

and since by the hypothesis

Min $\Delta(X|\pi^k) \geq 0$, it follows that

$\text{Min } \Delta(X|\pi^k) = 0.$

Now, from (25.131) we have,

$$\text{Min } \Delta(X|\pi^k) + Z^k \leq \text{Min } \phi_0(X) \leq \phi_0(\hat{X}^k) \leq Z^k, \text{ by (25.132)} \tag{25.134}$$

From (25.133) and (25.134) it then follows

That

$$-\pi^k_0 = Z^k = \phi_0(\hat{X}^k) = \text{Min } \phi_0(X)$$

which establishes the theorem

If however, $\underset{X \in R}{\text{Min}}\ \Delta(X|\pi^k) < 0$, a new vector P_{p+k+1} is generated for the restricted master problem for the (k + 1)th iteration.

We define $X^{k+1} \varepsilon R$ and the new vector P_{p+k+1} by

$$\Delta(X^{k+1}|\pi^k) = \underset{X \in R}{\text{Min}}\ \Delta(X|\pi^k) < 0 \tag{25.135}$$

$$P^T_{p+k+1} = (1, \phi, (X^{k+1}), \ldots \phi_m(X^{k+1}), \phi_0(X^{k+1})) \tag{25.136}$$

and the process is repeated

It can be shown that $\underset{X \in R}{\text{Min}}\ \Delta(X|\pi^k)$ tends to zero on some subsequence of the k's, so that $\phi_0(X^k)$ converges to the desired minimum and finally we have

Theorem 25.17

$$\lim_{k \to \infty} \phi_0(\hat{X}^k) = \text{Min } \phi_0(X)$$

for x satisfying the conditions of the problem (25.116)

Proof : Dantzig [107, 109]

Thus in Dantzig's method for convex programming, we obtain at each iteration an approximate solution by the simplex method, which is then checked whether optimal or not by solving a problem of the type (25.129), Theorem 25.17 then states that the process, though it may be infinite converges to a solution.

25.8 Exercises

1. Convert the following problems into separable programs. Replace the non-linear functions by their piecewise linear approximations and then solve the approximating problems.

   (i) Minimize $2x_1^2 - x_1x_2 + x_2^2 - 6x_1 + 2x_2$

   Subject to $2x_1 + x_2 \leq 12$

   $x_1^2 + x_2^2 = 8$

   $x_1, x_2 \geq 0$

(ii) Minimize $3x_1 + 2x_2$
Subject to $x_1x_2 \geq 16$
$x_2 \geq 4$

(iii) Minimize $x_1 - 2x_2$
Subject to $x_1x_2 \geq 2$
$x_1 + x_2 \leq 4$
$x_1, x_2 \geq 0$

2. Solve the following problems by the separable programming method. Can the restricted basis entry rule be dropped for these problems?

(i) Maximize $2x_1 + 3x_2$
Subject to $x_1^2 + 4x_2^2 \leq 16$
$x_1, x_2 \geq 0$

(ii) Minimize $(x_1 - 3)^2 + (x_2 - 3)^2$
Subject to $x_1 + x_2 \leq 4$
$x_1, x_2 \geq 0$

(iii) Minimize $x_1^2 + x_2^2 - 8x_1 - 6x_2 - 1/2x_3$
Subject to $-x_1 + x_2^2 \leq 3$
$x_1 + x_2 + x_3 \leq 5$
$x_1, x_2, x_3 \geq 0$

3. Solve the following problems by Kelley's cutting plane algorithm

(i) Minimize $x_1 - x_2$
Subject to $x_1 + x_2 \leq 5$
$x_1 . x_2 \geq 4$
$x_1, x_2 \geq 0$

(ii) Maximize $x_1 + 3x_2$
Subject to $x_1 + x_2 \leq 3$
$x_1 + x_2^2 \leq -1$
$x_1, x_2 \geq 0$

(iii) Minimize $5x_1 + 4x_2$
Subject to $2x_1^2 + x_2^2 + 2x_1x_2 \leq 4$
$-x_1^2 - x_2^2 + 4x_2 \geq 3$

4. Prove that Kelley's cutting plane algorithm is still globally convergent if it is modified by discarding at each stage all cutting plane constraints that are not binding on the optimal solution to the corresponding linear program.

5. Solve the following problem by Zoytendijk's feasible direction method

(i) Maximize $3x_1 + 2x_2 - x_1^2 + x_1x_2 - x_2^2$

Subject to $x_1 - 2x_2^2 \geq 0$

$$5x_1 + x_2 \le 5$$
$$x_1, x_2 \ge 0$$

(ii) Maximize $8x_1 + 6x_2 - x_1^2 - x_2^2$

Subject to $x_1^2 + 4x_2^2 \le 16$
$$5x_1 + 3x_2 \le 15$$
$$x_1, x_2 \ge 0$$

(iii) Maximize $x_1 + 2x_2 - x_2^2$

Subject to $3x_1^2 + 2x_2^2 \le 6$
$$x_1, x_2 \ge 0$$

6. Solve the following problem:

Minimize $(x_1 - 1)^2 + (x_2 - 2)^2$
$$x_1 - x_2^2 \ge 0$$
$$2x_1 - x_2 = 1$$

by Zoutendijk's method using the normalization constraint

(a) $S^T S \le 1$

(b) $-1 \le s_j \le 1$, for $j = 1,2$.

7. Solve the following nonlinear programs with linear constraints by Zoutendijk's feasible direction method and by Rosen's gradient projection method.

(i) Maximize $x_1 + 2x_2 - x_2^2$

Subject to $x_1 + 2x_2 \le 4$
$$3x_1 + 2x_2 \le 6$$
$$x_1, x_2 \ge 0$$

(ii) Minimize $4x_1^2 + 3x_2^2$

Subject to $x_1 + 3x_2 \ge 5$
$$x_1 + 4x_2 \ge 4$$
$$x_1, x_2 \ge 0$$

8. Solve the following problem by Rosen's gradient projection method

Maximize $2x_1 + 4x_2 - x_1^2 - x_2^2$

Subject to $2x_1 + 3x_2 \le 6$
$$x_1 + 4x_2 \le 5$$
$$x_1, x_2 \ge 0$$

9. Solve the problem of Exercise 8 by Zontendijk's feasible direction method using the normalization constraint $S^T S \le 1$

10. Solve the following problem by the gradient projection method using the origin as the starting point

Maximize $2x_1 + x_2 + x_1^3 x_2$

Subject to $2x_1 + x_2 \le 8$

$x_1 - 2x_2 \le 4$

$x_1 - x_2 \ge 0$

$x_1, x_2 \ge 0$

11. Determine the projection of the gradient of the function

$$F(x) = 5x_1 - 3x_2 + 6x_3$$

onto the x_1x_2- plane. Sketch the gradient and its projection.

12. Show that the gradient projection method will solve a linear program in a finite number of steps.

13. Solve the following problem by (a) the reduced gradient method and (b) the convex simplex method, starting at the origin in each case.

(i) Maximize $10x_1 + 20x_2 - x_1^2 + x_1x_2 - x_2^2$

Subject to $x_1 - x_2 \le 6$

$x_1 + x_2 \le 12$

$x_1, x_2 \ge 0$

(ii) Maximize $10\log_e(x_1+1) + (x_2+1)^2 + 4x_3$

Subject to $x_1 + x_2 + x_3 \le 12$

$3x_2 + x_3 \le 24$

$4x_1 + x_2 + 2x_3 \le 28$

$x_1, x_2, x_3 \ge 0$

14. Consider the problem

Minimize $f(X)$

Subject to $AX = b$

$X \ge 0$

Where A is an $m \times n$ matrix of rank m , f is a continuously differentiable function on R^n and that the nondegeneracy assumption holds.

Suppose that the reduced gradient method discussed in section 25.5 is modified [McCormick] with the modifications that (a) the basic variables are, at the beginning of an iteration, always taken as the m largest variables and (b) the direction vector $S=(S_B,S_N)$ is now defined as

$$S_{Nj} = \begin{cases} -r_{Nj}, \text{ if } r_{Nj} \le 0 \\ -x_{Nj}r_{Nj}, \text{ if } r_{Nj} > 0 \end{cases}$$

$$S_B = -B^{-1} N S_N$$

Using this modification, solve the following problem by the reduced gradient method:

$$\begin{array}{ll} \text{Minimize} & x_1^2 + x_2^2 - x_1x_2 - 3x_1 - 2x_2 \\ \text{Subject to} & 5x_1 + x_2 + x_3 = 5 \\ & x_1 + x_2 + x_4 = 2 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{array}$$

15. Show that the convex–simplex method reduces to the simplex method if the objective function is linear.

Chapter 26

Duality in Nonlinear Programming

26.1. Introduction

The duality theory in nonlinear programming is a natural extension of duality in linear programming. Noting the success of the application of the duality principles in linear programming many authors became interested in generalizing the well known duality properties of linear programs to nonlinear cases. The Lagrangian saddle point problem and the Kuhn–Tucker theory also incited a great deal of interest in duality in nonlinear programming. As a result, a large number of papers on duality in nonlinear programming appeared in the literature.

Consider the problem

$$\begin{aligned} &\text{Minimize} && \phi(X) \\ &\text{Subject to} && g(X) \leq 0 \end{aligned} \tag{26.1}$$

Where ϕ and g are scalar and m-vector functions respectively both defined on an open set $\mathcal{X} \subset R^n$.

It is assumed that both ϕ and g are differentiable on χ.

The problem (26.1) is called the primal problem. From the saddle point problem of the associated Lagrangian function we form the following problem called the dual problem

$$\begin{aligned} &\text{Maximize} && \psi(X,U) = \phi(X) + U^T g(X) \\ &\text{Subject to} && \nabla\phi(X) + U^T \nabla g(X) = 0 \\ &&& U \geq 0 \end{aligned} \tag{26.2}$$

where $X \in \chi$ and $U \in R^m$

Let the constraint set of the primal be denoted by C_p and that of the dual problem by C_D.

26.2. Duality Theorems

Several duality formulations have been evolved by various authors under different conditions which satisfy many of the properties of linear dual programs. In this section, we discuss the relationships between the nonlinear programs (26.1) and (26.2), under the assumptions of convexity of the functions involved.

Theorem 26.1: Weak Duality Theorem [525]

Let $\phi(X)$ and $g(X)$ be diffrentiable convex function of $X \in \mathfrak{X}$. Then

$\text{Inf } \phi(X) \geq \text{Sup } \psi(X,U)$

Proof : Let $X^1 \in C_p$ and $(X^2, U^2) \in C_D$

$$\begin{aligned}
\text{Then}\quad \phi(X^1) - \phi(X^2) &\geq \phi(X^2)^T (X^1 - X^2), && \text{since } \phi(X) \text{ is convex}\\
&= -U^{2T}\, \nabla g(X^2)\, (X^1 - X^2), && \text{since } \nabla\phi(X^2) + U^{2T} g(X^2) = 0\\
&\geq U^{2T}\, [g(X^2) - g(X^1)], && \text{since } g(X) \text{ is convex and } U^2 \geq 0\\
&\geq U^{2T}\, g(X^2), && \text{since } g(X^1) \leq 0 \text{ and } U^2 \geq 0
\end{aligned}$$

thus $\phi(X^1) \geq \phi(X^2) + U'^2 g(X^2) = \psi(X^2, U^2)$

Using the convection that

$$\text{Inf } \phi(X) = +\infty, \quad \text{if } C_p \text{ is empty}$$

And $$\text{Sup } \psi(X, U) = -\infty \quad \text{if } C_D \text{ is empty}$$

we have $\text{Inf } \phi(X) \geq \text{Sup } \psi(X, U)$.

Theorem 26.2: Duality Theorem [525, 238]

Let $\phi(X)$ and $g(X)$ be diffrentiable convex functions of $X \in \mathfrak{X}$. Let X^0 be an optimal solution of the primal problem (26.1) and let $g(X)$ satisfy the Kuhn–Tucker constraint qualification. Then there exists a $U^0 \in R^m$ such that (X^0, U^0) is an optimal solution of the dual problem (26.2) and

$$\phi(X^0) = \psi(X^0, U^0).$$

Proof: Since $g(X)$ satisfies the Kuhn–Tucker constraint qualification, there exists a $U^0 \in R^m$ such that (X^0, U^0) satisfies the Kuhn–Tucker conditions.

$$\begin{aligned}
&\nabla\phi(X^0) + U^{0T}\, \nabla g(X^0) = 0 \qquad (26.3)\\
&U^{0T}\, g(X^0) = 0\\
&g(X^0) \leq 0\\
&U^0 \geq 0
\end{aligned}$$

It is clear that (X^0, U^0) is a feasible solution of the dual problem. Let (X, U) be any arbitary feasible solution of the dual, Then

$$\begin{aligned}
\psi(X^0, U^0) - \psi(X, U) &= \phi(X^0) - \phi(X) + U^{0T} g(X^0) - U^T g(X)\\
&\geq \nabla\phi(X)^T (X^0 - X) - U^T g(X),\\
&\qquad \text{since } \phi(X) \text{ is convex and } U^{0T} g(X^0) = 0\\
&\geq \nabla\phi(X)^T (X^0 - X) - U^T [g(X^0) = \nabla g(X)^T (X^0 - X)]\\
&\qquad \text{since } g(X) \text{ is convex and } U \geq 0\\
&= [\nabla\phi(X)^T + U^T\, \nabla g(X)^T\,]\, (X^0 - X) - U^T g(X^0)\\
&= 0 - U^T g(X^0)\\
&\geq 0, \quad \text{since } U \geq 0 \text{ and } g(X^0) \leq 0
\end{aligned}$$

and thus $\psi(X^0, U^0) \geq \psi(X, U)$

which implies that (X^0, U^0) is an optimal solution of the dual problem

Moreover, since $U^{0T}g(X^0) = 0$

$\psi(X^0, U^0) = \phi(X^0) + U^{0T}g(X^0) = \phi(X^0)$

Theorem 26.2 can alternatively be proved by direct application of weak duality theorem. Since g(X) satisfies the Kuhn–Tucker constraint qualification, there exists $U^0 \in R^m$ such that (X^0, U^0) satisfies the condition (26.3), which implies that (X^0, U^0) is a feasible of the dual problem.

$$
\begin{aligned}
\text{Now,} \quad \psi(X^0, U^0) &= \phi(X^0) + U^{0T}g(X^0) \\
&= \phi(X^0), \qquad \text{since } U^{0T}g(X^0) = 0 \\
&\geq \psi(X, U), \text{ by Theorem 26.1}
\end{aligned}
$$

Hence (X^0, U^0) is an optimal solution of the dual problem and $\phi(X^0) = \psi(X^0, U^0)$.

Theorem 26.3: Converse Duality Theorem [238]

Let $\phi(X)$ and $g(X)$ be diffrentiable convex functions of $\mathfrak{X} \in \chi$ and let (X^0, U^0) be an optimal solution of the dual problem (26.2). If $\psi(X, U^0)$ is twice continuously differentiable at X^0 and the $n \times n$ Hessian matrix $\nabla^2_x\psi(X^0, U^0)$ is nonsingular then X^0 is an optimal solution of the primal problem (26.1) and $\phi(X^0) = \psi(X^0, U^0)$.

Proof : Since by assumption, the Hessian matrix $\nabla^2_x\psi(X^0, U^0)$ is nonsingular, the implict function theorem is applicable to $\nabla_x\psi(X, U)$. Hence there is an open set $W \in R^m$ containing U^0 and an n-dimensional differentiable vector function h(U) defined on W such that

$$X^0 = h(U^0)$$

$$H(U) \in \chi, \text{ for } U \in W$$

$$\nabla\Psi_x[h(U), U] = 0, \text{ for } U \in W \tag{26.4}$$

Since (X^0, U^0) is an optimal solution of the dual problem and $\psi(X^0, U^0) = \psi[h(U^0), U^0]$, U^0 is an optimal solution of the problem

$$
\begin{aligned}
&\text{Maximize} \quad \psi[h(U), U] && (26.5)\\
&\text{Subject to} \quad U \geq 0, \qquad U \in W
\end{aligned}
$$

Since W is open and the constraints of the problem (26.5) are linear, the Kuhn–Tucker constraint qualification is satisfied. The associated Lagrange function is

$$L(U, V) = \psi[h(U), U] + V^TU$$

where $V \in R^m$ is the vector of Lagrange multipliers. Thus, there is some $V^0 \in R^m$ such that (U^0, V^0) satisfy the Kuhn–Tucker conditions

$$
\begin{aligned}
&\nabla_U \psi[h(U^0), U^0] + V^0 = 0 \qquad && \text{(i)} \\
&V^{0T}U^0 = 0 && \text{(ii)} \qquad (26.6)\\
&U^0 \geq 0, V^0 \geq 0
\end{aligned}
$$

By chain rule of diffrentiation applied to $\nabla_U[h(U^0), V^0]$ we get

$$
\begin{aligned}
-V^0 &= \nabla_U \psi[h(U^0), U^0] \\
&= \nabla_u h(U^0)^T \nabla_x\psi(X^0, U^0) + \nabla_U \psi(X^0, U^0),
\end{aligned}
$$

$$= \nabla_u \psi (X^0, U^0), \qquad \text{since } h(U^0) = X^0$$
$$= g(X^0) \qquad \text{since } \psi_x(X^0, U^0) = 0 \tag{26.7}$$

Thus $g(X^0) = -V^0 \leq 0$ and $X^0 \in \mathfrak{X}$, so that $X^0 \in C_p$. Further, we note that $U^0 \geq 0$, $U^{T0}g(X^0) = - U^{0T}V^0 = 0$ and $\nabla_x \psi (X^0, U^0) = 0$. Hence (X^0, U^0) satisfies the Kuhn–Tucker conditions for the problem (26.1) and therefore X^0 is an optional solution of the primal problem.

Moreover,

$$\phi(X^0) = \phi(X^0) = U^{0T}g(X^0),$$
$$= \psi(X^0, U^0). \qquad \text{since } U^{oT}g(X^0) = 0$$

26.3. Special Cases

In this section we shall discuss the duality theory for special cases of nonlinear programs. In particular, we shall discuss duality in linear constrained nonlinear programs and in quadratic programming.

26.3.1. Duality in Nonlinear Programs with Linear Constraints

Consider the problem

$$\begin{aligned} &\text{Minimize} && \phi(X) \\ &\text{Subject to} && AX \leq b \\ &&& X \geq 0 \end{aligned} \tag{26.8}$$

Where $\phi(X)$ is a differentiable convex function of $X \in R^n$.

The problem (26.8) will be called the primal problem. Since the constraints of (26.8) are linear, the Kuhn–Tucker constraint qualification is satisfied and the Kuhn–Tucker conditions are given by

$$\begin{aligned} &\nabla\phi(X) + A^TU \geq 0 \\ &X^T \nabla\phi(X) + U^TAX = 0 \\ &AX \leq b \\ &U^TAX - b^TU = 0 \\ &X \geq 0, U \geq 0 \end{aligned} \tag{26.9}$$

The dual problem to (26.8), formed according to (26.2) is then given by

$$\begin{aligned} &\text{Maximize} && \psi (X, U) = \phi(X) - X^T \phi(X) - b^TU \\ &\text{Subject to} && \nabla\phi(X) + A^TU \geq 0 \\ &&& U \geq 0 \end{aligned} \tag{26.10}$$

Duality Results

The following theorems follow directly from the theorems proved in the previous section .

Theorem 26.4: Weak Duality Theorem:

$$\text{Inf } \phi(X) \geq \text{Snp } \psi (X, U)$$

Proof: Follows the Theorem 26.1.

Theorem 26.5: Duality Theorem

If X^0 is an optimal solution of the primal problem (26.8), then there exists a $U^0 \in R^m$ such that (X^0, U^0) is an optional solution of the dual problem (26.10) and $\phi(X^0) = \psi(X^0, U^0)$.

Proof: Follows from theorem 26.2.

Theorem 26.6: Converse Duality Theorem

If (X^0, U^0) is an optimal solution of the dual program (26.10) and if $\phi(X)$ is twice continuously diffrentiable at X^0 with $\nabla^2_x\phi(X)$ nonsingular, then X^0 is an optimal solution of the problem (26.8) and $\phi(X^0) = \psi\ (X^0, U^0)$

Proof: Follows from Theorem 26.3

The following additional result can also be derived

Theorem 26.7: Unbounded Dual

If the primal problem is infeasible and the dual problem is feasible then

$$\text{Snp } \psi\ (X, U) = +\infty$$

Proof: Since the primal problem is infeasible the system

$$AX \leq b$$
$$X \geq 0$$

has no solution

By the inequality Theorem 7.11 then there exists a $Y \in R^m$ satisfying

$$A^TY \geq 0 \tag{26.11}$$
$$b^TY < 0\ ,\ Y \geq 0.$$

Let (X, U) be a feasible solution of the dual problem (26.10). Then (X, U + ty) is also feasible to the dual problem for all $t \geq 0$ and

$$\psi\ (X, U+tY) = \phi(X) - X^T\phi(X) - b^T(U + tY)$$
$$\rightarrow +\infty \text{ as } t \rightarrow +\infty \text{ since } b^TY < 0$$

26.3.1.1. Duality in Quadratic Programming

The duality theory of quadratic programming was first studied by Dennis [117] and then by Dorn [122]. As in section 26.1, we form the dual to a quadratic programming problem from the saddle point problem associated with the Lagrangian function of the given problem and establish their duality relations.

Consider the quadratic programming problem

$$\begin{aligned} \text{Minimize} \quad & \phi(X) = p^TX + \tfrac{1}{2} X^TCX \\ \text{Subject to} \quad & Ax \leq b \\ & x \geq 0 \end{aligned} \tag{26.12}$$

where $X \in R^n$, A is an m × n matrix, b an m-vector and C is an n × n symmetric positive semidefinite matrix so that $\phi(X)$ is a diffrentiable convex function of X $\in R^n$.

The Kuhn–Tucker conditions associated with the problem is then given by

$$\begin{aligned}&p + CX + A^TU \geq 0\\&p^TX + X^TCX + U^TAX = 0\\&AX = b\\&U^TAX - b^TU = 0\\&X \geq 0,\ U \geq 0\end{aligned} \tag{26.13}$$

According to (26.10), the dual problem to (26.12) is given by

$$\begin{aligned}\text{Maximize}\quad & \psi(X, U) = -1/2X^TCX - b^TU\\ \text{Subject to}\quad & C^TX + A^TU \geq -p\\ & u \geq 0\end{aligned} \tag{26.14}$$

The problem (26.12) is then called the primal problem. Let C_{pa} and C_{Dq} denote the constaint sets of the primal and the dual problem respectively.

Duality Results

Theorem 26.8: Weak Duality Theorem

$$\text{Inf } \phi(X) \geq \text{Sup } \psi(X, U)$$

Proof: It follows directly from Theorem 26.4

Theorem 26.9: Duality Theorem

The existence of an optimal solution of either the primal or the dual problem implies the existence of an optimal solution of the other and then their extreme values are equal.

Proof: Let X^0 be an optimal solution of the primal problem. Since by Theorem 26.8, the quadratic function $\psi(X, U)$ is bounded from above, it attains its maximum (see Theorem 2.1) and hence there exists an optimal solution of the dual problem.

Conversely, if $(\hat{X}, \hat{U})$ is an optimal solution of the dual problem, then by the same argument the primal problem also has an optimal solution.

Furthermore, if X^0 ia an optimal solution of the primal problem, then for some U^0, (X^0, U^0) satisfy the Kuhn–Tucker conditions (26.13) and we have (X^0, U^0) feasible for the dual problem and

$$\begin{aligned}\phi(X^0) &= p^TX^0 + 1/2X^{0T}CX^0\\ &= p^TX^0 + 1/2X^{0T}CX^0 - (p^TX^0 + X^{0T}CX^0 + b^TU^0) && \text{since the expression in the bracket is equal to zero.}\\ &= -1/2X^{0T}CX_0 - b^TU^0\\ &= \psi(X^0, U^0) \leq \psi(\widehat{X}, \widehat{U}), && \text{since } \hat{X}, \hat{U} \text{ is an optional solution of the dual}\end{aligned}$$

Hence, by Theorem 26.8 $\phi(X^0) = \psi(\hat{X}, \hat{U})$

Theorem 26.10: Existence Theorem

If both the primal and the dual problems have feasible solutions, then both have optimal solutions.

Proof: Since both the primal and dual problems are feasible, then by Theorem 26.8, $\phi(X)$ is bounded below on C_{PQ} and $\psi(X, U)$ is bounded above on C_{DQ} and since $\phi(X)$ and $\psi(X, U)$ are quadratic functions, there exists points $X_0 \in C_{pQ}$ and $(\hat{X}, \hat{U}) \in C_{DQ}$ such that

$$\phi(X^0) = \underset{X \in C_{PQ}}{\text{Min}} \phi(X) \text{ and } \psi(\hat{X}, \hat{U}) = \underset{(X,U) \in C_{DQ}}{\text{Max}} \psi(X, U)$$

Theorem 26.11: Unboundedness Theorem

If one of the primal and dual problems is feasible while the other is not, then on its constraint set, the objective function of the feasible problem is unbounded in the direction of optimization.

Proof: Let the primal problem is feasible, so that there exists a vector X satisfying $AX \leq b$, $X \geq 0$. Now, the dual problem is infeasible means that there is no solution of

$$\begin{pmatrix} C & A^T \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ U \end{pmatrix} \geq \begin{pmatrix} -p \\ 0 \end{pmatrix}$$

By inequality Theorem 7.10 then there exists vector (Z, V) satisfying

$$\begin{pmatrix} C & 0 \\ A & 1 \end{pmatrix} \begin{pmatrix} Z \\ V \end{pmatrix} = 0, \ (p^T, 0) \begin{pmatrix} Z \\ V \end{pmatrix} = 1$$

$$Z \geq 0, V \geq 0$$

Or,

$$CZ = 0$$
$$AZ \leq 0$$
$$p^T Z = -1$$
$$Z \geq 0$$

It then follows that (X + tZ) is feasible for the primal problem for all $t \geq 0$ and

$$\begin{aligned} \phi(X + tZ) &= p^T(X + tZ) + \tfrac{1}{2}(X + tZ)^T C(X + tZ) \\ &= p^T X + tp^T Z + 1/2 X^T CX \qquad \text{because } CZ = 0 \\ &\to -\infty \text{ as } t \to +\infty, \text{ since } p^T Z = -1 \end{aligned}$$

The other case can be proved similarly.

Several approaches to duality formulation in nonlinear programming are available in the literature. For early results on duality, see Dennis [124], Dorn [129], Wolfe [525], Hanson [221], and Mangasarian [314]. In our presentation, the formulation is based on the saddle point problem of the associated Lagrange function and for the proofs to establish duality relations we have followed Wolfe [525] and Huard [238]. For quite a different approach, which is based on the concept of conjugate function of Fenchel [157], the reader may consult Rockafellar [379, 380] and Whinston [510, 511]. For symmetric duality, see Cottle [83], Dantzig, Eisenberg and Cottle [111], Mond [348] and for minimax approach to duality, Stoer [445], Mangasarian and Ponstein [323].

For other studies on Duality, see Karamardian [266], Geoffrion [197] and Lasdon [296].

Exercises

1. Obtain a dual to the problem

   Minimize $f(X)$

   Subject to $g_i(X) \leq 0, \; i = 1, 2, \ldots m$

   $h_i(X) = 0, \; i = m + 1, \ldots k$

   $X \geq 0$

   where the function f and g_i $(i = 1, 2, \ldots m)$ are differentiable convex function and h_i $(i = m + 1, \ldots k)$ are linear.

2. Write the dual of the problem

   Minimize $f(X)$

   Subject to $g_i(X) \leq 0, \; i = 1, 2, \ldots m$

   $AX = b$

   $X \geq 0$

   Where f and g_i $(i = 1, 2, \ldots m)$ are differentiable convex function and A is a $k \times n$ matrix.

3. Give the dual of the problem

   Minimize $2x_1 + x_1^2 + x_2^2$

   Subject to $x_1^2 + x_2^2 \leq 1$

   $3x_1 + 2x_1 \geq 1$

   $x_1, x_2 \geq 0$

4. Show that the dual of the quadratic programming problem

   Minimize $C^TX + X^TBX$

   Subject to $X \geq 0$

   Where B is a positive definite matrix, can be considered to be

   Maximize $-Y^TB^{-1}Y$

   Subject to $-Y \leq C$

5. Obtain the dual of the quadratic programming problem

   Minimize $Z = p^TX + \frac{1}{2} X^TCX$

   Subject to $AX = b$

   $X \geq 0$

   Where C is a positive semidefinite matrix

6. Is it true that the dual of the dual of the problem in Q5 above is the primal?

CHAPTER 27

Stochastic Programming

27.1. Introduction

In linear programming, the coefficients of the linear functions are assumed to be constants. However, this is frequently not a very realistic assumption. A problem of stochastic linear programming arises when the coefficients of the linear functions, i.e. the parameters of the linear programming model are random variables. The linear programming model for such a case, however, has no meaning and it is necessary to formulate a new model to deal with such cases.

Expected Volume Model [102]

The initial approach to reduce the effect of uncertainty in the problem was to replace the random variables by their expected values or by some good estimates of them and then to solve the resulting linear program. In many practical situations, however, the solution under such a formulation may not be feasible to the original problem (Exercise 2) and even if it is feasible it may lead to a misleading result.

For example, consider the linear programming problem

$$\begin{array}{ll} \text{Minimize} & z = c^T X \\ \text{Subject to} & AX \geq b \\ & X \geq 0 \end{array}$$

Where A is an mxn matrix, b is an m-vector and c, X are n-vectors.

Suppose that only the elements of the cost vector c are random variables. Then for a given X, z is a random variable and we solve the problem by replacing z by its expected value $\bar{z}$ (X) = Ez (X) = $\bar{c}^T X$. Now, suppose that the feasible points X^1 and X^2 are both optimal so that Ez(X^1) = Ez (X^2) but var z (X^1) > var z (X^1). If var z (X) is very much greater than var z (X^2), it may be dangerous to select X^1 as our desired optimal solution.

Minimum Variance Model [328]

The above discussion on the expected volume solution procedure for a linear programming problem where only the cost elements are random variables, shows that it may be desirable to control the variance of z for a fixed tolerance value of the expected cost. The problem then reduces to

$$\begin{aligned} \text{Minimize var } z \quad & = X^T B X \\ \text{Subject to} \quad & AX \geq b \\ & Ez \leq \alpha \\ & X \geq 0 \end{aligned}$$

Where α is an upper bound which we want to improve on the expected cost and B is a positive semidefinite matrix.

The problem is a quadratic programming problem and can be solved by Wolfe's method.

Alternatively, α may be taken a parameter and the problem is solved as a parametric programming problem.

There are essentially two different approaches to deterministic formulation of stochastic linear programs, namely the 'wait and see' and the 'here and now' models.

In the 'wait and see' model, the decision maker waits for the realization of the random variables and then solves the resulting linear program. The optimal solution and the optimal value of the objective function being functions of random variables are also random variables and the question arises as to what are their expectations and variances. It is therefore natural to determine the probability distribution of the optimal solution or the optimal value of the problem. This problem of finding the distribution is known as the **distribution problem.**

In 'here and now' model, a decision has to be taken at the very beginning before the realization of the random variables. In a linear programming problem, where some or all the parameters (A,b,c) are random variables with a known joint probability distribution, the problem is to determine an $X \geq 0$ which satisfies the constraints with a certain preassigned probability α and minimize the expected value of the objective function.

The problem is then reduced to

$$\begin{aligned} \text{Minimize} \quad & E\, C^T X \\ \text{Subject to} \quad & P\{AX \geq b\} \geq \alpha \\ & X \geq 0 \end{aligned}$$

where P denotes probability

This is called the **chance constrained programming problem**. The chance constrained programming technique was originally developed by Charnes and Cooper [70].

Two stage problems [102, 114]

A Special 'here and now' approach to deal with the stochastic linear program is to solve the problem in two stages. In the first stage, a vector $X \geq 0$ is determined which is feasible for the problem for some estimated values of the random vectors (A, b, c). After observing the realization of the random vectors a recourse or second stage activity vector $Y \geq 0$ is introduced in the constraint to compensate for any discrepancies between AX and b at an additional cost f called the penalty cost. The decision maker then wants to minimize the expected value of the sum of the original costs and the penalty costs (which are random variables) over the modified constraints.

The two stage problem or the stochastic linear program with recourse where only b is random then takes the form

$$\begin{aligned} &\text{Minimize} && c^TX + E \min_{Y} f^TY \\ &\text{Subject to} && AX + BY = b \\ & && A, X = b^1 \\ & && X, Y \geq 0 \end{aligned}$$

where A is an $m \times n$ matrix, B is $m \times \bar{n}$ A_1 is $\bar{m} \times n$, b^1 is $\bar{m} \times 1$ and b is random m-vector with known distribution and c and f are known n and $\bar{n}$ dimensional vectors. The problem is said to be complete [502] when the matrix B (after an appropriate rearrangement of rows and columns) can be partitioned as $B = (I, -I)$. The problem then becomes

$$\begin{aligned} &\text{Minimize} && c^TX + E \min_{Y} (f^{+T}Y^{+}+f^{-T}Y^{-}) \\ &\text{Subject to} && AX + IY^{+} - IY^{-} = b \\ & && A_1X = b^1 \\ & && X \geq 0, Y^{+} \geq 0, Y^{-} \geq 0 \end{aligned}$$

As an illustration of the two stage problem, consider the following simple example due to Dantzig [102]

Example 27.1

Suppose that a factory has 100 items on hand which may be shipped to an outlet at the cost of \$1 a piece to meet an uncertain demand d. If the demand exceeds the supply it is necessary to meet the unsatisfied demand by purchases from the local market at \$2 a piece. Let the demand d be uniformly distributed between 70 and 80. The problem is to find the quantity to be shipped so that the total cost of shipping is minimum and can be stated as

$$\begin{aligned} &\text{Minimize } z = x_1 + 2E \max(0, d - x_1) \\ &\text{Subject to } \quad x_1 + x_2 = 100 \\ &\qquad\qquad\qquad x_1 + x_3 - x_4 = d \\ &\qquad\qquad\qquad x_j \geq 0, j = 1, 2, 3, 4 \end{aligned}$$

where x_1 = number shipped from the factory

x_2 = number stored at the factory

x_3 = number purchased from local market

x_4 = excess of supply over demand

d = unknown demand uniformly distributed between 70 and 80.

It can be easily verified that

$$z = \begin{cases} 150 - x_1 & \text{if } 0 \le x_1 \le 70 \\ (x_1 + 1/10(80 - x_1)^2 = 1/10(75 - x_1)^2 + 77.5, & \text{if } 70 \le x_1 \le 80 \\ x_1 & \text{if } x_1 \ge 80 \end{cases}$$

It then follows that min z = 77.5 at x_1=75 = E (d). This means that the factory should ship 75 items which is the expected demand in this case. It should however, be noted that it is not always best to ship the expected demand.

A large number of papers with variations and extensions of the above models have appeared in the literature. An excellent bibliography on stochastic programming has been prepared by Roger J–B wets (private circulation).

In section 27.2, we present the deterministic formulation of the general stochastic linear program as suggested by Sinha [422], under the assumption that at least the means, variances and covariances of the random variables are known, which can be dealt with in a wholly constructive manner.

27.2 General Stochastic Linear Program [422, 423]

We consider a linear programming problem where all the parameters (A, b, c) of the problem are random variables. With the assumption that at least the means, variances and covariances of the random variables are known, the stochastic linear programming problem is reduced to a deterministic convex programming problem, which can be dealt with in a wholly constructive manner.

27.2.1. Mathematical Formulation

Consider the linear programming problem

$$\text{Minimize} \quad \sum_{j=1}^{n} d_j^0 x_j$$

$$\text{Subject to} \quad \sum_{j=1}^{n} a_{ij}^0 x_j \le b_i^0 x_{n+1}, \; i=1, 2 \ldots m$$

$$x_{n+1} = 1$$

$$x_j \ge 0, \; j = 1, 2, \ldots n + 1 \tag{27.1}$$

Where the parameters d_j^0, a_{ij}^0 and b_i^0 (i = 1, 2…m; j = 1, 2,…n) are random variables whose joint distribution for fixed i is independent of the choice of x_j. Under the situation, the problem in the usual linear programming format is no longer meaningful and a reformulation of the problem is necessary. The uncertainty aspect of the problem, however suggests that it can only be solved probabilistically and hence a reasonable formulation of the stochastic programming problem requires that our activity levels should be such that with a certain preassigned high probability β_i $(0 < \beta_i < 1)$, i = 1, 2,...m the total quantities required for each item should not exceed the available quantities and at the same time should guarantee a minimum objective with a preassigned high probability β_0 $(0 < \beta_0 < 1)$.

So we require

$$P\left[\sum_{j=1}^{n} a_{ij}^0 x_{ij} - b_i^0 x_{n+1} \le 0\right] \ge \beta_i, \quad i = 1, 2,\ldots m \tag{27.2}$$

$$x_{n+1} = 1 \tag{27.3}$$

$$x_j \ge 0, \;\; j = 1, 2,\ldots n+1 \tag{27.4}$$

Where P denotes probability

Among the feasible vectors X (i.e. vectors $X = (x_1, x_2 \ldots x_{n+1})^T$ satisfying (27.2), (27.3) and (27.4) we should then select one for which the upper $(1 - \beta_0)$ probability point of the corresponding distribution of the objective function is a minimum.

So we define a preference functional

$$Z(X) = \bar{e}_0(X) + t_0 \sigma_{e_0(X)} \tag{27.5}$$

and minimize it with respect to X satisfying (27.2), (27.3) and (27.4), where

$$e_0(X) = \sum_{j=1}^{n} d_j^0 x_j, \bar{e}_0(X) = Ee_0(X) = \text{Expectation of } e_0(X)$$

$$\sigma^2_{e_0(X)} = E\,[e_0(X) - \bar{e}_0(X)]^2 = \text{variance of } e_0(x) \tag{27.6}$$

and t_0 is a constant determined by

$$P\,[e_0(X) \le Z(X)] \ge \beta_0 \tag{27.7}$$

We consider the following two cases:

(i) The random variables d_j^0, a_j^0, b_i^0 (j = 1, 2,...n) have known joint normal distributions for each i = 1, 2,...m.

(ii) Only their means, variances and covariances are known.

Case (i): Linear programming coefficients have known joint normal distributors.

$$\text{Let } e_i(X) = \sum^{n} a_j^0 x_j - b_i^0 x_{n1}, \; i = 1, 2,\ldots m \tag{27.8}$$

Since e_i (x), (i = 0, 1, 2,...m) are linear combinations of normal variables, they themselves are normally distributed [90] with mean $\bar{e}_i(X) = Ee_i(X)$, and

$$\text{variance } \sigma^2_{e_i(X)} = E[e_i(X) - \bar{e}_i(X)]^2, \quad i = 0, 1, 2 \ldots m \tag{27.9}$$

From the Normal Probability table [367], we then determine the t_i's so that

$$P\left[e_i(X) \le \bar{e}_i(X) + t_i \sigma_{e_i(X)}\right]$$

$$= P\left[\frac{e_i(X) - \bar{e}_i(X)}{\sigma_{e_i(X)}} \le t_i\right] = \beta_i, \quad i = 0, 1, 2 \ldots m \tag{27.10}$$

and our requirement (27.2): $P[e_i(X) \le 0] \ge \beta_i, i = 1, 2, \ldots m$
is thus equivalent to the condition that

$$\bar{e}_i(X) + t_i \sigma_{e_i(X)} \le 0, \; i = 1, 2, \ldots m \tag{27.11}$$

Case (ii): Only the means, variances and covariances of the coefficients in the linear programming problem are known.

Even if it is not known that the coefficients in the linear programming problems are jointly normally distributed, but only their means, variances and covariances are known, we proceed as follows:

By Techebysheff's extended lemma [165]

$$P\left[\frac{e_i(X) - \bar{e}_i(X)}{\sigma_{e_i(X)}} \ge t_i\right] \le \frac{1}{1 + t_i^2}, \quad i = 0, 1, 2, \ldots m$$

or $$P\left[e_i(X) \le \bar{e}_i(X) + t_i \sigma_{e_i(X)}\right] \ge \frac{t_i^2}{1 + t_i^2}, \; i = 0, 1, 2, \ldots m \tag{27.12}$$

We now take

$$t_i^2 = \frac{\beta_i}{1 - \beta_i} \qquad i = 0, 1, 2 \ldots m \tag{27.14}$$

and it is seen that any X satisfying

$$\bar{e}_i(X) + t_i \sigma_{e_i(X)} \le 0, \qquad i = 1, 2, \ldots m \tag{27.14}$$

will satisfy our requirement (27.2)

Thus our problem reduces to

$$\begin{aligned} \text{Minimize} \quad & \bar{e}_0(X) + t_0 \sigma_{e_0(X)} \\ \text{Subject to} \quad & \bar{e}_i(X) + t_i \sigma_{e_i(X)} \le 0, \; i = 1, 2, \ldots m \\ & X_{n+1} = 1 \\ & X_j \ge 0, \qquad j = 1, 2 \ldots n+1 \end{aligned} \tag{27.15}$$

which can be expressed as

$$\begin{aligned}
\text{Minimize} \quad & D^T X + t_0 (X^T B^0{}_1 X)^{1/2} \\
\text{Subject to} \quad & \bar{A}_i X + t_i (X^T B^i{}_1 X)^{1/2} \leq 0, \quad i = 1, 2, \ldots m \\
& x_{n+1} = 1 \\
& X \geq 0
\end{aligned} \tag{27.16}$$

where

$D^T = (d_1, d_2 \ldots d_n, d_{n+1})$; $d_j = Ed_j^0$, $j = 1, 2, \ldots n + 1$, and $d^0{}_{n+1} = 0$

$\bar{A}_i = (a_{i1}, a_{i2}, \ldots a_{in}, -b_i)$; $a_{ij} = Ea_{ij}^0$, $j = 1, 2 \ldots n$ and

$b_i = Eb^0{}_i$; $\quad i = 1, 2, \ldots m$

$X^T = (x_1, x_2, \ldots x_{n+1})$ a $1 \times (n + 1)$ matrix

$B^i{}_1 = (b^i{}_{jk})$, a$(n + 1) \times (n + 1)$ symmetric positive semidefinite matrix, where $b^i{}_{jk} = E(a^0{}_{ij} - a_{ij})(a^0{}_{ik} - a_{ik})$ and $a^0{}_{in+1} = -b^o{}_i$; $j, k = 1, 2, \ldots n, (n + 1)$; $i = 1, 2 \ldots m$.

$B^0{}_i = (b^0{}_{jk})$, a $(n + 1) \times (n + 1)$ symmetric positive semidefinite matrix, where

$b^0{}_{jk} = E(d^0{}_j - d_j)(d^0{}_k - d_k)$; $j, k = 1, 2, \ldots n + 1$; and

t_i, $i = 0, 1, 2, \ldots m$ are known constant.

For convenience, we now write (27.16) in the form:

$$\begin{aligned}
\text{Minimize} \quad & F(X) = D^T X + (X^T B^0 X)^{1/2} \\
\text{Subject to} \quad & f_i(X) = A_i X + (X^T B^i X)^{1/2} \leq b_i, \; i = 1, 2, \ldots m \\
& F_{m+1}(X) = A_{m+1} X = 1 \\
& x \geq 0
\end{aligned} \tag{27.18}$$

Where

$$\begin{aligned}
& A_i = (a_{i1}, a_{i2}, \ldots a_{in+1}), \; a_{in+1} = 0, \; i = 1, 2, \ldots m \\
& A_{m+1} = (0, 0 \ldots 1), \text{ a } 1 \times (n + 1) \text{ matrix and} \\
& B^i = t^2{}_i B^i{}_1 . \quad i = 0, 1, 2, \ldots m
\end{aligned} \tag{27.19}$$

The Stochastic linear programming problem is therefore reduced to the case of nonlinear programming, where the nonlinearity occurs in the objective function as well as in the constraints in the form of square roots of positive semidefinite quadratic forms. It can be shown that the functions $F(X)$, $f_i(X)$, $(i = 1, 2, \ldots; m + 1)$ in (27.18) are convex functions.

Generalizing the Cauchy–Schwartz inequality, we have

Lemma 27.1. If C is a real symmetric positive semidefinite matrix, then

$$(X^T C Y)^2 \leq (X^T C X)(Y^T C Y), \text{ for all } X, Y \in R^n$$

and equality holds if and only if CX and CY are linearly dependent, i.e. say $CX = \lambda CY$ where λ is a real number.

For a short proof see [146]

Lemma 27.2. $F(X)$, $f_i(X)$, $(i = 1, 2, \ldots m + 1)$ in (27.18) are convex functions of $X \in R^n$.

Proof: For any $X_1, X_2 \in R^n$ and $\alpha, \beta \geq 0, \alpha + \beta = 1$

$$f_i(\alpha X_1 + \alpha X_2) = \alpha A_i X_1 + \beta A_i X_2 + [(\alpha X_1 + \beta X_2)^T B^i (\alpha X_1 + \beta X_2)]^{1/2}$$
$$= \alpha A_i X_1 + \beta A_i X_2 + [\alpha^2 X^T_1 B^i X_1 + 2\alpha\beta X^T_1 B^i X_2 + \beta^2 X^T_2 B^i X_2]^{1/2}$$
$$\leq \alpha A_i X_1 + \beta A_i X_2 + [\alpha^2 X_1^T B^i X_1 + 2\alpha\beta (X^T_1 B^i X_1)^{1/2} (X_2 B^i X_2)^{1/2}$$
$$+ \beta^2 X^T_2 B^i X_2]^{1/2} \qquad \text{(by Lemma 27.1)}$$
$$= \alpha A_i X_1 + \beta A_i X_2 + (\alpha X^T_1 B^i X_1)^{1/2} + \beta (X^T_2 B^i X_2)^{1/2}$$

or $\quad f(\alpha X_1 + \beta X_2) \leq \alpha f_i(X_1) + \beta f_i(X_2), \quad i = 1, 2, \ldots m$

The proof that F(x) is convex is analogous and $f_{m+1}(X)$, being linear is convex.

Our problem therefore, is to minimize a convex function subject to convex constraints , which ensures that any local minimum encountered in the problem will give the global solution desired. But our functions need not be differentiable (as they involve positive semidefinite forms and hence can vanish) and we cannot therefore apply any of the various methods of convex programming known in the literature based on differentiability assumptions. However, Dantzig [109] has developed a method for solving convex programming problems without the assumption of differentiability and his method can be suitably applied to solve our problem.

It is interesting to note that when all the corelation coefficients involved in the problem are unity, the problem reduces to a linear programming problem.

27.3. The Sochastic Objective Function

Consider the case where the coefficients in the constraints and the resources in a linear programming problem are constants but the coefficients in the objective funtion are random variables. The problem is reduced to a nonlinear program where nonlinearity occurs only in the objective function. Since it is difficult to solve the problem directly, we first obtain a solution of a dual problem. A solution of the (primal) problem is then obtained with the help of the solution obtained for the dual problem.

27.3.1. The problem and its Dual

Our problem can be stated as

$$\begin{aligned} &\text{Maximize} && F(X) = D^T X - (X^T B X)^{1/2} \\ &\text{Subject to} && AX \leq b \\ &&& X \geq 0 \end{aligned} \qquad (27.20)$$

where A is a mxn matrix, b a m × 1, D, X are nx1 matrices and B is a mxn positive semidefinite matrix.

It can be shown that a formal dual problem to (27.20) is given by

$$\begin{aligned} &\text{Minimize} && G(Y) = b^T Y \\ &\text{Subject to} && A^T Y + BW \geq D \\ &&& W^T B W \leq 1 \end{aligned}$$

$$Y \geq 0 \tag{27.21}$$

Where we assume that if (27.21) is feasible then there exist (Y,W) feasible for (27.21) with $W^TBW < 1$. (Assumption P)

This assumption however, does not seem to be a serious restriction. In fact, if it is not satisfied, we can suitably perturb the variance matrix B, so that there exists a feasible solution of the dual problem with $W^TBW < 1$.

The problem (27.20) is called the primal problem and (27.21), the dual problem. Let the constraint set of the primal problem be devoted by C_p and that of the dual by C_D. The symbol Ø denotes the empty set.

27.3.2. Duality

It will now be shown that a duality relation holds between (27.20) and (27.21) in the sense that

(a) Sup F(X) = Inf G(Y) (Inequality or Weak Duality Theorem)
(b) The existence of an optimal solution of one of these problems implies the existence of an optimal solution of the other in which case their extreme values are equal. (Duality Theorem)
(c) If one problem is feasible, while the other is not, then on its constraint set, the objective function of the fesible problem is unbounded in the direction of optimization (Unboundedness theorem)
(d) If both problems are feasible, then both have an optional solution (Existence Theorem)

Theorem 27.1. Sup F(X) = Inf G(Y)

Proof: Using the convention that

$$\text{Sup } F(X) = -\infty, \text{ if } C_p \text{ is empty.}$$

and $$\inf G(Y) = +\infty, \text{ if } C_D \text{ is empty.}$$

it remains to prove the inequality under the assumption that both problems are feasible.

Let X and (Y,W) be feasible solutions of the primal and the dual problems, respectively.

We then have

$$\begin{aligned} F(X) = D^TX - (X^TBX)^{1/2} &\leq D^TX - (X^TBX)^{1/2}(W^TBW)^{1/2} \\ &\leq D^TX - W^TBX \text{ (by Lemma 27.1)} \\ &\leq Y^TAX \leq b^TY = G(Y) \end{aligned} \tag{27.22}$$

which proves the theorem.

We recall the result obtained by Eisenberg in [146], which we give here as a lemma.

Lemma 27.3. Let C be a real symmetric nxn positive semidefinite matrix and A be a real mxn matrix, so that $G = R^n \cap \{X | AX \leq 0\}$ is a polyhedral convex cone, Let

$U = R^n \cap \{u | X\, G \geq u^T X \leq (X^T C X)^{1/2}\}$, and

$V = \{v \mid \exists\, Y \in R^m{+}\; X \in G$ and

$v = A^T Y + CX,\; X^T C X = 1\}$,

Where $R^m_+ = R^m \cap \{Y | Y \geq 0\}$,

Then $U = V$

Theorem 27.2. If X_0 is an optimal solution of the primal problem, there exists a (Yo,Wo) so that (Yo,Wo) is an optimal solution of the dual problem and the extrema are equal.

Proof: Let X_0 be an optimal solution of the primal problem and $F(X_0) = M$.

Consider the set

$$K = R^{n+1} \cap \{(X, \lambda) \mid AX - b\lambda \leq 0,\; X \geq 0,\; \lambda \geq 0\} \tag{27.23}$$

so that K is a polyhedral convex cone. The set can be rewritten as

$$K = R^{n+1} \cap \{(X, \lambda) \mid A_1 X - b_1 \lambda \leq 0\}$$

where

$$A_1 \begin{bmatrix} A \\ -1 \\ 0 \end{bmatrix}, b_t^T = (b^T, 0, \ldots 0, 1) \tag{27.24}$$

are (m+n+1) x n and 1x(m+n+1) matrices respectively and I is a (n × n) identity matrix.

Now, if $(X, \lambda) \in K$ and $\lambda > 0$, then $\lambda^{-1} X \in C_p$

and

$$F(\lambda^{-1} X) = \lambda^{-1} F(X) \leq M$$

or $F(X) \leq \lambda M$

while if $\lambda = 0$, then $AX \leq 0$, $X \geq 0$ and therefore $X_0 + tX \in C_p$ for all $t \geq 0$

Hence

$$M \geq F(X_0 + tX) \quad F(X_0) + tF(X)$$

since F(X) is a homogenous concave function.

Now, if $F(X) > 0$, taking large value of t, we can make the right side of the above expression greater than M. Hence we must have $F(X) \leq 0$.

Thus,

$$(X, \lambda) \in K \Rightarrow F(X) \leq \lambda M.$$

i.e.

$$(X, \lambda) \in K \Rightarrow (D^T, -M)(X^T, \lambda)^T \leq [(X, \lambda)^T B_1 (X^T, \lambda)^T]^{1/2} \tag{27.25}$$

where $B_1 = \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix}$

From Lemma 27.3 it follows that there exists a $Y \in R_+^{m+n+1}{}_+$ and $(X, \mu) \in K$, such that

$$(D^T, -M)^T = (A_1, -b_1)^T Y + BX,\ X^T BX \le 1$$

$$\text{or}\quad \begin{aligned} &D = (A^T, -I, 0)\, Y + BX \\ &-M = -\,(b^T, 0, \ldots 0, 1)Y \\ &X^T BX \le 1 \end{aligned} \tag{27.26}$$

which implies that there exists a (Yo, Wo) such that

$$\begin{aligned} &A^T Yo + BWo \ge D \\ &Wo^T\, BWo \le 1 \\ &Yo \ge o \end{aligned} \tag{27.27}$$

$$\text{and}\quad b^T Yo \le M$$

From theorem 27.1, it follows that $b^T Yo = M$ and then (Yo, Wo) is an optimal solution of the dual problem which proves the theorem.

Converse

Now, suppose that (Yo, Wo) is an optimal solution of the dual problem and consider the following linear program:

$$\begin{aligned} \text{Minimize}\quad & g(Y) = b^T Y \\ \text{Subject to}\quad & A^T Y + BW \ge D \\ & Y \ge 0 \end{aligned} \tag{27.28}$$

There are three possibilities

(a) (27.28) may have an optimal solution with $W^T BW \le 1$, or

(b) every optimal solution of (27.28) may satisfy $W^T BW > 1$, or

(c) (27.28) may have unbounded solution.

It is noted that the system of inequalities

$$\begin{aligned} &A^T Y \ge 0 \\ &b^T Y < 0 \\ &Y \ge 0 \end{aligned} \tag{27.29}$$

has no solution. For, if Y is a solution of (27.29), then (Yo + tY, Wo) is a feasible solution of the dual problem and G (Yo + tY) $\to -\infty$ for $t \to \infty$, contradicting the assumption that (Yo, Wo) is optimal for the dual problem.

Now, (27.28) has an unbounded solution if and only if there exists a solution (Y_2, W_2) of

$$\begin{aligned} &A^T Y + BW \ge 0 \\ &b^T Y\ < 0 \\ &Y \ge 0 \end{aligned} \tag{27.30}$$

so that for any feasible solution (Y_1, W_1) of (27.28), (Y_1+tY_2, W_1+tW_2) is feasible for (27.28) for all $t \ge 0$ and $g(Y_1 + tY_2) = b^T(Y_1+tY_2) \to -\infty$ for $t \to \infty$.

Then,

$$(W_1 + tW_2)^T B(W_1+tW_2) = t^2W^T_2BW_2+2tW^T_1BW_2+W^T_1BW_1$$

Since the system of inequalities (27.29) has no solution, BW_2 cannot be a null vector and since B is positive semidefinite $W^T_2BW_2 > 0$.

Hence $(W_1+tW_2)^T B(W_1+tW_2) \to \infty$, for $t \to \infty$.

Thus in the case that g(Y) is unbounded on the constraint set of (27.28), W^TBW also tends to infinity.

Summarizing the above, we have the following two cases

Case (i): (27.28) may have an optimal solution with $W^TBW \leq 1$ (i.e. solution satisfies the constraints of the dual),

Case (ii): it may have a solution with $W^TBW > 1$ (i.e. solution does not satisfy the constraints of the dual).

Case (i):

Lemma 27.4: If (Yo, Wo) is an optimal solution of the dual problem and (27.28) has an optimal solution satisfying the constraints of the dual then (Yo, Wo) is an optimal solution of (27.28).

Proof: Let (Y_1,W_1) be an optimal solution of (27.28) such that $W^T_1BW_1 \leq 1$. Since (Y_1,W_1) is a feasible solution of the dual problem,

$$b^TY_0 \leq b^TY_1 \tag{27.31}$$

Further, since (Yo, Wo) is a feasible solution of the linear program (27.28),

$$b^TY_1 \leq b^TY_0 \tag{27.32}$$

Hence $b^TY_0 = b^TY_1$

Theorem 27.3: If (Yo, Wo) is an optimal solution of the dual problem and the linear program (27.28) has an optimal solution satisfying the constraints of the dual, there exists an Xo, so that Xo, is an optimal solution of the primal problem and the extrema are equal.

Proof: By Lemma 27.4, (Yo, Wo) is an optimal solution of (27.28). The duality theorem of linear programming states that there exists an Xo, such that

$$AXo \leq b$$
$$BXo = 0$$
$$Xo \geq 0$$

and $D^TXo = b^TYo$.

Since $BXo = o$, we have

$$F(Xo) = D^TXo - (X^To\, BXo)^{1/2} = b^TYo = G(Yo) \tag{27.34}$$

It then follows from Theorem 27.1 that Xo is an optimal solution of the primal problem and the proof is complete.

Case (ii):

Consider the following quadratic programming problem.

Minimize $G_1(\theta, Y, W) = \theta b^T Y + \frac{1}{2}(1-\theta) W^T B W$

Subject to $A^T Y + BW \geq D$

$$Y \geq 0 \tag{27.35}$$

Where θ is a single real parameter, which can be chosen as convenient between 0 and 1 $(0 \leq \theta \leq 1)$

A method of solving such a problem for all θ, $0 \leq \theta \leq 1$ is discussed in section 27.3.3.

Lemma 27.5: If (Y_θ, W_θ) is a solution of (27.35) for $0 \leq \theta \leq 1$, then $W_\theta^T \; BW_\theta$ is a monotonically increasing function of θ.

Proof: Take any θ_1 and θ_2, where $0 \leq \theta_1 < \theta_2 \leq 1$. Since $\left(Y_{\theta_1}, W_{\theta_1}\right)$ minimizes G_1 (θ_1, Y, W) we have

$$\theta_1 b^T Y_{\theta_1} + \frac{1}{2}(1-\theta_1) W_{\theta_1}^T B W_{\theta_1} \leq \theta_1 b^T Y_{\theta_2} + \frac{1}{2}(1-\theta_1) W_{\theta_2}^T B W_{\theta_2} \tag{27.36}$$

and since $\left(Y_{\theta_2}, W_{\theta_2}\right)$ minimizes G_1 (θ_2, Y, W), we have

$$\theta_2 b^T Y_{\theta_2} + \frac{1}{2}(1-\theta_2) W_{\theta_2}^T B W_{\theta_2} \leq \theta_2 b^T Y_{\theta_1} + \frac{1}{2}(1-\theta_2) W_{\theta_1}^T B W_{\theta_1} \tag{27.37}$$

Multiplying (27.36) by θ_2 and (27.37) by θ_1, we get

$$\theta_1 \theta_2 b^T Y_{\theta_1} + \frac{1}{2}\,\theta_2 (1-\theta_1) W_{\theta_1}^T B W_{\theta_1} \leq \theta_1 \theta_2 b^T Y_{\theta_2} + \frac{1}{2}\,\theta_2 (1-\theta_1) W_{\theta_2}^T B W_{\theta_2}$$

$$\theta_1 \theta_2 b^T Y_{\theta_2} + \frac{1}{2}\,\theta_1 (1-\theta_2) W_{\theta_2}^T B W_{\theta_2} \leq \theta_1 \theta_2 b^T Y_{\theta_1} + \frac{1}{2}\,\theta_1 (1-\theta_2) W_{\theta_1}^T B W_{\theta_1}$$

Adding these two inequalities and rearranging, we get

$$\frac{1}{2}(\theta_2 - \theta_1) W_{\theta_1}^T B W_{\theta_1} \leq \frac{1}{2}(\theta_2 - \theta_1) W_{\theta_2}^T B W_{\theta_2}$$

and hence

$$W_{\theta_1}^T B W_{\theta_1} \leq W_{\theta_2}^T B W_{\theta_2}$$

Lemma 27.6: If for some $\theta = \theta^*$, $0 < \theta^* < 1$, $(Y_{\theta^*}, W_{\theta^*})$ is a solution of (27.35) with $W_{\theta^*}^T B W_{\theta^*} = 1$, then $(Y_{\theta^*}, W_{\theta^*})$ is a solution of the dual problem (27.21)

Proof: Let (Y, W) be any feasible solution of (27.21), then (Y, W) is also feasible for (27.35). Since $(Y_{\theta*}, W_{\theta*})$ minimizes $G_1(\theta^*, Y, W)$, we get

$$\theta^* b^T Y_{\theta*} + \tfrac{1}{2}(1-\theta^*)\, W_{\theta*}^T B W_{\theta*} \leq \theta^* b^T Y + \tfrac{1}{2}(1-\theta^*)\, W^T B W$$

or $$\theta^* b^T Y_{\theta*} \leq \theta^* b^T Y - \tfrac{1}{2}(1-\theta^*)\,(W_{\theta*}^T B W_{\theta*} - W^T B W)$$

Since $W_{\theta*}^T B W_{\theta*} = 1$ and $W^T B W \leq 1$, we obtain

$$\theta^* b^T Y_{\theta*} \leq \theta^* b^T Y$$

or $$b^T Y_{\theta*} = b^T Y$$

and since $(Y_{\theta*}, W_{\theta*})$ is feasible for the dual problem (27.21), the lemma is proved.

Corollary 27.6.1. If $(Y_{\theta*}, W_{\theta*})$ is a solution of (27.35) for $\theta = \theta^*$ with $W_{\theta*}^T B W_{\theta*} = 1$, then $(Y_{\theta*}, W_{\theta*})$ is optimal for (27.35) for $\theta = \theta^*$.

Proof: Since $W_{\theta*}^T B W_{\theta*} = 1$, the unboundedness of $G_1(\theta^*, Y, W)$ will imply the unboundedness of $b^T Y_{\theta*}$, which contradicts the assumption that the dual problem has an optimal solution.

It will now be shown that we can in fact find a $\theta = \theta^*$, $0 < \theta^* < 1$, so that (27.35) has an optimal solution $(Y_{\theta*}, W_{\theta*})$ with $W_{\theta*} B W_{\theta*} = 1$.

Lemma 27.7: There exists a $\theta = \theta^*$, $0 < \theta^* < 1$, for which (27.35) has an optimal solution (Y^*, W^*) with $W^{*T} B W^* = 1$

Proof: For $\theta = 0$. (27.35) reduces to

$$\begin{aligned} \text{Minimize} \quad & W^T B W \\ \text{Subject to} \quad & A^T Y + BW \geq D \\ & Y \geq 0 \end{aligned} \tag{27.38}$$

According to our assumption (P), (27.38) has a feasible solution (Y, W) with $W^T B W < 1$. Hence there exists an optimal solution (Y_1, W_1) of (27.38), [173, Appendix i], with $W_1^T B W_1 < 1$.

Further, for $\theta = 1$, (27.35) reduces to

$$\begin{aligned} \text{Minimize} \quad & b^T Y \\ \text{Subject to} \quad & A^T Y + BW \geq D, \\ & Y \geq 0 \end{aligned}$$

and case (ii) states that there is a solution of this problem with $W^T B W > 1$.

From Lemma 27.5, it then follows that a $\theta = \theta^*$, $0 < \theta^* < 1$ can be obtained for which (27.35) has a solution (Y^*, W^*) with $W^* B W^* = 1$. (This will be illustrated in section (27.3.4.), which by corollary 27.6.1, is in fact an optimal solution.

Lemma 27.8: If (Yo, Wo) is an optimal solution of the dual problem, then $W_o^T BWo = 1$

Proof: Let (Yo,Wo) be an optimal solution of the dual problem (27.21) with $W_o^T BWo < 1$.

From Lemma 27.7, it follows that for some $\theta = \theta^*$, $0 < \theta^* < 1$, (27.35) has an optimal solution (Y*,W*) with $W^{*T}BW^* = 1$ and (Y*,W*) is then feasible for (27.21)

Since (Y*,W*) minimizes $G_1(\theta^*, Y, W)$ and (Yo,Wo) is a feasible solution of (27.35), we have

$$\theta^* b^T Y^* + \tfrac{1}{2}(1-\theta^*)\, W^{*T} BW^* \le \theta^* b^T Yo + \tfrac{1}{2}(1-\theta^*)\, W_o^T BWo$$

or $$\theta^* b^T Y^* \le \theta^* b^T Yo - \tfrac{1}{2}(1-\theta^*)\,(W^{*T} BW^* - W_o^T BWo)$$

and since $W^{*T} BW^* = 1$ and $W_o^T BWo < 1$, we get

$$\theta^* b^T Y^* < \theta^* b^T Yo$$

or $$b^T Y < b^T Yo$$

contradicting the assumption that (Yo,Wo) is an optimal solution of the dual problem.

Lemma 27.9: If (Y*,W*) is an optimal solution of the quadratic problem (27.35) for a $\theta = \theta^*$, $0 < \theta^* < 1$, so that $W^{*T} BW^* = 1$ (and hence (Y*,W*) is also an optimal solution of the dual problem), then (Y*,W*) is an optimal solution of the following linear programming problem.

$$\begin{aligned} &\text{Minimize} && g_1(Y,W) = \theta^* b^T Y + (1-\theta^*)\, W^{*T} BW \\ &\text{Subject to} && A^T Y + BW \ge D \\ &&& Y \ge 0 \end{aligned} \tag{27.39}$$

Proof: The problems (27.35) and (27.39) have the same constraints. Suppose there exists an $(\hat{Y}, \widehat{W})$ satisfying the constraints such that

$$g_1(\hat{Y}, \widehat{W}) < g_1(Y^*, W^*)$$

which means

$$[\theta^* b^T \hat{Y} + \tfrac{1}{2}(1-\theta^*)\, W^{*T} B \widehat{W}] - [\theta^* b^T Y^* + (1-\theta)\, W^{*T} BW^*] < 0$$

or $$\theta^* b^T(\hat{Y} - Y^*) + (1-\theta^*)\, W^{*T} B(\widehat{W} - W^*) < 0 \tag{27.40}$$

Let $0 < \alpha < 1$ and define

$$Y_1 = (1-\alpha)\, Y^* + \alpha\, \hat{Y} = Y^* + \alpha\,(\hat{Y} - Y^*)$$

$$W_1 = (1-\alpha) W^* + \alpha\, \widehat{W} = W^* + \alpha\,(\widehat{W} - W^*)$$

Since the constraints set of (27.35) or (27.39) is convex, (Y_1, W_1) is a feasible solution for (27.35) or (27.39), Consider

$$G_1(\theta^*, Y_1, W) - G_1(\theta^*, Y^*, W^*)$$
$$= \alpha\,[\theta^* b^T(\hat{Y} - Y^*) + (1-\theta^*)\, W^{*T} B(\widehat{W} - W^*)]$$
$$+ \tfrac{1}{2}(1-\theta^*)\,\alpha^2\,(\widehat{W} - W^*)^T B\,(\widehat{W} - W^*) \tag{27.41}$$

From (27.40) and positive semidefiniteness of B, it follows that the right-hand side of (27.41) can be made negative for sufficiently small positive α. This conradicts the assumption that (Y^*, W^*) is an optimal solution of (27.35) for $\theta = \theta^*$. Therefore, (Y^*, W^*) must be an optimal solution of the linear program (27.39)

Theorem 27.4: If (Y^*, W^*) is an optimal solution of the dual problem (27.21) and if the linear program (27.28) has a solution which does not satisfy the constraints of the dual, then there exists an X^*, so that X^* is an optimal solution of the primal problem and the extrema are equal.

Proof: By Lemma 27.8 it follows that (Y^*, W^*) is an optimal solution of the dual with $W^{*T}BW^* = 1$, and hence is an optimal solution of (27.35) for $\theta = \theta^*$. By Lemma 27.9 then, (Y^*, W^*) is an optimal solution of the linear program (27.39). By the duality theorem of linear programming, there exists az, such that

$$\begin{aligned} AZ &\le \theta^* b \\ BZ &= (1-\theta^*)\, BW^* \\ Z &\ge 0 \end{aligned} \tag{27.42}$$

and $D^T Z = \theta^* b^T Y^* + (1-\theta^*)\, W^{*T} B W^*$

or $D^T Z - W^{*T} BZ = \theta^* b^T Y^*$

By Lemma 27.1 and since $W^{*T} B W^* = 1$, we get

$$D^T Z - (Z^T BZ)^{1/2} = \theta^* b^T Y^* \tag{27.43}$$

Let $X^* = (1/\theta^*)\, Z$, then X^* is a feasible solution of the primal problem and

$$D^T X^* - (X^2 BX^*)^{1/2} = b^T Y^* \tag{27.44}$$

From Theorem 27.1, it then follows that X^* is an optimal solution of the primal problem and the theorem is proved.

Theorem 27.5

(a) If $C_p \neq \emptyset$ and $C_D = \emptyset$, then $\sup\limits_{X \in C_p} F(X) = +\infty$

(b) If, $C_p = \emptyset$ and $C_D \neq \emptyset$, then $\inf\limits_{Y \in C_D} G(Y) = -\infty$

Proof: (a) Let $X_o \in C_p$ so that $AX_o \le b$, $X_o \ge 0$ Since $C_D = \emptyset$, the inequalities

$$\begin{aligned} A^T Y &\ge D \\ Y &\ge 0 \end{aligned} \tag{27.45}$$

have no solution, for otherwise, a solution of (27.45) with $W = 0$, will be feasible for the dual problem, contradicting the assumption that $C_D = \emptyset$.

Hence there exists [Theorem 7.11] a solution of

$$AX \leq 0$$
$$D^TX > 0 \tag{27.46}$$
$$X \geq 0$$

Now, if $X \geq 0$, $AX \leq 0 \Rightarrow F(X) \leq 0$

i.e.,

$$\begin{bmatrix} A \\ -I \end{bmatrix} X \leq 0 \Rightarrow F(X) \leq 0$$

then by Lemma 27.3, there exist $Y \in R_+^m$, $Z \in R_+^n$

and $W \in R_+^n$, such that

$$A^TY - Z + BW = D$$
$$W^TBW = 1$$

i.e. there exists a solution of

$$A^TY + BW \geq D$$
$$W^TBW \leq 1$$
$$Y \geq 0$$

which contradicts the assumption that

$C_D = \emptyset$. Hence

$$X \geq 0,\ AX \leq 0 \Rightarrow F(X) > 0$$

For $\lambda \geq 0$, we then have

$$A(Xo + \lambda X) \leq AXo \leq b$$

and $Xo + \lambda X \geq 0$

Hence $Xo + \lambda X \in C_P$ But

$$F(Xo + \lambda X) \geq F(Xo) + \lambda F(X),$$

since F(X) is a homogenous concave function.

And then

$$\lim_{\lambda \to \infty} (Xo + \lambda X) = +\infty$$

(b) Let $(Yo, Wo) \in C_D$, so that

$$A^TYo + BWo \geq D$$
$$W^ToBWo \leq 1$$
$$Yo \geq 0$$

Now, $CP = \emptyset$ implies [Theorem 7.11] that there exists a solution Y of

$$A^TY \geq 0$$
$$b^TY < 0$$

$$Y \geq 0$$

and for $\mu \geq 0$, we have

$$A^T(Yo+\mu Y) + BWo \geq A^T Yo + BWo \geq D$$
$$W^T_o BW_o \leq 1$$
$$Yo + \mu Y \geq 0$$

Hence $(Yo + \mu Y, Wo) \in C_D$, but

$$G(Yo + \mu Y) = b^T Yo + \mu b^T Y$$

and $\lim\limits_{\mu \to \infty} G(Yo + \mu Y) = -\infty$

since $b^T Y < 0$

Corallary 27.5.1

(a) If $Cp \neq \emptyset$, and F(x) is bounded above on Cp, then $C_D \neq \emptyset$.
(b) If $C_D \neq \emptyset$ and G(Y) is bounded below on C_D then $Cp \neq \emptyset$.
(c) If $Cp \neq \emptyset$ and $C_D \neq \emptyset$, then F(x) is bounded above on Cp and G(Y) is bounded below on C_D

Proof: (a), (b): If one problem is feasible and the other is not then by theorem 27.5, the objective function of the feasible problem is unbounded in the direction of optimization contradicting the assumption.

(c): Is an immediate consequence of Theorem 27.1

Theorem 27.6: If both primal and dual problems are feasible, then both have optimal solutions.

Proof: Since C_P and C_D are both nonempty, F(X) is bounded above on C_P and G(Y) is bounded below on C_D.

Let us know consider the problem (27.28), i.e. the problem

Minimize$g(Y) = b^T Y$

Subject to $A^T Y + BW \geq D$

$$Y \geq 0$$

If (27.28) has an optimal solution (Yo,Wo) with $W_o^T BWo \leq 1$, then (Yo,Wo) is also optimal for the dual problem (follows from Lemma 27.4) If there is no optimal solution of (27.28) with $W^T BW \leq 1$, consider the problem (27.35), i.e. the problem.

Minimize$G_1(\theta, Y, W) = \theta b^T Y + ½ (1-\theta) W^T BW$

Subject to $A^T Y + BW \geq D$

$$Y \geq 0$$

Where θ is chosen as convenient between 0 and 1.

Since G(Y) is bounded below on C_D, it follows from Lemma 27.7 that there exists a $\theta = \theta^*$, $0 < \theta^* < 1$ for which (27.35) has an optimal solution (Y^*,W^*)

with $W^{T*}BW^*=1$, which is also an optimal solution of the dual problem (Lemma 27.6)

It then follows from Theorem 27.3 and Theorem 27.4 that there also exists an optimal solution of the primal problem.

The duality relation is now fully established.

27.3.3. Solution of the quadratic program (27.35)

We will now discuss a method for solving the quadratic programming problem (27.35) for all θ, $0 \le \theta \le 1$.

For $\theta = 1$. The problem reduces to a linear programming problem and this can be solved by the usual simplex method.

For $0 \le \theta < 1$, we note that if we set $\lambda = \dfrac{\theta}{1-\theta} \ge 0$, (27.35) is equivalent to the problem.

$$\begin{aligned} &\text{Minimize} \quad G_2(\lambda, Y, W) = \lambda b^T Y + \tfrac{1}{2} W^T B W \\ &\text{Subject to} \quad A^T Y + BW \ge D \\ &\qquad\qquad\quad Y \ge 0 \end{aligned} \tag{27.46}$$

Now, let $W=W_1-W_2$, where $W_1 \ge 0$, $W_2 \ge 0$ are nx1 matrices. (27.46) can then be written in the form

$$\begin{aligned} &\text{Minimize} \quad G_2(\lambda, Z) = \lambda\, b_1^T Z + \tfrac{1}{2} Z^T C Z \\ &\text{Subject to} \quad A_1^T Z = D \\ &\qquad\qquad\quad Z \ge 0 \end{aligned} \tag{27.47}$$

where $A_1^T = (A^T, -I, B, -B)$ is an $n \times (m + 3n)$ matrix,

$b_1^T = (b^T, 0, 0\ldots 0)$, a $1 \times (m + 3n)$ matrix,

$$C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & B & -B \\ 0 & -B & B \end{bmatrix} \text{an } (m + 3n) \times (m + 3n)$$

symmetric positive semidefinite matrix and

$Z^T=(Y^T, y_{m+1}\ldots y_{m+n}, W_1^T, W_2^T)$ is a

$1 \times (m + 3n)$ matrix, y_{m+i} $(1 \le i \le n)$ being scalors.

It may be noted that for a feasible solution of (27.47) with $Z^TCZ = 1$, there is a feasible solution (Y, W) of (27.35) with $W^TBW \le 1$ and hence of the dual problem (27.21) and conversely.

We now assume that (27.47) has a feasible solution and apply Wolfe's method [524] discussed in Chapter 24 to solve this problem for all $\lambda \ge 0$.

Theorem 27.7 [Wolfe]: The problem (27.47) has a solution if and only if there exists a $V \geq 0$, (V is $(m + 3n \times 1)$ and U (U is $n \times 1$) such that

$$\begin{aligned} &V^T Z = 0 \\ &A_1^T Z = D \qquad\qquad (27.48) \\ &CZ - V + A_1 U + \lambda b_1 = 0 \\ &Z \geq 0 \end{aligned}$$

Proof: Suppose Zo,Vo,Uo is a solution of (27.48) and Z is any feasible solution of (27.47). Since C is positive semi-definite, we have

$$(Z - Zo)^T C(Z - Zo) \geq 0$$

or $$Z^T CZ - Z_o^T CZo \geq 2 Z_o^T C (Z - Zo) \qquad (27.49)$$

Now

$$\begin{aligned} G_2(\lambda, Z) - G_2(\lambda, Zo) &= \lambda b_1^T (Z - Zo) + ½ (Z^T CZ - Z_o^T CZo) \\ &\geq \lambda b_1^T (Z - Zo) + Z_o^T C (Z - Zo) \qquad \text{by (27.49)} \\ &= (\lambda b_1^T + Z_o^T C) (Z - Zo) \end{aligned}$$

From (27.48) we know that

$$\lambda b_1^T + Z_o^T C = V^T - U^T A_1^T$$

Hence,

$$\begin{aligned} G_2(\lambda, Z) - G_2(\lambda, Zo) &\geq (V^T - U^T A_1^T) (Z - Zo) \\ &= (V^T Z - V^T Z_o - U^T A_1^T Z + U^T A_1^T Z_o \\ &= V^T Z - o - U^T D + U^T D \\ &\geq 0, \text{ since } V \geq 0, Z \geq 0 \end{aligned}$$

Conversely if (27.47) has a solution, then it must satisfy the Kuhn-Tucker conditions

$$\begin{aligned} &V^T Z = 0, \; A_1^T Z = D, \\ &CZ - v + A_1 U + \lambda b_1 = 0 \\ &V \geq 0, Z \geq 0 \end{aligned}$$

Hence there exists Z, U, V satisfying (27.48)

Theorem 27.8 [Wolfe]: Let A_1, D, C be as before, let the matrix Q be (m+3n) by n_1, q be 1 by x^{n1} and g be (m+3n) by 1. Let $Z \geq 0$, $V \geq 0$ such that $V^T Z = 0$ be given, Denote by Z_z, those components of Z which are positive and by V_z the corresponding components of V (note $V_z = 0$); denote by V_v the positive components of V and by Z_v the corresponding components of Z (note $Z_v = 0$)

If the linear form qw

Is minimal under the linear constraints

$$Vz = 0,$$

$$Zv = 0, \tag{27.50}$$

and

$$A_1Z = D \tag{27.51}$$

$$CZ - IV + A_1U + QW = g$$

then there exists an r such that

$$Cr = 0,\ A_1^T r = 0 \text{ and } qw = rg.$$

Let us now minimize the linear form

$$-\lambda \tag{27.52}$$

subject to (27.48)

where Theorem 27.8 can be applied with

$$Q = b_1,\ q = -1,\ g = 0 \text{ and } W = \lambda$$

Two cases may arise

Case (i): $-\lambda$ has a finite minimum.

Case (ii): λ is unbounded.

Case (i): If (27.52) has a finite minimum, the hypothesis of Theorem (27.8) is satisfied and we note that the minimum value of $-\lambda$ is

$$-\lambda = qW = rg = 0 \text{ and}$$

there exists an r such that

$$A_1^T r = 0,\ Cr = 0$$

Further from the proof of Theorem 27.8, (See section 24.2) we have

$$b_1^T r \leq -1,\ r \geq 0$$

Thus for any $t \geq 0$ and aZ feasible for (27.47), we have

$$A_1^T\ (Z + tr) = D$$

$$Z + tr \geq 0$$

and $\quad G_2(\lambda, Z + tr) = \lambda\, b_1^T Z + \lambda t\, b_1^T r + ½\, Z^T CZ$

Since $\quad b_1^T r \leq -1$

$$G_2\ (\lambda, Z + tr) \rightarrow -\infty \text{ as } t > \infty \text{ for any } \lambda > o \tag{27.53}$$

Note: If there exists an optimal solution of the dual problem (27.21), $-\lambda$ in (27.52) cannot have a finite minimum.

Case (ii): λ is not bounded, since in the minimizing problem (27.52), only a finite number of bases are available, a sequence of basic solutions $(Z^i, V^i, U^i, \lambda^i)$, i = 1, 2,...k will be produced and finally $(Z^{k+1}, V^{k+1}, U^{k+1})$ such that $(Z^k+tZ^{k+1}, V^k+tV^{k+1}, U^k+tU^{k+1}, \lambda^k+t)$, is a feasible solution for all $t \geq 0$. [96] Due to the restriction (27.50), we will have

$$0 = V^iZ^i = V^i Z^{i+1} = V^{i+1} Z^i = V^{i+1} Z^{i+1} \text{ and } \lambda^i < \lambda^{i+1} \text{ for } i = 1, 2,... k \tag{27.54}$$

Now, for given $\lambda^i \leq \lambda \leq \lambda^{i+1}$, let

$$Z = \frac{\lambda^{i+1} - \lambda}{\lambda^{i+1} - \lambda^i} Z^i + \frac{\lambda - \lambda^i}{\lambda^{i+1} - \lambda^i} Z^{i+1}, \quad i = 1, 2,...k\text{-}1 \tag{27.55}$$

and let V and U be respectively the same convex combination of V^i, V^{i+1} and U^i, U^{i+1}. Then (Z, V, U) is a feasible solution of (27.52) and hence satisfies the conditions of Theorem 27.7, so that Z yields the desired minimum in (27.47).

If, on the other hand, $\lambda \geq \lambda^k$, then

$$Z = Z^k + (\lambda - \lambda^k)Z^{k+1},\ V = V^k + (\lambda - \lambda^k)\, V^{k+1},\ U = U^k + (\lambda - \lambda^k)\, U^{k+1} \tag{27.56}$$

satisfies the conditions of Theorem 27.7 and Z is a solution of our problem (27.47).

27.3.4 Solution of the Dual Problem

Suppose that the dual problem is feasible. We then first solve the linear programming problem (27.28)

Case (a): If (27.28) has an optimal solution (Yo,Wo) satisfying the constraints of the dual, then (Yo,Wo) is an optimal solution of the dual problem. (Lemma 27.4)

Case (b): If not, then either G (Y) is unbounded (see (27.53) or from Lemma 27.7 it follows that we find $\lambda^{i+1} > \lambda^i > 0$ in the sequence of the basic solutions of (27.52), such that for $\lambda = \lambda^i$, we obtain an optimal solution $(Y^i_\lambda, W^i_\lambda)$ of (27.46) with

$$0 \leq W^T_{\lambda i}\, BW_{\lambda i} < 1$$

and for $\lambda = \lambda^{i+1}$, an optimal solution

$$(Y_{\lambda^{i+1}}, W_{\lambda^{i+1}}) \text{ with}$$

$$W^T_{\lambda^{i+1}} BW_{\lambda^{i+1}} > 1$$

If $\lambda^i \leq \lambda^{k-1}$, we define

$$Y^* = \alpha\, Y_{\lambda^i} + (1 + \alpha)\ Y_{\lambda^{i+1}}$$

$$W^* = \alpha\, W_{\lambda^i} + (1 + \alpha)\ W_{\lambda^{i+1}}$$

$$0 < \alpha < 1$$

such that $W^{*T}BW^* = 1$, where α is determined from the quadratic equation

$$\alpha^2 W_{\lambda^i}^T BW_{\lambda^i} + 2\alpha(1-\alpha)W_{\lambda^i}^T BW_{\lambda^{i+1}} + (1-\alpha)^2 W_{\lambda^{i+1}}^T BW_{\lambda^{i+1}} - 1 = 0$$

solving $\lambda = \lambda^*$ from

$$\alpha = \frac{\lambda^{i+1} = \lambda}{\lambda^{i+1} - \lambda^i}$$

we note from (27.55) that (Y^*,W^*) is an optimal solution of (27.46) for $\lambda = \lambda^*$ and hence of (27.35) for $\theta = \theta^* = \lambda^*/1+ \lambda^*$ with $W^{T}{}^*BW^* = 1$. (Y^*,W^*) is therefore an optimal solution of the dual problem (Lemma 27.6)

If $\lambda^i \geq \lambda^k$, we define

$$Y^* = Y_{\lambda^k} + \alpha Y_{\lambda^{k+1}}$$

$$W^* = W_{\lambda^k} + \alpha W_{\lambda^{k+1}}$$

Such that $W^{*T}BW^* = 1$, where α is determined from the quadratic equation

$$\alpha^2 W_{\lambda^{k+1}}^T BW_{\lambda^{k+1}} + 2\alpha W_{\lambda^{k+1}}^T BW_{\lambda^k} + W_{\lambda^k}^T BW_{\lambda^k} - 1 = 0$$

Now, solving $\lambda = \lambda^*$ from $\lambda - \lambda^k = \alpha$, we note from (27.56) that (Y^*,W^*) is an optimal solution of (27.35) for $\theta = \theta^* = \lambda^*/1+ \lambda^*$ with $W^{*T}BW^* = 1$ and hence is an optimal of the dual problem.

27.3.5. Solution of the Primal Problem

For obtaining a solution of the primal problem, we first solve the dual problem.

(i) If the dual problem is feasible and G(Y) is unbounded, the primal problem is infeasible (Theorem 27.5)

(ii) If the dual problem is infeasible but the primal feasible, then F(X) is unbounded. This can however, only happen if there exists a solution of

$$\begin{aligned} &AZ \leq 0 \\ &BZ \leq 0 \\ &D^TZ > 0 \\ &Z \geq 0 \end{aligned} \tag{27.57}$$

in which case, for any feasible solution X of the primal problem X + tZ is also feasible for all $t \geq 0$ and $F(X+tZ) \to \infty$ for $t \to \infty$

(27.57) can be solved by the simplex method.

(iii) In case the dual problem does have an optimal solution (Yo,Wo), there exists an optimal solution Xo of the primal problem (Theorem 27.3 and 27.4), such that

$$F(Xo) = G\ (Y0)$$

i.e.

$$D^TXo - (X_0^T BXo)^{1/2} = b^TYo \tag{27.58}$$

We also know from (27.22) that

$$F(Xo) = D^TXo - (X_0^T BXo)^{1/2} \le D^TXo - (X_0^T BXo)^{1/2} (W_0^T BWo)^{1/2}$$

$$\le D^TXo - W_0^T BXo \tag{27.59}$$

$$\le Y_0^T AXo = b^TYo = G(Yo)$$

and by (27.58), equality must hold throughout in (27.59). Hence

$$F(Xo) = D^TXo - (X_0^T BXo)^{1/2} = D^TXo - (X_0^T BXo)^{1/2} (W_0^T BWo)^{1/2}$$

$$= D^TXo - W_0^T BXo \tag{27.60}$$

$$= Y^TAXo = b^TYo = G(Yo)$$

and this is true if any only if

$$BXo = \alpha BWo$$
$$A^TYo = D - BWo$$
$$AXo = b$$
$$Y_0^T AXo = b^TYo$$

where $\alpha = 0$ if the solution of the dual problem is obtained by solving the linear program (27.28) and $\alpha = \frac{1-\theta^*}{\theta^*}$, if the solution is obtained by solving (27.35), we therefore obtain a solution of

$$\begin{aligned} &AX \le b \\ &BX = \alpha BWo \\ &X \ge 0 \end{aligned} \tag{27.62}$$

by linear programming techniques.

A solution of (27.62) will then give us an optimal solution of the primal problem.

27.4 The General Case

Consider the case when all the parameters (A, b, c) in a linear programming problem are random variables. Assuming that the distributions of the random variables are not known but only their means, variances and covariances are known, a reasonable deterministic formulation of the problem can be obtained (see Section 27.2) as,

$$\begin{aligned} \text{Minimize} \quad & F(X) = D^TX + (X^TB^oX)^{1/2} \\ \text{Subject to} \quad & f_i(X) = A_iX + (X^TB^iX)^{1/2} \le b_i,\ i = 1, 2, \ldots m \\ & f_{m+1}(x) = A_{m+1}X = 1 \\ & X \ge 0 \end{aligned} \tag{27.63}$$

where B^i,(i = 0, 1, 2,...m) are symmetric positive semidefinite matrices.

The problem is a convex programming problem but the functions F(X), $f_i(X)$, (i = 1, 2,...m) need not be differentiable and we cannot therefore apply any of the various methods of convex programming based on differentiability assumptions to solve our problem. However, Dantzig [109] has shown that even if the functions are not differentiable, a method can be developed for solving a general convex programming problem, provided some mild regularity conditions are satisfied. It can be seen that the burden of the work in this iterative procedure shifts to a sub-problem, which must be solved afresh at each iteration. This itself is a convex programming problem which may or may not be easy to solve for general convex functions.

We now impose the restriction of boundedness on X in (27.63) (in fact, in many practical situations, the decision maker knows beforehand the upper bounds of the levels of activities that he can employ) and further assume that there exists an X such that $f_i(X) < b_i$ (i = 1, 2,...m) and $f_{m+1}(X) = 1$, so that the regularity assumptions in Dantzig's method are satisfied.

The general problem then becomes

$$\begin{aligned}
\text{Minimize}\quad & F(X) = D^TX + (X^TB^oX)^{1/2} \\
\text{Subject to}\quad & f_i(X) = A_iX + (X^TB^iX)^{1/2} \le b_i,\ i = 1, 2,\ldots m \\
& f_{m+1}(X) = X_{n+1} = 1 \qquad (27.64)\\
& X \in R
\end{aligned}$$

where $R = \{0 \le X \le S\}$ (27.65)

and S is a vector with positive elements.

27.4.1. The subproblem and its Dual

We now apply Dantzig's method (see Section 25.7) to solve our problem (27.64). At each iteration we are then required to solve a convex programming problem called the 'subproblem' which in our case can be stated as

$$\begin{aligned}
\text{Maximize}\quad & \phi(X) = D^TX - \sum_{i=1}^{t}(X^TB^iX)^{1/2} \\
\text{Subject to}\quad & IX \le S \qquad (27.66)\\
& X \ge 0
\end{aligned}$$

where B^i are symmetric positive semidefinite matrices.

As in section 27.3, we first obtain a solution of a dual problem to (27.66), with the help of which a solution of the subproblem is then obtained.

It can be shown that a dual problem to (27.66) is given by

$$\begin{aligned}
\text{Minimize}\quad & \psi(Y) = S^TY \\
\text{Subject to}\quad & IY + \sum_{i=1}^{t} B^iW^i \ge D \qquad (27.67)\\
& W^{iT}B^iW^i \le 1,\quad i = 1, 2,\ldots t, \\
& Y \ge 0
\end{aligned}$$

27.4.2. Duality

It will now be shown that the dual relations hold between (27.66) and (27.67). The sub problem (27.66) will also be called the primal problem and (27.67) the dual.

Theorem 27.9: Sup Ø (X) $\leq$ Inf ψ(Y)

Proof: Let X and (Y, W^i), i = 1, 2,...t be any feasible solution of the primal and the dual problem respectively.

We then have

$$\phi(X) = D^T X - \sum_{i=1}^{t} (X^T B^i)^{1/2} (W^T B^i W^i)^{1/2}$$

$$\leq D^T X - \sum_{i=1}^{t} W^{iT} B^i X \tag{27.68}$$

$$\leq Y^T I X \leq S^T Y = \psi(Y).$$

Hence sup $\phi(X) \leq$ Inf $\psi(Y)$ in the case that both the problems are feasible. The theorem then follows, if we assume the convention that

Sup $\phi(X) = -\infty$, if the primal constraint set is empty

Inf $\psi(Y) = +\infty$, if the dual constraint set is empty.

Extending Lemma 27.3, we now prove the following

Theorem 27.10: Let C^i (i = 1, 2,...t) be real symmetric nxn positive semidefiniee matrices and A be a real mxn matrix. Let

$G = R^n \cap \{X | AX \leq 0\}$ be a polyhedral convex cone and

$$U = R^n \cap \left\{ u \mid X \in G \Rightarrow u^T X \leq \sum_{i=1}^{t} (X^T C^i X)^{1/2} \right\}$$

Let $\quad R^m + = R^m \cap \{\pi | \pi \geq 0\}$ and consider the set

$$V = \left\{ v \mid \exists \pi \in R_+^m \; X^i \in G \text{ with } v = A^T \pi + \sum_{i=1}^{t} C^i X^i ,\; X^{it} C^i X^i \leq 1,\; i = 1, 2, \ldots t \right\}$$

Then U = V

Proof: See [425]

Theorem 27.11: There exists an optimal solution of the primal problem if and only if there exists an optimal solution of the dual problem, in which case their respective extreme values are equal.

Proof: It is clear that the primal problem (27.66) is equivalent to the problem

Maximize $\quad$ Ø(X)

Subject to $\quad Y^TIX \le \psi(Y)$ for all $Y \ge 0$ (27.69)

$X \ge 0$

Consider the problem

Minimize $\quad \psi(Y) = S^TY$

Subject to $\quad X^TIY \ge Ø(X)$ for all $X \ge 0$ (27.70)

$Y \ge 0$

Since $Ø(X)$ and $\psi(Y)$ are positively homogenous continuous, concave and convex functions respectively and since

$$IX \le 0,\ X \ge 0,\ \phi\,(X) \ge 0 \Rightarrow X = 0$$

$$IY \ge 0,\ Y \ge 0,\ \psi\,(Y) \le 0 \Rightarrow Y = 0$$

it follows from a theorem of Eisenberg on duality in homogenous programming [145] that there exists an optimal solution of the problem (27.69) if and only if there exists an optimal solution of the problem (27.70) and then their extreme values are equal.

Our theorem will then follow, if we show that the dual problem (27.67) is equivalent to (27.70).

Let (Y, W^i), $i = 1, 2, \ldots t$ be any feasible solution of the dual problem. Then for all $X \ge 0$,

$$X^TIY \ge D^TX - \sum_{i=1}^{t} W^{i^T}B^iX$$

$$\ge D^TX - \sum_{i=1}^{t}(W^{iT}B^iW^i)^{1/2}(X^TB^iX)^{1/2}, \quad \text{by Lemma 27.1}$$

$$\ge D^TX - \sum_{i=1}^{t}(X^TB^iX)^{1/2} = \phi(X)$$

Hence Y is a feasible solution of (27.70)

Conversely, let Y_1 be any feasible solution of (27.70), then

$$X^TIY_1 \ge D^TX - \sum_{i=1}^{t}(X^TB^iX)^{1/2} \quad \text{for all } X \ge 0,\ Y_1 \ge 0$$

which means

$$-IX \ge 0 \Rightarrow (D^T - Y_1^T)X \le \sum_{i=1}^{t}(X^TB^iX)^{1/2}$$

By Theorem 27.10, it then follows that there exist

II, W^i, $i = 1, 2, \ldots t$ such that

$$-I\ II + \sum_{i=1}^{t} B^iW^i = D - IY_1$$

$W^{iT}B^iW^i \leq 1,$

$II \geq 0,\ W^i \geq 0,\ i = 1, 2,\ldots t.$

Which implies that

$$-I\,Y_1 + \sum_{i=1}^{t} B^i W^i \geq D$$

$W^{iT}B^iW^i \leq 1,$

$Y_1 \geq 0,\ W^i \geq 0,\ i = 1, 2,\ldots t$

i.e. (Y_1, W^i), i = 1, 2,...t is a feasible solution of the dual problem.

Hence (27.70) and the dual problem are equivalent.

Theorem 27.12: Both primal and dual problems have optimal solutions.

Proof: Since the constraint set of the primal problem (27.66) is bounded, it has an optimal solution and then from Theorem 27.11, it follows that the dual problem (27.67) also has an optimal solution.

The duality relations between (27.66) and (27.67) are thus established.

27.4.3. Solution of the "Subproblem"

The subproblem (27.66) is a convex programming problem, where the objective function is nonlinear which may not be differentiable and in practice it may or may not be possible to solve this problem directly.

If it is not possible to solve the 'subproblem' directly, we first try to obtain a solution of the dual problem which again is a convex programming problem, where the nonlinearity now occurs in the constraints. The functions here are however, differentiable and it may therefore be possible to solve this problem by one of the various methods available for convex programming problems.

Suppose then that (Yo, Wo^i), i = 1, 2,...t is an optimal solution of the dual problem. Then there exists an optimal solution Xo of the primal problem (Theorem 27.11) such that

$$\emptyset\,(Xo) = \psi(Yo)$$

i.e.
$$D^T X_0 - \sum_{i=1}^{t} (X_0^T B^i X_0) = S^T Y_0 \tag{27.71}$$

We also know

$$\phi(X_0) = D^T X_0 - \sum_{i=1}^{t} (X_0^T B^i X_0)^{1/2} \leq D^T X_0 - \sum_{i=1}^{t} (X_0^T B^i X_0)^{1/2} (W_0^{iT} B^i W_0^i)^{1/2}$$

$$\leq D^T X_0 - \sum_{i=1}^{t} w_0^{iT} B^i X_0 \tag{27.72}$$

$$\leq Y^T{}_0 I X_0 \leq S^T Y_0 = \psi(Y_0)$$

and by (27.71) equality must hold throughout in (27.72)

This is, however, true if and only if

$$B^iX_0 \leq \alpha^i B^i W^i_0,\ i = 1, 2, \ldots t$$

$$IY_0 = D - \sum_{i=1}^{t} B^i W_0^i$$

$$IX_0 \leq S$$

$$Y_0^T IX_0 = S^T Y_0$$

We therefore obtain a solution of

$$IX \leq S$$

$$B^iX - \alpha^i B^i W^i_0 = 0,\ i = 1, 2, \ldots t$$

$$Y_0^T IX = S^T Y_0 \tag{29.73}$$

$$X \geq 0$$

By linear programming techniques.

A solution of (27.73) will then yield an optimal solution of the (primal) subproblem.

27.4.4. Solution of the General Problem

In solving our general problem we use Dantzig's method [109] for convex programming.

At each step of iteration an approximate solution is obtained by the simplex method. It is then checked whether the optimal value of the objective function of the corresponding subproblem (of the type (27.66)) is nonnegative.

If it is nonnegative, the approximate solution obtained is the optimal solution of the general problem (Theorem 25.16)

If the optimal value of the objective function is negative, the iteration process is continued and from Theorem 25.17, it follows that for some iteration k, the optimal value of the objective function of the corresponding subproblem tends to zero and the kth approximate solution tends to the optimal solution of the general problem.

27.5. Exercises

1. Consider the problem

$$\begin{aligned} \text{Maximize}\quad & z = 4x_1 + 2x_2 + 3x_3 + c_4x4 \\ \text{Subject to}\quad & x_1 + x_3 + x_4 \leq 24 \\ & 3x_1 + x_2 + 2x_3 + 4x_4 \leq 48 \\ & 2x_1 + 2x_2 + 3x_3 + 2x_4 \leq 16 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

where c_4 is a discrete random variable which takes 4, 5, 6, or 7 with probabilities $\cdot 1, \cdot 2, \cdot 3$ and $\cdot 4$ respectively.

Find a solution that maximizes the expected value of z

2. Consider the problem [262]

$$\begin{array}{ll} \text{Minimize} & x_1 + x_2 \\ \text{Subject to} & ax_1 + x_2 \geq 7 \\ & bx_1 + x_2 \geq 4 \\ & x_1, x_2 \geq 0 \end{array}$$

Where a and b are random variables uniformly distributed between 1 and 4, and 1/3 and 1 respectively

Show that the solution obtained by the expected value solution procedure is infeasible to the original problem with probability '75.

3. Solve the same problem as in Example 27.1 when d is a discrete random variable taking the values 70, 71, 72,...80 each with probability 1/11.

4. Consider the problem of minimizing with respect to x, the function

$$Z(x) = 2x + E(\min_y \; 10\, y)$$

$$\begin{array}{ll} \text{Subject to} & x + y \geq b \\ & x, y \geq 0 \end{array}$$

where b is a random variable having normal distribution with mean 100 and standard deviation 12.

5. A manufacturing firm produces two types of products P_1, P_2 using three machines M_1, M_2 and M_3. The processing times on these machines and the available time per week are given below. The profit per unit of each of the products are random variables whose joint distribution is not known but only their means, variances and the covariance are known.

Type of machines	Time required per unit (minutes)		Available time per week (minutes)
	Product P_1	Product P_2	
M_1	10	5	2500
M_2	4	10	2000
M_3	1	1.5	450

The mean values of profit per unit for Product 1 and Product 2 are \$50 and \$100 respectively and the variance–covariance matrix is given by

$$B = \begin{pmatrix} 400 & 200 \\ 200 & 2500 \end{pmatrix}$$

Formulate the deterministic equivalent of the problem and use the method discussed in section 27.3 to find the quantities of Product 1 and Product 2 to be produced to obtain the maximum profit with probability '9.

6. Consider the problem in Exercise 5 above, where now the machine times required for the two products are independent random variables with known means a_{ij} and variances σ^2_{ij} as given below

Type of machines	Time required per unit (minutes)		Available time per week (minutes)
	Product 1	Product 2	
M_1	$\bar{a}_{11} = 10, \sigma^2_{11} = 36$	$\bar{a}_{12} = 5,\ \sigma^2_{12} = 16$	2500
M_2	$\bar{a}_{21} = 4,\ \sigma^2_{21} = 16$	$\bar{a}_{22} = 10,\ \sigma^2_{22} = 49$	2000
M_3	$\bar{a}_{31} = 1,\ \sigma^2_{31} = 4$	$\bar{a}_{32} = 1.5,\ \sigma^2_{32} = 9$	450

The mean values of profit per unit and the variance–covariance matrix are the same as in the previous problem. Formulate the deterministic equivalent of the problem and use the method discussed in section 27.4 to find the quantities of Product 1 and Product 2 to be manufactured per week to obtain the maximum profit with probability '9.

Chapter 28

Some Special Topics in Mathematical Programming

In this chapter, we discuss some special cases in mathematical programming which frequently arise in real world problems, namely, goal programming, multiobjective programming and fractional programming.

28.1. Goal Programming

Goal programming is a relatively new concept that was conceived by Charnes and Cooper [71, 74] to deal with certain linear programming problems in which multiple conflicting objectives (goals) exist. In the present day business environment, management may no longer be satisfied with profit maximization only but is also concerned with market share, labour stability or other business and social factors. The management therefore sets a desired or acceptable level of achievement (aspiration level, target or goal value) for each objective under consideration. However, it might be impossible to satisfy all goals of management exactly and therefore the management is interested in finding a solution that achieves the goals "as closely as possible". This may be achieved by considering all the goals as constraints and minimizing a suitable function of the sum of the absolute values of the deviations from such goals. To illustrate the concept, we consider the following simple example.

Example. A manufacturing company produces two types of products: A and B from the same raw material. The production of a single unit of A requires 2 units and a single unit of B requires 3 units of the raw material. Total quantity of raw material available in stock is 120 units. The profit per unit of A is \$4 and that of the product B is \$3. The market share of the products are estimated to be 40 units of A and 30 units of B.

The profit maximization problem is then

$$\begin{aligned} \text{Maximize} \quad & 4x_1 + 3x_2 \\ \text{Subject to} \quad & 2x_1 + 3x_2 \leq 120 \\ & x_1 \leq 40 \end{aligned}$$

$$x_2 \leq 30$$
$$x_1, x_2 \geq 0$$

This is a linear-programming problem and can be solved by the usual simplex method.

Suppose now that the management sets the goals:

1. A profit target of $240 should be met.
2. The purchase of the material from the open market should be minimized.

   It is clear that there would be no feasible point that would satisfy all the goals. We therefore try to find a feasible point that achieves the goals as closely as possible. We may therefore reformulate the problem as

$$\text{Minimize} \quad d_1^- + d_2^+$$

$$\begin{aligned} \text{Subject to} \quad & 4x_1+3x_2+d_1^- - d_1^+ = 240 \\ & 2x_1+3x_2+d_2^- - d_2^+ = 120 \\ & x \leq 40 \\ & x \leq 30 \\ & x_1, x_2, d_1^-, d_1^+, d_2^-, d_2^+ \geq 0 \end{aligned} \tag{28.1}$$

Where d_i^- represents the negative deviation from the ith goal, i = 1, 2.

(under-achievement)

d_i^+ epresents the positive deviation from I the goal , i = 1,2.

(over-achievement)

This problem is still in a linear programming form and the usual simplex method can be used to find a solution.

However, goals are rarely of equal importance and even if they are equally important, unit of measurement may be different. Therefore, deviations from these goals are not additive (as in the example above). It is therefore, necessary to derive an equivalent common measure or to use a conversion factor.

There are two basic models in goal programming

(a) The Archimedian model.
(b) The preemptive model.

28.1.1. The Archimedian Goal Programming

In the Archimedian model, weights are assigned to undesirable deviations according to their relative importance and is minimized as an Archimedian sum. This is known as weighted goal programming. (W.G.P.)

Mathematically the problem can be stated as:

$$\text{Minimize} \quad z = \sum_{i=1}^{k} (w_i^- d_i^- + w_i^+ d_i^+)$$

$$\begin{aligned} \text{Subject to} \quad & f_i(X) + d_i^- - d_i^+ = b_i, \; i = 1, 2, \ldots k \\ & f_i(x) \qquad\qquad = b_i, \; i = k+1, \ldots m. \\ & x \geq 0, \; d_i^-, \; d_i^+ \geq 0, \; i = 1, 2, \ldots k \end{aligned} \tag{28.2}$$

where

$f_i(X)$, $i = 1, 2, \ldots m$ are linear functions of $X \in R^n$,
b_i, $i = 1, \ldots m$ are the target values,

d_i^-, d_i^+, $i = 1, 2, \ldots k$ are the deviational variables associated with undesirable deviations from the target values.

w_i^-, w_i^+, $i = 1, 2, \ldots k$ are the positive weights attached to the respective deviations in the achievement function z.

The problem is in a linear programming form and can therefore, be solved by the usual simplex method.

28.1.2. Preemptive Goal Programming

In Preemptive (lexicographic) goal programming, the goals of equal importance are grouped together and priorities are assigned to them. The goals of highest importance are assigned priority level 1, designated as p_1, the goals of the next highest importance are assigned priority level 2, designated as P_2 so that $P_1 >> P_2$ and so on. In general, $Pr >> P_{r+1}$, which means that there exists no real number β such that $\beta P_{r+1} \geq P_r$. Any rigid constraints are assigned priorirty1 and all goals within a given priority must be commensurable.

The mathematical representation of the problem is given as:

$$\text{Lex minimize} \quad a = \{ g_1(d^-, d^+), g_2(d^-, d^+) \ldots g_k(d^-, d^+) \}$$

$$\begin{aligned} \text{Subject to} \quad & \sum_{j=1}^{n} c_{ij} x_j + d_i^- - d_i^+ = b_i, \qquad i = 1, 2, \ldots m. \\ & xj \geq 0, \qquad j = 1, 2, \ldots n \\ & d_i^-, \; d_i^+ \geq 0, \quad i = 1, 2, \ldots m \end{aligned} \tag{28.3}$$

where

x_j, $j = 1, 2 \ldots n$ are the decision variables,

k is the total number of priority levels,

a is the achievement vector; an ordered row vector measure of the attainment of the objectives at each priority level.

d_i^-, d_i^+ are the negative and the positive deviations associated with the ith goal constraint, i = 1, 2,...m.

$g_r(d^-,d^+)$, r = 1, 2...k are linear functions of the deviational variables associated with the goal constraints at priority level r.

c_{ij} is a constant associated with variable j in goal i (a technological coefficient) i = 1, 2...m, j =1, 2...n,

b_i, i = 1, 2,...m are the target values.

One method to solve the above problem (28.3) is to find the sequential solution to a series of conventional linear programming problems. This is accomplished by solving the single objective linear programming problem one for each priority level while maintaining the minimal values reached by all higher priority level minimization.

This is known as sequential goal programming (SLGP) method.

In the first stage, we solve the problem of minimizing the first term in the achievement function subject only to the goals associated with the priority level 1, that is, we solve the problem.

$$\begin{aligned} &\text{Minimize} \quad a_1 = g_1(d^-, d^+) \\ &\text{Subject to} \quad \sum_{j=1}^{n} c_{ij}x_j + d_i^- - d_i^+ = b_i, \; i \in P_1 \\ &\qquad X, d^-, d^+ \geq 0 \end{aligned} \tag{28.4}$$

Let a_1^0 denote the optimal value of a_1 from stage 1.

In the next stage, we are to minimize the second term in the achievement function subject to all goals at priority levels P_1 and P_2 and an extra goal that the achievement at the priority level one must be equal to the optimal achievement a_1^0, already obtained in the first stage minimization.

The problem in the second stage therefore, is

$$\begin{aligned} &\text{Minimize} \quad a_2 = g_2(d^-, d^+) \\ &\text{Subject to} \quad \sum_{j=1}^{n} c_{ij}x_j + d_i^- - d_i^+ = b_i, \; i \in P_1 \\ &\qquad \sum_{j=1}^{n} c_{ij}x_j + d_i^- - d_i^+ = b_i, \; i \in P_2 \\ &\qquad g_1(d^-, d^+) = a_1^0 \\ &\qquad X, d^-, d^+ \geq 0 \end{aligned} \tag{28.5}$$

We continue the process until all priorities have been considered. The optimal solution at the last stage of the process id then the optimal solution to the original goal-programming problem.

It should be noted that as soon as a unique optimal solution is obtained at any stage of the process, it is the optimal solution to the original problem and the process is terminated. This is so because in that case, goals with lower priority levels do not influence the solution obtained at the previous stage.

28.1.3. Multiphase Goal Programming

Another approach for solving the lexicographic goal programming problem was presented by Lee[298]. The method, known as the multiphase (or modified simplex) method is an extension of the well-known two phase method of the simplex algorithm and is an improvement over the sequential goal programming technique as it generally requires fewer computations.

28.1.4. Incommensurability Goal Programming

In general, units of measurements for deviation variables from different goals are incommensurable. Thus, a simple summation of these variables in the objective functions of a W.G.P. or within a priority level of an Lex.G.P. is not valid. To overcome this difficulty a normalization procedure is adopted. Each objective is divided throughout by a constant pertaining to that objective which ensures that all objectives have roughly the same magnitude. Such a constant is known as normalization constant. There are several different normalization constants according to different norms used. For example, L_1-norm. L_2-norm (Euclidean), L_∞-norm (Tehelbycheff) or percentage norm.

28.1.5. Goal Efficiency

Definations 28.1. Let $(\bar{X}, \bar{d})$ be a feasible solution to the goal problem. Then $\bar{X}$ is goal-efficient if there does not exist another feasible point (X, d) such that $d \leq \bar{d}$, $d \neq \bar{d}$, where $\bar{d}$ and d are the corresponding vectors of undesirable deviational variables.

28.1.6. Sensitivity Analysis in Goal Programming

It should be noted that assigning weights in goals programs is often very difficult. Moreover, in reality, most processes and organizations need to change the importance or priorities of their goals with time. Analysing the impact of such changes in the model is therefore, important in the total decision making process.

In Archimedian goal programming, we first solve the problem with a set of reasonable weights and then perform sensitive experiments with other sets of weights to see if a better solution can be obtained.

In preemptive goal programming, we change the order of the priorities and then solve the problem again.

The impact of discrete changes or range variations in other parameters of the linear goal programming model may also be analysed in a manner quite similar to that performed in linear programming.

28.1.7. Interactive Goal Programming

The central theme of interactive goal programming is to actively involve the decision maker in the decision making process. For example in preemptive goal programming, after specifying the priority levels and assigning weights within those priority levels associated with more than one goal, the problem is first solved with the sequential programming procedure. After obtaining a solution, various data and results are presented to the decision maker. The decision maker may then refine the formulation with changes in any goal type, target value, priority level and weight that he considers appropriate. Then, the new problem is solved generating the second stage solution and so forth. If the feasible region is bounded, a goal efficient solution is produced.

28.1.8. Duality and Extensions

The duality in linear goal programming has been discussed by Ignizio [242]. In SLGP approach, since each problem in the sequence is a conventional linear program, there is a corresponding sequence of conventional linear programming duals. In multiphase approach, the dual of a linear goal programming problem is a linear programming problem with multiple, prioritized right hand sides. Ignizio has designated this dual as the multidimensional dual.

A method for solving goal programming problems using fuzzy sets, known as fuzzy programming has been given by Zimmerman [560], Hannan [220]. Extensions of linear goal programming to the integer and nonlinear cases are given by Ignizio [242] and Lee and Marris [299].

28.2. Multiple Objective Linear Programming

After the development of the simplex method by Dantzig for solving linear programming problems, various aspects of single objective mathematical programming have been studied quite extensively. It was however realized that almost every real-life problem involves more than one objective. For such problems, the decision makers are to deal with several objectives conflicting with one another, which are to be optimized simultaneously. For example, in a transportation problem, one might like to minimize the operating cost, minimize the average shipping time, minimize the production cost and maximize its capacity. Similarly, in production planning, the plant manager might be interested in obtaining a production programme which would simultaneously maximize profit, minimize the inventory of the fuished goods, minimize the overtime and minimize the back orders. Several other problems in modern management can also be identified as having multiple conflicting objectives.

Mathematically, a multi-objective linear programme (MOLP) can be stated as:

$$\begin{aligned} &\text{Maximize} \quad z_1 = c^1X \\ &\text{Maximize} \quad z_2 = c^2X \\ &\qquad\vdots \\ &\text{Maximize} \quad z_k = c^kX \end{aligned}$$

Subject to $X \in S$

or Maximize $Z = CX$

Subject to $X \in S$ (28.6)

where $X \in R^n$ is the decision vector.

k is the number of objectives

c^i is the vector of the coefficients of the ith objective function

z_i is the value of the ith objective.

C is a $k \times n$ matrix. The components of the column of CX are the k objectives.

S is the feasible region defined by

$S=\{X \in R^n \mid AX = b, X \geq 0\}$ A being an $m \times n$ matrix of full rank, $m \leq n$.

Z is the vector valued objective function (criterion vector)

The problem (28.6) is also called a linear vector maximization problem. (LVMP)

There are three basic approaches to deal with multi-objective linear programs.

(a) Weighting or utility methods.
(b) Ranking or prioritizing methods.
(c) Efficient solution methods.

28.2.1. Weighting or Utility Methods

In this method each objective is assigned a positive weight and the weighted sum of the k objectives is then maximized over S. The problem is thus converted into a single objective programming problem which can be solved by the conventional simplex method.

Assuming that the weights λ_i are normalized so that $\sum_{i=1}^{k} \lambda_i = 1$, we have the problem

Maximize $\lambda^T CX$

Subject to $X \in S$ (28.7)

Where $\lambda \in R^k, \lambda_i > 0, i = 1, 2 \ldots k, \sum_{i=1}^{k} \lambda_i = 1$

However, the obvious drawback to such an approach is that it is extremely difficult to obtain the necessary weights. To determine the weights, we are to make use of the utility function, which itself is very difficult to construct and if constructed (in general nonlinear) is valid only for one point in time. (see [258])

28.2.2. Ranking or Prioritizing Methods

This approach tries to circumvent the difficulties faced in determining the weights to be assigned to the objective functions. Instead of attempting to find a numerical weight for each objective, the objectives are ranked according to their importance. In fact, ranking is a concept that seems inherent to much of decision making.

We then specify aspiration levels (target values) for the objectives and solve the problem lexicographically (see preemptive goal programming in §28.1.2.)

28.2.3. Efficient Solution Methods

Since in general there does not exist a point in S, which will simultaneously maximize all the objectives in the multi-objective linear program, we can only try to maximize each objective as best as possible. We therefore seek to obtain efficient solutions of the problem in the sense of the following definition.

Definition 28.2. A point $X^0 \in S$ is said to be an efficient point if and only if there is no $X \in S$, such that

$$CX \geq CX^0 \text{ and } CX \neq CX^0$$

that is, a point $X^0 \in S$ is efficient if and only if its criterion vector is not dominated by the criterion vector of some other point in S.

An efficient point is often called a Pareto-Optimal point, an admissible point or a nondominated solution.

A slightly restricted concept of efficiency called the proper efficiency is proposed by Geoffrion[195].

Definition 28.3. An efficient solution X^0 of (26.6) is said to be properly efficient if there exists a scalar $M > 0$ such that for each

$p \in \{1, 2 \ldots k\}$ and each $X \in S$, $C_pX > C_pX^0$, there is at least

one $q \in \{1, 2, \ldots k\}$, $q \neq p$ with $C_qX < C_qX^0$ and

$$\frac{C_pX - C_pX^0}{C_qX^0 - C_qX} \leq M$$

Geoffrion has shown that X^0 is a properly efficient solution of (28.6) if X^0 is an optimal solution of the problem.

Maximize $\lambda^T CX$

Subject to $X \in S$

for some $\lambda \in \Lambda = \left\{ \lambda \in R^k \mid \lambda_i > 0, \sum_{i=1}^{k} \lambda_i = 1 \right\}$

The multi-objective linear program which is sometimes called linear vector maximization problem can therefore be solved by finding the set E of all efficient solutions and then choosing between them on an entirely subjective basis. Finding an initial efficient extreme point is crucial in developing algorithms for enumerating all efficient points. Several algorithms have been developed for finding the set of efficient extreme points of S, see for example, Refs[368, 151, 143, 251]. Efficient faces are then generated and the set E is obtained as the union of the maximal faces.

The method for obtaining the set of all efficient solutions thus proceeds in three phases. In phase 1, an initial efficient extreme point is determined or it is ascertained that the set E of efficient points is empty. If E is not empty, the set of all efficient extreme points is generated in phase II. Finally, in phase III, all maximal efficient faces are determined. The set E is then obtained as union of all these faces.

Before we discuss the algorithms, we first consider the following results.

A well known theorem in multi-objective linear programming is the following.

Theorem 28.1. A point $X^0 \in S$ is efficient if and only if X^0 maximizes the problem

$$\text{Max}\{\lambda^T CX \mid X \in S\} \tag{28.8}$$

for some $\lambda \in \Lambda = \left\{\lambda \in R^k \mid \lambda_i > 0, \sum_{i=1}^{k} \lambda_i = 1\right\}$

For the proof of the theorem, we follow Steuer[441]. We first consider the following lemmas.

Lemma 28.1. (Tucker's Theorem of the Alternative).

Let G, H and K be given matrices of order $p \times n$, $q \times n$ and $r \times n$ respectively with G nonvacuous.

Then either

System I: $GX \geq 0$ $GX \neq 0$, $HX \geq 0$ $KX = 0$ has a solution $X \in R^n$.

or the

System II: $G^T Y_2 + H^T Y_3 + K^T Y_4 = 0$

$Y_2 > 0$, $Y_3 \geq 0$

has a solution $Y_2 \in R^p$, $Y_3 \in R^q$ and $Y_4 \in R^r$

Proof: See Chapter 7.

Lemma 28.2. Let $X^0 \in S$ and D be an $n \times n$ diagonal matrix with

$$d_{jj} = \begin{cases} 1 & \text{if } x_j^0 = 0 \\ 0 & \text{otherwise} \end{cases}$$

Then $X^0 \in E$ if and only if the system

$$CU \geq 0, CU \neq 0, DU \geq 0, AU = 0$$

has no solution $U \in R^n$

Proof: Suppose that U satisfies the system. Let $\bar{X} = X^0 + \alpha U$. Then there exists an $\bar{\alpha} > 0$, such that for all $\alpha \in [0, \bar{\alpha}]$, $\bar{X} \in S$. But $CX - CX^0 = \alpha CU \geq 0$ and $\alpha CU \neq 0$. This implies that X^0 is not efficient.

Conversely, suppose that the system is inconsistent and let X be any point in S. Then for $U = X - X^0$, $AU = 0$ and $DU \geq 0$.Therefore since the system is inconsistent, it is not true that $CU \geq 0$, $CU \neq 0$. It is thus not true that $CX \geq CX^0$, $CX \neq CX^0$. Hence X^0 is an efficient point.

Lemma 28.3. Let $X^0 \in S$ and D be a diagonal matrix as defined in lemma 28.2.Then X^0 is efficient if and only if there exist $\pi \in R^k$. $Y_3 \in R^n$ and $Y_4 \in R^m$ such that

$$c^T\pi + D^TY_3 + A^TY_4 = 0$$
$$\pi > 0,\ Y_3 \geq 0.$$

Proof: Follows from lamma 28.1 and 28.2

Proof of the theorem

Let $X^0 \in S$ be an efficient point. Then by lemma 28.3, there exists $\pi \in R^k$, $Y_3 \in R^n$ and $Y_4 \in R^m$ such that

$$C^T\pi + D^TY_3 + A^TY_4 = 0$$
$$\pi > 0,\ Y_3 \geq 0 \text{ is consistent} \tag{28.9}$$

Now, taking $\alpha = \sum_{i=1}^{k} \pi_i > 0$, the system (28.9)

can be written as

$$(C^T\lambda)\alpha + D^TY_3 + A^TY_4 = 0$$
$$\lambda \in \Lambda,\ Y_3 \geq 0.$$

But by lemma 28.1, the system

$$(\lambda^TC)\,U \geq 0$$
$$(\lambda^TC)\,U \neq 0$$
$$DU \geq 0$$
$$AU = 0$$

has no solution.

Now, for any $X \in S$, $U = X - X^0$, $DU \geq 0$, $AU = 0$. Hence it is not true that $\lambda^TCU \geq 0$, $\lambda^TCU \neq 0$. This implies that $\lambda^TCX \leq \lambda^TCX^0$, which means that X^0 maximizes. (28.8)

Conversely, suppose that X^0 maximizes (28.8) but is not an efficient point. Then there exists an $X \in S$ such that $CX \geq CX^0$, $CX \neq CX^0$. Since $\lambda > 0$, this implies that $\lambda^TCX > \lambda^TCX^0$, which contradicts that X^0 maximizes (28.8).

Corollary 28.1. If S has an efficient point, then at least one extreme point of S is efficient.

Proof : Follows from Theorem 28.1 and the fact that if a linear program has an optimal solution, it has an optimal extreme point.

28.2.4. Finding Efficient Extreme Points

We now discuss methods for finding an initial efficient extreme point and then to generate all efficient extreme points of S.

(a) Evans–Stener Method[151]

Based on Theorem 28.1 an initial efficient extreme point can be obtained by solving the linear program

Max $\{\lambda^T CX \mid X \in S\}$

$$\text{for an arbirary } \lambda \in \Lambda = \left\{\lambda \in R^k \mid \lambda_i > 0, \sum_{i=1}^{k} \lambda_i = 1\right\}. \tag{28.10}$$

If S is nonempty and bounded, an optimal solution is obtained, which by Theorem 28.1 is an efficient extreme point. However, the method is not fail-safe when S is unbounded.

Now, in the sequence of basic feasible solutions (extreme points) generated before an optimal solution of the problem is reached, there may be one or more efficient points. Therefore, a test of efficiency is applied to basic feasible solutions along the way they are generated and time to find an initial efficient extreme point may be reduced.

Test for Efficiency of an extreme point

To determine if a given extreme point of S is efficient we proceed as follows:

Let X be an extreme point of S with associated basis B and let A and C be partitioned into basic and nonbasic parts so that we have

$$\begin{aligned} X_B &= B^{-1}b - B^{-1}NX_N \\ Z &= C_B B^{-1}b + (C_N - C_B B^{-1}N)X_N \\ &= C_B B^{-1}b + WX_N \end{aligned}$$

Where B denotes the basic columns of A and N the nonbasic columns,

X_B, X_N denote the vectors of nonbasic variables respectively,

C_B $(k \times m)$ and C_N $(k \times (n - m))$ denote the submatrices of C corresponding to basic and nonbasic vectors X_B, X_N.

Z is the k-vector of criterion values corresponding to the basic feasible solution X and

$W = C_N - C_B B^{-1}N$ is the $k \times (n - m)$ reduced cost matrix.

Note that B is an optimal basis of (28.10) if and only if

$$\lambda^T W \leq 0 \text{ for } \lambda > 0 \tag{28.11}$$

We call B an efficient basis if and only if B is an optimal basis of the weighted sum linear program (28.10)

Theorem 28.2. Let $\bar{X}$ be an extreme point of S with corresponding basis B. Let

$Q = \{i \mid \bar{X}_{B_i} = 0\}$ and $(B^{-1}N)_i$, $i \in Q$ be the rows of $(B^{-1}N)$ associated with degenerate basic variables. Then $\bar{X}$ is efficient if and only if the subproblem

$$\begin{aligned} &\text{Maximize} && e^T V \\ &\text{Subject to} && -WY + IV = 0 \\ &&& (B^{-1}N)_i\, Y + s_i = 0,\ i \in Q \\ &&& Y \geq 0,\ V \geq 0,\ s_i \geq 0 \text{ for } i \in Q. \end{aligned}$$

where e is the sum vector of ones, $Y \in R^{n-m}$, $V \in R^k$ is consistant bounded with optimal objective value equal to zero.

Proof.[441]: Let A and C be partitioned into basic and nonbasic parts. Then by Lemma 28.2 $\bar{X}$ is efficient if and only if the system

$$\begin{aligned} C_B U_B + C_N U_N &\geq 0 \\ C_B U_B + C_N U_N &\neq 0 \\ DU &\geq 0 \\ BU_B + NU_N &= 0 \end{aligned} \tag{28.13}$$

is inconsistent.

If we let $U_B = -B^{-1}NU_N$ and since for degenerate basic variables $(-B^{-1}N)_i$ $U_N \geq 0$ for $i \in Q$, $\bar{X}$ is efficient if and only if the system

$$\begin{aligned} C_B B^{-1}NU_N - C_N U_N &\leq 0 \\ C_B B^{-1}NU_N - C_N U_N &\neq 0 \\ (B^{-1}N)_i\, U_N &\leq 0, \quad i \in Q \\ U_N &\geq 0 \end{aligned} \tag{28.14}$$

is inconsistent. In other words, $\bar{X}$ is efficient if and only if the system

$$\begin{aligned} -WU_N + IV &= 0 \\ (B^{-1}N)_i U_N &\leq 0, \quad i \in Q \\ U_N &\geq 0 \end{aligned} \tag{28.15}$$

does not have a solution such that $V \geq 0$, $V \neq 0$. Thus the above system enables us to test the efficiency of an extreme point. Hence $\bar{X}$ is efficient if and only if the subproblem (28.12) has a solution with the optimal objective value equal to zero.

We now consider the problem of finding the set of all efficient extreme points. After obtaining an initial efficient extreme point, we introduce a nonbasic variables into the current efficient basis converting one of the basic variables to nonbasic and thus obtain an adjacent extreme point which is then tested for efficiency. This test is conducted for each nonbasic variable in each efficient tableau (that is, the tableau corresponding to an efficient basis) and thus each adjacent extreme point is classified as efficient or nonefficient. If it is efficient, the entering nonbasic variable is called an efficient nonbasic variable. A series of subproblems are therefore solved to enumerate all efficient extreme points and hence involves significant computation.

For testing the efficiency of an extreme point adjacent to a given efficient extreme point, we solve the subproblem (28.16)

Theorem 28.3. Let x^0 be an efficient extreme point of S and x_j be a nonbasic variable with respect to the efficient basis B. Then the adjacent extreme point with x_j a basic variable, is efficient if and only if the subproblem

$$\begin{aligned}
&\text{Maximize} && e^T V \\
&\text{Subject to} && -WY + W^j u + IV = 0 \qquad (28.16)\\
& && Y, V \geq 0,\ u \in R^1
\end{aligned}$$

where W^j is the jth column of the reduced cost matrix $W = C_N - C_B B^{-1} N$ and $Y \in R^{n-m}$, $V \in R^k$

is consistent and bounded with optimal objective value equal to zero. We then call x_j, an efficient nonbasic variable.

Proof: Clearly, the adjacent extreme point is efficient if and only if the problem

$$\begin{aligned}
&\text{Minimize} && 0^T \lambda \\
&\text{Subject to} && W^T \lambda \leq 0 \\
& && (W^j)^T \lambda = 0 \\
& && I\lambda \geq e \qquad (28.17)\\
& && \lambda \geq 0
\end{aligned}$$

has an optimal solution where the value of the objective function is zero.

Since $(W^j)^T\lambda = 0$ implies $(W^j)^T\lambda \geq 0$ and $-(W^j)^T\lambda \geq 0$, the problem reduces to

$$\begin{aligned}
&\text{Minimize} && 0^T \lambda \\
&\text{Subject to} && -W^T \lambda \geq 0 \\
& && (W^j)^T \lambda \geq 0 \qquad (28.18)\\
& && I\lambda \geq e \\
& && \lambda \geq 0
\end{aligned}$$

Its dual is given by

$$\begin{aligned}
&\text{Maximize} && e^T V \\
&\text{Subject to} && -WY + W^j u + IV + It = 0 \qquad (28.19)\\
& && Y \geq 0,\ u \in R^1,\ V \geq 0,\ t \geq 0
\end{aligned}$$

Thus the adjacent extreme point is efficient if and only if the problem (28.19) has an optimal solution with the value of the objective function equal to zero.

The problem (28.19) can be expressed as (28.16) since the slack vector t is not necessary because if there exists a $t_i > 0$,we can increase the value of the objective function by setting $t_i = 0$.

This completes the proof.

Thus we note that since the subproblem is always consistent, the adjacent extreme point is inefficient if and only if the subproblem (28.16) is unbounded.

Echer–Kouda Method

Echer and Kouda[142] provides us with a method for finding an initial efficient point or showing that the set E is empty.

Let $X^0 \in S$ and consider the problem

$$\begin{aligned} &\text{Maximize} && e^T s \\ &\text{Subject to} && CX = Is + CX^0 \\ &&& AX = b \\ &&& 0 \le X \in R^n,\ 0 \le s \le R^k \end{aligned} \tag{28.20}$$

Theorem 28.4. If $(\overline{X}, \overline{s})$ is an optimal solution of (28.20), then $\overline{X}$ is efficient.

Proof: Suppose that $\overline{X} \in E$. Then there exists an $\hat{X} \in S$ such that $c\hat{X} \geq c\overline{X}$, $c\hat{X} \neq c\overline{X}$

Now, $c\overline{X} - cX^0 = I\overline{s}$ and since we can always find a $\hat{s}$ such that $(\hat{X}, \hat{s})$ is feasible for (28.20), $c\hat{X} - cX^0 = I\hat{s}$. Hence $I\hat{s} \geq I\overline{s}$, $I\hat{s} \neq I\overline{s}$ and then $c^T\hat{s} > c^T\overline{s}$, which contradicts that $(\overline{X}, \overline{s})$ is an optimal solution of (28.20). Hence $\overline{X}$ is efficient.

Theorem 28.5. If (28.20) does not have a finite maximum value, then the set E is empty.

Proof: If (28.20) does not have a finite maximum value, then its dual

$$\begin{aligned} &\text{Minimize} && (CX^0)\,P + b^T Y \\ &\text{Subject to} && C^T P + A^T Y \geq 0 \\ &&& -IP \geq e \\ &&& P, Y \text{ unrestricted} \end{aligned} \tag{28.21}$$

is infeasible.

Setting $\pi = -P$ and $Y_4 = -Y$, the dual problem can be written as

$$\begin{aligned} &\text{Minimize} && -(CX^0)^T \pi - b^T Y_4 \\ &\text{Subject to} && C^T \pi + IY_3 + A^T Y_4 = 0 \\ &&& \pi > 0,\ Y_3 \geq 0 \end{aligned} \tag{28.22}$$

If we now suppose that $E \neq \phi$ and $X \in E$, then by lemma 28.3, there exist (π, Y_3, Y_4) such that

$$\begin{aligned} & C^T\pi + D^T Y_3 + A^T Y_4 = 0 \\ & \pi > 0,\ Y_3 \geq 0 \end{aligned} \tag{28.23}$$

Thus there exists a $Y_3 \geq 0$ such that (28.22) has a feasible solution, which contradicts that the dual problem is infeasible. Hence, it is not possible for (28.20) to have an unbounded value of the objective function.

Of course, the efficient point thus obtained need not to be an extreme point. However, it can be shown[140] that if it is not an extreme point, a simple procedure for performing pivots on the optimal tableau for (28.20) will yield an efficient extreme point.

Now, to enumerate all efficient extreme points, Ecker and Kouda[143] proposed to find a solution of the linear system

$$
\begin{aligned}
& W^T V + Y = -W^T e \\
& V \geq 0, Y \geq 0
\end{aligned}
\tag{28.24}
$$

Where W is the reduced cost matrix of an efficient basis and $V \in R^k$, $Y \in R^{n-m}$.

Then X_j is an efficient nonbasic variable if and only if there exists a solution

$$(\bar{V}, \bar{Y}) \text{ of } (28.24) \text{ such that } \bar{y}_j = 0$$

Given an efficient extreme point of (28.6) a routine on simplex tableau of the linear system is applied and Ecker-Kouada's procedure to determine the set of efficient nonbasic indices J an be stated as follows:

Step1: Let L = Set of nonbasic indices = $\{1, 2, \ldots n - m\}$ and J = ϕ. Consider the tableau.

$$
\begin{array}{c|c|c|c|}
 & & V_1 \ldots V_k & y_1 \ldots y_{n-m} \\
\hline
y_1 & & & \\
\vdots & -W^T e & W^T & 1 \\
y_{n-m} & & & \\
\hline
\end{array}
$$

Step 2: For each $j \in L$ such that y_j row in the tableau has a positive left constant and nonpositive nonbasic entries, drop the row and set $L = L - \{j\}$.

Step 3: Perform pivots as necessary on the tableau to obtain a new tableau with a nonnegative constant column.

Step 4: For each $j \in L$ such that y_j is nonbasic or basic with value zero, set $J = J \cup \{j\}$ and $L = L - \{j\}$.

Step 5: For each $j \in L$ such that y_j is currently basic but can be made nonbasic in one pivot, set $J = J \cup \{j\}$ and $L = L - \{j\}$

Step 6: If $L = \phi$, stop. Otherwise, select a $j \in L$ and set $L = L - \{j\}$. Add to the current tableau an objective row to minimize y_j. In the course of minimization, check for steps 2,4 and 5 after each pivot.

Step 7: If y_j has a minimum value of zero, set $J = J \cup \{j\}$ and go to step 6.

In the above procedure it has been assumed that the problem is nondegenerate. For the degenerate case, the above procedure has to be slightly altered (see [143]).

(c) Isermann's Method [251]

Isermann's procedure for finding an initial efficient extreme point of S is a two step procedure.

In the first step, the problem

$$\begin{array}{ll} \text{Minimize} & U^T b \\ \text{Subject to} & U^T A - V^T C \geq O^T \\ & V \geq e \end{array} \tag{28.25}$$

is solved. Isermann then states that the multi-objective linear program (28.6) has an efficient solution if and only if (28.25) has an optimal solution. Next, provided that an optimal solution (U^0, V^0) for (28.25) has been determined, an initial efficient extreme point for (28.6) is obtained by solving the linear program

$$\begin{array}{ll} \text{Maximize} & (V^0)^T CX \\ \text{Subject to} & X \in S \end{array} \tag{28.26}$$

However, Eeker and Hegner [140] have shown by a counter example that, in fact (28.25) may be feasible and unbounded even though (28.6), has an efficient solution. Thus, for certain problems Isermann's method may fail to generate an efficient extreme point solution even though such solutions exist. Benson [48] however establishes that if the set $Q = \{X \in S \mid CX \geq 0\}$ is nonempty, Isermann's procedure is valid. Benson has also proposed a new method for finding an initial efficient extreme point.

Let us now consider the method proposed by Isermann for generating all efficient extreme points of S.

It is clear that, x_j is an efficient nonbasic variable with respect to an efficient basis if and only if there exists a $\lambda \in \Lambda$ such that

$$\begin{array}{l} \lambda^T W \leq 0 \\ \lambda^T W^j = 0 \end{array} \tag{28.27}$$

where W is the reduced cost matrix of the efficient basis and W^j is the j the column of W.

Hence if there is an optimal solution $\bar{\lambda}$ of the problem

$$\begin{array}{ll} \text{Minimize} & e^T \lambda \\ \text{Subject to} & -W^T \lambda \geq 0 \\ & (W^J)^T \lambda = 0 \\ & \lambda \geq e \end{array} \tag{28.28}$$

where W^J is the matrix of columns of W, J denoting the index set of the nonbasic indices, then the basis pertaining to W is an optimal basis for the problem (28.28) corresponding to $\lambda = \bar{\lambda}$ and each nonbasic variable x_j, $j \in J$ is efficient.

Now, the dual to (28.28) can be written as

$$\begin{array}{ll} \text{Maximize} & e^T V \\ \text{Subject to} & -WY + W^J U + IV = e \end{array}$$

$$Y \geq 0, U \geq 0, V \geq 0 \qquad (28.29)$$

where $Y \in R^{n-m}$, $V \in R^k$ and $U \in R^{|J|}$, $|J|$ denoting the number of elements in J.

Thus each non basic variable x_j, $j \in J$ is efficient if and only if the subproblem (28.29) has an optimal solution.

Consider that an efficient extreme point X is available for (28.6). We then start the algorithm by solving the subproblem (28.29) with J = { r }. If the subproblem has an optimal solution, then x_r, is an efficient nonbasic variable. Isermann's method then seeks to enlarge the set J in the following way in order to reduce the number of subproblems to be solved to classify all nonbasic variables.

If y_j is in the optimal basis of the subproblem with J = {r}, then x_j is an efficient nonbasic variable and J is enlarged by each such j. Now, let y_j be a nonbasic variable in the last optimal solution with $J \neq \phi$. We then successively drop the sign restriction on y_j and test if we can introduce y_j into the basis in exchange for a basis variable v_i and still the basis is optimal. If the converted basis is optimal, xj is an efficient nonbasic variable and J is enlarged to JU {j}.

At the end, the index of each nonbasic variable y_j of the last optimal solution for which the reduced cost is zero, is included in J. The resulting index set J is a maximal index set of efficient nonbasic variables.

The algorithm is continued with each efficient extreme point and the corresponding maximal set of efficient nonbasic variables are obtained from which the set of all efficient extreme points is determined. Note that, we may have several maximum index sets for a given efficient extreme point.

28.2.5. Determining the Set of All Efficient Points

The problem which now remains is to determine the set E of all efficient solutions for (28.6). We therefore find all maximal efficient faces of S and the union of these faces will then give the set E.

Several authors have investigated the problem of computing efficient faces, see e.g. Yu and Zeleny [544], Gal [180], Isermann [251] and Ecker, Hegner and Kouada [141]. While Yu–Zeleny and Gal make use of multiparametric linear programming to find all efficient extreme points and then develop their respective algorithms for generating the efficient faces, Isermann's approach identifies the efficient extreme points through a problem dual to the parametric linear program and then by extending his algorithm generates all maximal efficient faces. Ecker, Hegner, Konada's method is similar to Isermann's.

Isermann's Method

Let us assume that S is bounded and let Ix denote the index set of all efficient bases and J^{ij}, the maximal index sets of efficient nonbasic variables at efficient basis i. (as described in section 28.2.4(c)).

To compute all maximal efficient faces, Isermann's procedure is as follows:

For each $i \in Ix$, let T^i denote the set of all indices which belong to at least one of the maximal sets $J^{i,j}$, at efficient basis i, i.e. $T_i = U_j J^{i,j}$ and let g denote the number of indices in T^i.

For each T^i, a directed graph $\Gamma(T^i)$ is formed whose nodes are the $\begin{pmatrix} g \\ h \end{pmatrix}$ combinations of indices in T^i, where h = g, g-1, ...1. The $\begin{pmatrix} g \\ g \end{pmatrix}$ node is the source and the $\begin{pmatrix} g \\ i \end{pmatrix}$ nodes are the sinks. Since each node of $\Gamma(T^i)$ represents a potential maximal index set at efficient basis i, the graph has to be adjusted as one or more maximal sets $J^{i,j}$ have already been determined. In adjusting $\Gamma(T^i)$ for maximal index sets that are known, the respective nodes and all predecessors and successors are deleted. If the adjusted graph $\Gamma(T^i)$ has no node, all maximal index sets of efficient nonbasic variables have been identified. If however, the adjusted graph is nonempty, a sink of the adjusted graph $\Gamma(T^i)$ is selected and the linear program (28.29) is solved for the respective $J^{i,j}$. If (28.29) has no optimal solution, the node corresponding to $J^{i,j}$ and all its predecessors are deleted. If, however (28.29) has an optimal solution, a maximal set $J^{i,j}$ is determined and the respective node, its predecessors and all its successors are deleted.

Let us now form the index sets

$$Q^{i,j} = J^{i,j} \, U D^i \quad \text{for all } i, j$$

where D^i denotes the set of indices of the basic variables at $i \in Ix$.

Now, the same index set may be constructed several times and therefore we form the minimal number of minimal index sets U^α, $(\alpha = 1, 2, \ldots \bar{\alpha})$ that subsume all of the $Q^{i,j}$ and have the property that for each $Q^{i,j}$ there exists an $\alpha \in \{1, 2, \ldots \bar{\alpha})$ such that $Q^{i,j} \subset U^\alpha$ and for each U^α there exists at least one index set $Q^{i,j}$ such that $Q^{i,j} = U^\alpha$ and moreover $U\alpha' \not\subset U\alpha''$ for any $\alpha', \alpha'' \in \{1, 2, \ldots \bar{\alpha}\}$, $(\alpha' \neq \alpha'')$.

Now, for each $\alpha \in \{1, 2, \ldots \bar{\alpha}\}$, we form the index set

$$I_x^\alpha = \{i \in I_X \mid D^i \subset U^\alpha\}$$

Thus, the αth maximal efficient face is characterized by I_x^α.

Then the union of these maximally efficient faces gives the set of all efficient points.

28.3. Fractional Programming

In this section we discuss a special class of nonlinear programming problems where we are concerned with optimizing a ratio of real valued functions over a convex set. Such problems are known as fractional programs. Occasionally, they are also called hyperbolic programs. If the objective functions of the problem is a ratio of two linear (or affine: linear plus a constant) functions and the constraint set is a convex polyhedron, the problem is called a linear fractional program.

Mathematically, the fractional programming problem can be stated as

$$\text{Maximize} \quad F(X) = \frac{f(X)}{g(X)}$$

$$\text{over} \quad X \in S,$$

where S is a compact convex sub-set of R^n and f(X) and g(X) are continuous real valued functions of $X \in S$.

In most cases, it is further assumed that f(X) is nonnegative and concave on S and g(X) is positive and convex on S.

Fractional programs arise in various contexts

(a) The stock cutting problem:
It is often required in paper (or steel) industry to cut the rolls of paper (or steel sheet) from the stock into narrower rolls of specified widths to satisfy the orders. The problem is that cutting should be done in such a way that the amount of wastage is minimum. Linear programming technique can be used to find an optimal cutting pattern, but it may be more appropriate [190] to minimize the ratio of wasted and used amount of raw material and thus a linear fractional program arises.

(b) Investment problem
In investment problems, the firm wants to select a number of projects on which money is to be invested so that the ratio of the profits to the capital invested is maximum subject to the total capital available and other economic requirements which may be assumed to be linear[346, 347]. If the price per unit depends linearly on the output and the capital is a linear function then the problem is reduced to a nonlinear fractional program with a concave quadratic function in the numerator of the objective function.

(c) Stochastic problem
Nonlinear fractional program also appears in stochastic linear programming. Suppose that in the problem of maximizing a profit function say c^TX, subject to linear constraints, the coefficients of the profit function are random variables whose means and variance–covariance matrix are only known. Under the situation, a reasonable deterministic formulation of the problem is to maximize the probability that the profit function attains at least the desired value subject to the linear constraints [see section 27.2]. Then the problem of maximization of return on investment reduces to a nonlinear fractional program of the form Max $\text{Max}\dfrac{P^TX - (X^TBX)^{1/2}}{D^TX}$, subject to $AX \le b$, $X \ge 0$ where B is a positive semidefinite matrix.

Also see [50] and [360].

Applications of fractional programming are also found in inventory problems [231] information theory [7, 343], game theory [409] and in many other areas.

In the recent years various aspects of fractional programming have been studied and a large number of papers on theory and methods of solution have appeared in the literature.

The fractional programming problem (28.30) is said to be a concave–convex fractional program if f is concave and g is convex on S. Concave–convex fractional programs have some important properties in common with concave programs.

It can be shown that (see section 9.3)

1. If f is nonnegative and g is strictly positive then F = f/g is explicitly quasi-concave on S and hence a local maximum is global maximum.
2. If f and g are also differentiable, then F = f/g is pseudoconcave and hence a point satisfying the Kuhn–Tucker optimality conditions is a global maximum.
3. For a linear fractional program, since the objective function is quasiconcave (and quasiconvex), if the feasible region is bounded, the maximum (minimum) is attained at an extreme point of the feasible region.

Based on these properties, it has been possible to solve concave–convex fractional programs by some available techniques in mathematical programming.

28.3.1. Linear Fractional Programming

An example of linear fractional programming was first identified and solved by Isbell and Marlow [249] in 1956. Their algorithm generates a sequence of linear programs whose solutions coverage to the solution of the fractional program in a finite number of iterations. Since then several methods of solutions were developed. Gilmore and Gomory [200] modified the simplex method to obtain a direct solution of the problem. Martos [330] and Swarup [446] have suggested a simplex-line procedure, while by making a transformation of variables, Charnes and Cooper [72] have shown that a solution of the problem can be obtained by solving at most two ordinary linear programs. Algorithms based on the parametric form of the problem have been developed by Jagannathan [253] and Dinkelbach [125].

We now present below the method suggested by Charnes and Cooper for solving a linear fractional program.

The method of Charnes and Cooper

Consider the problem

$$\text{Maximize} \qquad F(X) = \frac{C^T X + \alpha}{D^T X + \beta}$$

$$\begin{aligned} \text{Subject to} \quad & AX \le b \\ & X \ge 0 \end{aligned} \tag{28.31}$$

where C and D are n-vectors, b is an m-vector, A is an mxn matrix and α and β are sealors.

It is assumed that the constraints set

$$S = \{X \mid AX \le b, X \ge 0\} \tag{28.32}$$

is nonempty and bounded and further

$$D^TX + \beta \ne 0 \text{ over } S \tag{28.33}$$

Note that the assumptions (28.33) implies that either $D^TX + \beta > 0$ for all $X \in S$ or $D^TX + \beta < 0$ for all $X \in S$. It cannot have both positive and negative values in the constraints set. If there exists an $X_1 \in S$ for which $D^TX_1 + \beta > 0$ and an $X_2 \in S$ for which $D^TX_2 + \beta < 0$, then for some convex combination X of X_1 and X_2, $D^TX + \beta = 0$ contradicting our assumption.

Suppose $D^TX + \beta > 0$ for every $X \in S$ and

$$\text{let} \qquad Y = tX, \text{ where } t = \frac{1}{D^TX + \beta} > 0 \tag{28.34}$$

The problem (28.31) is then transformed into the following linear program

$$\begin{aligned} \text{Maximize} \quad & G(Y, t) = C^TY + \alpha t \\ \text{Subject to} \quad & AY - bt \le 0 \\ & D^TY + \beta t = 1 \\ & Y \ge 0, t \ge 0 \end{aligned} \tag{28.35}$$

Lemma 28.4. For every (y, t) feasible to the problem (28.35), $t > 0$.

Proof: Suppose, $(\bar{Y}, \bar{t} = 0)$ is feasible to the problem (28.35) Obviously, $\bar{Y} \ne 0$. Let $\bar{X} \in S$. Then $X = \bar{X} + \mu\bar{Y}$ is in S for all $\mu \ge 0$, which implies that the set S is unbounded contradicting our assumptions.

Theorem 28.6. If (Y_0, t_0) is an optimal solution to the problem (28.35), then Yo/ to is an optimal solution to the problem (28.31)

Proof: Let X be any feasible solution to the problem (28.31). Then there exists a $t > 0$, namely $t = \frac{1}{D^TX + \beta}$ such that (tX, t) is feasible to the problem (28.35).

Since (Y_0, t_0) is an optimal solution to (28.35) we have

$$C_T(Y_0, t_0) = C^TY_0 + \alpha t_0 \ge C^T(tX) + \alpha t = t(C^TX + \alpha) = \frac{C^TX + \alpha}{D^TX + \beta} \tag{28.36}$$

Now, since $t_0 > 0$, it is clear that

Y_0/t_0 is feasible to the problem (28.31) and

$$F(Y_0/t_0) = \frac{C^TY_0/t_0 + \alpha}{D^TY_0/t_0 + \beta} = \frac{C^TY_0 + \alpha t_0}{D^TY_0 + \beta t_0} \geq \frac{C^TX + \alpha}{D^TX + \beta} \tag{28.36}$$

which implies that yo/to is an optimal solution to the problem (28.31)

Now, if $D^TX + \beta < 0$ for all $X \in S$, then letting $Y = tx$, $t = -\dfrac{1}{D^TX + \beta} > 0$,

we get the following linear program

$$\begin{aligned}
\text{Maximize} \quad & -C^TY - \alpha t \\
\text{Subject to} \quad & AY - bt \leq 0 \\
& -D^TY - \beta t = 1 \\
& y, t \geq 0
\end{aligned} \tag{28.37}$$

As in the previous case, it can be shown that if (yo, to) is an optimal solution to (28.37), then yo/to is an optimal solution to the fractional programming problem (28.31).

Example

Consider the problem

$$\begin{aligned}
\text{Maximize} \quad & F = \frac{5x_1 + 6x_2}{2x_2 + 7} \\
\text{Subject to} \quad & 2x_1 + 3x_2 \leq 6 \\
& 2x_1 + x_2 \leq 3 \\
& x_1, x_2 \geq 0
\end{aligned}$$

We solve the problem using the method of Charnes and Cooper. Note that the denominator $2x_2 + 7$ is positive over the entire feasible region.

Let $Y = tX$ where $t = \dfrac{1}{2x_2 + 7}$. The equivalent linear program is then given by

$$\begin{aligned}
\text{Maximize} \quad & G = 5y_1 + 6y_2 \\
\text{Subject to} \quad & 2y_1 + 3y_2 - 6t \leq 0 \\
& 2y_1 + y_2 - 3t \leq 0 \\
& 2y_2 + 7t = 1 \\
& y_1, y_2, t \geq 0
\end{aligned}$$

It can be verified that $y_1^0 = \dfrac{3}{40}, y_2^0 = \dfrac{3}{20}, t^0 = \dfrac{1}{10}$

is an optimal solution to the above linear program.

Hence the optimal solution to the fractional program is

$$x_1^0 = \frac{y_1^0}{t^0} = \frac{3}{4},\ x_2^0 = \frac{y_2^0}{t^0} = \frac{3}{2} \text{ and } F^0 = \frac{51}{40}$$

28.3.2. Nonlinear Fractional Programming

As mentioned before, nonlinear fractional programs arise in many applications. Based on the nature of the functions and their properties several algorithms have been suggested for solving nonlinear fractional programs. The interested readers may refer to Jagannathan [253], Swarup [447], Dinkelbach [125], Bector [42], Mangasarian [317], Almogy and Levin [11].

We present below the parametric algorithm of Dinkelbach.

Consider the nonlinear fractional program

$$\text{Maximize} \quad \left\{ \frac{N(X)}{D(X)} \middle| X \in S \right\} \tag{28.38}$$

where S is a closed, bounded and connected subset of R^n and N(X), D(X) are real valued continuous functions of $X \in S$.

Further, it is assumed that $D(X) > 0$ for all $X \in S$.

It will be shown that a solution of the problem can be obtained by solving the following parametric problem associated with (28.38)

$$F(q) = \text{Max } \{N(X) - qD(X) | X \in S\}, \tag{28.39}$$

where $q \in R^1$ is a parameter.

From the assumptions it follows that the problems (28.38) and (28.39) have solutions, if S is nonempty.

Lemma 28.5. $F(q) = \text{Max } \{N(X) - qD(X)\}$ is convex over R^1

Proof: Let $0 \le \lambda \le 1$ and $q' \neq q''$. Then

$$\begin{aligned} F(\lambda q' + (1-\lambda)\, q'') &= \underset{X\in S}{\text{Max}} \{N(X) - (\lambda q' + (1-X)\, q'')\, D(X)\} \\ &= \underset{X\in S}{\text{Max}} \{\lambda(N(X) - q' D(X)) + (1-X)\, N(X) - q''\, D(X)\} \\ &= \lambda \underset{X\in S}{\text{Max}} (N(X) - q' D(X)) + (1-\lambda) \underset{X\in S}{\text{Max}} (N(X) - q''\, D(X)) \\ &= \lambda\, F(q') + (1-\lambda)\, F(q'') \end{aligned}$$

Lemma 28.6. $F(q)'$ is continuous for $q \in R^1$.

Proof: The proof follows from the result that if f is a convex function on a convex set, it is continuous on its interior. (See Theorem 9.10)

Lemma 28.7. F(q) is a strictly decreasing function of $q \in R^1$

Proof: Let $q' < q''$. $q', q'' \in R^1$ and suppose that $X'' \in S$ maximizes $N(X) - q''$ $D(X)$. Then

$$
\begin{aligned}
F(q'') &= \text{Max}\ \{N(X) - q''\, D(X) \mid X \in S\} \\
&= N\,(X'') - q''\, D(X'') \\
&< N(X'') - q'\, D(X'') \\
&= \text{Max}\ \{N(X) - q'\, D(X) \mid X \in S\} \\
&= F(q')
\end{aligned}
$$

Lemma 28.8. F(q) = 0 has an unique solution say qo.

Proof: The proof follows from lemma 28.6 and lemma 28.7 and the fact that

$$\lim_{q \to -\infty} = +\infty \text{ and } \lim_{q \to -\infty} = F(q) = -\infty$$

Lemma 28.9. Let $X^+ \in S$ and $q^+ = \dfrac{N(X^+)}{D(X^+)}$, then

$$F(q^+) \geq 0.$$

Proof: $F(q^+) = \text{Max}\ \{ N(X) - q^+ D(X) \mid X \in S\}$

$\geq N(X^+) - q^+D(X^+) = 0$

For any $q = q^*$, let X^* be an optimal solution to the problem $\text{Max}\{N(X) - q^*D(X) \mid X \in S\}$ and the optimal value be denoted by $F(q^*, X^*)$

The following theorem now establishes a relationship between the nonlinear fractional and the nonlinear parametric programs.

Theorem 28.7

$$q_0 = \frac{N(X_0)}{D(X_0)} = \text{Max}\left\{\frac{N(X)}{D(X)}\middle| X \in S\right\}$$

If and only if

$$F(q_0) = F(q_0, X_0) = \text{Max}\ \{N(X) - q_0D(X)|X \in S\} = 0$$

Proof: Let X_0 be an optimal solution to the nonlinear fractional problem. (28.38)

We then have

$$q_0 = \frac{N(X_0)}{D(X_0)} \geq \frac{N(X)}{D(X)}, \quad \text{for all } X \in S. \tag{28.40}$$

Hence $N(X) - q_0D(X) \leq 0$, for all $X \in S$ (28.41)

and $N(X_0) - q_0D(X_0) = 0$ (28.42)

From (28.41) and (28.42), we have

$$\begin{aligned} F(q_0) &= \text{Max}\{N(X) - q_0D(X) \mid X \in S\} \\ &= N(X_0) - q_0D(X_0) \\ &= F(q_0, X_0) = 0 \end{aligned}$$

Hence X_0 is an optimal solution to the problem (28.39) for $q = q_0$.

To prove the converse, let X_0 be an optimal solution to the problem (28.39) such that, $N(X_0) - q_0D(X_0) = 0$

We then have,

$$N(X) - q_0D(X) \leq N(X_0) - q_0D(X_0) = 0, \text{ for all } X \leq S. \tag{28.43}$$

Hence

$$\frac{N(X_0)}{D(X_0)} \leq q_0, \text{ for all } X \in S \tag{28.44}$$

and

$$\frac{N(X_0)}{D(X_0)} = q_0 \tag{28.45}$$

which implies that X_0 is an optimal solution to the nonlinear fractional program (28.38) and

$$q_0 = \frac{N(X_0)}{D(X_0)} = \text{Max}\left\{\frac{N(X)}{D(X)} \middle| X \in S\right\}$$

It should be noted that X_0 may not be unique. Furthermore, the theorem is still valid, if we replace "max" by "min"

Dinkelbach's Algorithm

Let us now assume that N(X) is concave and D(X) is convex for all $X \in S$ and let S be a convex set. The problem (28.39) is then a problem of maximizing the concave function $N(X) - qD(X)$, $(q \geq 0)$ over the convex set S.

It is further assumed that

$$F(0) = \text{Max}\{N(X) | X \in S\} \geq 0 \tag{28.46}$$

Let X_0 be an optimal solution to the nonlinear fractional program (28.38). Based on the Theorem 28.7, we formulate the problem (28.38) as follows:

Find an $X_m \in S$, such that

$$q(X_0) - q(X_m) < \epsilon, \text{ for any given } \epsilon > 0, \tag{28.47}$$

where $q(X) = \dfrac{N(X)}{D(X)}$

Since F(q) is continuous, we can have an alternative formulations of the problem as

Find an $X_n \in S$ and $q_n = \dfrac{N(X_n)}{D(X_n)}$ such that

$$F(q_n) - F(q_0) = F(q_n) < \delta \text{ for any given } \delta > 0 \tag{28.48}$$

The algorithm can be started with $q = 0$ or by any feasible point $X_1 \in S$ such that $q(X_1) = \dfrac{N(X_1)}{D(X_1)} \geq 0$

The following steps are then followed

Step 1: Set $q_2 = 0$ or $q_2 = \dfrac{N(X_1)}{D(X_1)}, X_1 \in S$ and

proceed to step 2, with $k = 2$.

Step 2: By a suitable method of convex programming find a solution X_k to the problem

$$F(q_k) = \text{Max}\{N(X) - q_k\, D(X) \mid X \in S\}$$

Step 3: If $F(q_k) < \delta$, terminate the process

(a) If $F(q_k) > 0$, then $X_k = X_n$

and (b) If $F(q_k) = 0$, then $X_k = X_0$

X_0 is then an optimal solution and X_n, an approximate optimal solution to the nonlinear fractional program.

Step 4: If $(F(q_k) \geq \delta$, evaluate $q_{k+1} = \dfrac{N(X_k)}{D(X_k)}$

and repeat step 2 replacing q_k by q_{k+1}.

Proof of Convergence: We first prove that $q_{k+1} > q_k$, for all k with $F(q_k) \geq \delta$.

Lemma 28.9 implies that $F(q_k) > 0$

And by the definition of q_{k+1} we have

$$N(X_k) = q_{k+1}\, D(X_k) \tag{28.49}$$

Hence, $0 < F(q_k) = N(X_k) - q_k D(X_k)$

$$= (q_{k+1} - q_k)\, D(X_k) \tag{28.50}$$

and since $D(X_k) > 0$, we have $q_{k+1} > q_k$.

We now prove that

$$\lim_{k\to\infty} q_k = q(X_0) = q_0 \tag{28.51}$$

If this is not true, we must have

$$\lim_{k\to\infty} q_k = q^* < q_0 \tag{28.52}$$

Now, by construction of our procedure, we have a sequence $\{X_k^*\}$ with $\{q_k^*\}$ such that

$$\lim_{k\to\infty} (q_k^*) = F(q^*) = 0 \tag{28.53}$$

Since F(q) is strictly decreasing, we obtain from (28.52) and (28.53), that

$$0 = F(q^*) > F(q_0) = 0$$

which is a contradiction

Hence $\lim_{k\to\infty} q_k = q(X_0) = q_0$

and by continuity of F(q), we get

$$\lim_{k\to\infty} F(q_k) = F(\lim_{k\to\infty} q_k) = F(q_0) = 0$$

This completes the proof.

28.3.3. Duality in Fractional Programming

Several duals for concave–convex fractional programming have been suggested and duality relations proved. For example, see Gold'stein [205], Bector [43], Jagannathan [254], Rani and Kaul [373], Bitran and Megnanti [56], Schaible [396].

Most of them are, however, equivalent.

In contrast to linear programming, a dual to a linear fractional program is not necessarily a linear fractional program and is usually more complicated than the given fractional program. Hence in general, the dual does not give a computational advantage. However, the optimal dual variables are useful to measure the sensitivity of the maximal value of the primal objective function. To make the dual problem computationally more attractive than the primal fractional program, the functions in the primal problem should have certain special structure.

In the case of a linear fractional program, the equivalent problem is a linear program [72]. Hence by dualizing, the equivalent linear program, the dual of a linear fractional program can be obtained as a linear program, which itself is equavalent to a fractional program. Almost all duality approaches in linear fractional programming yield essentially this dual. See Kaska [271], Chadha [65], Kornbluth [285], Sharma and Swarup [413], Kydland [295], Craven and Mond [94]. Other duals that were suggested are nonlinear programs that seem to be less useful.

28.3.4. Other Fractional Programs

There are several applications that give rise to nonlinear fractional programs having different algebric structure of N(x) and D(x) in (28.38) A large number of papers on these special nonlinear fractional programs have appeared in the literature.

Quadratic Fractional Programs

A nonlinear fractional program is called a quadratic fractional program if N(x) and D(x) are quadratic and S is a convex polyhedron.

In 1962, Ritter [377] showed how a method for parametric quadratic programming can be used to solve a quadratic fractional program. With the help of linear transformation of variables Swarup [447] replaced the quadratic fractional program by two nonlinear programs, each with a quadratic objective function subject to linear and one quadratic constraints and showed that an optional solution, if it exists, can be obtained from the solutions of two associated nonlinear programs. Kaska and Pisek [271a] related the quadratic linear fractional program to parametric quadratic program, where the parameter appears in the objective function. Aggarwal and Swarup [6] developed a method for maximizing a ratio of linear functions subject to linear and one quadratic constraints.

Homogenous Fractional Programs

Sinha and Wadhwa [436] considered the problem of maximizing the ratio of concave and convex functions both being homogenous functions of degree one with a constant added to it subject to linear constraints and reduced the problem to a convex programming problem. A solution of the original problem can then be obtained from a solution of the convex programming problem.

Bradley and Frey. Jr [61] generalized the above homogenous fractional programming to the case of maximizing the ratio of nonlinear functions subject to nonlinear constraints, where the constraints are homogenous of degree one and the functions in the objective function are homogenous of degree one with a constraint added to it. Two auxiliary problems are developed and the relations between the solutions of the auxiliary problems and the solutions of the original problem are obtained.

Craven and Mond [95] obtained a dual to the homogenous fractional program with N(X), differentiable concave and D(X) differentiable convex functions both being homogenous of the same degree and the constraints are linear. Aylawadi [19a] generalized the results of Craven and Mond by replacing linear constraints by nonlinear constraints.

Nondifferentiable fractional programs

Aggarwal and Saxena [8a] considered the problem

$$\text{Minimize} \quad F(X) = [f(X) + (X^T BX)^{1/2}]\big/ g(X) \tag{28.54}$$

$$\text{Subject to} \quad h_i(x) \geq 0,\ i = 1, 2, \ldots m.$$
$$X \in S$$

where S is an open convex subset of R^n and $g(X) > 0$. B is a symmetric positive semidefinite matrix, f is a differentiable convex and g, h_i are differentiable concave functions on R^n.

They derived necessary and sufficient optimality conditions for the problem and also obtained duality results.

Singh [420], extended and generalized the results of Aggarwal and Saxena by considering the problem

$$\text{Minimize} \quad F(X) = \left[f(X) + \sum_{i=1}^{t} (X^T B_i X)^{1/2} \right] / g(X)$$

$$\begin{aligned} \text{Subject to} \quad & h_i(X) \geq 0, \ i = 1, 2, \ldots m \\ & K_i(X) = 0, \ i = m+1, \ldots m+p. \\ & X \in S \end{aligned} \qquad (28.55)$$

where S is an open convex subset of R^n, B_i are symmetric positive semidefinite matrices, all functions involved in the problem are differentiable and members of Hanson–Mond [224, 225] classes of functions and $g(X) > 0$. and also established the duality results.

Chandra and Gulati [66] obtained a dual and proved the duality theorems for the nondifferentiable fractional programming problem

$$\text{Minimize} \quad F(X) = \frac{C^T X - (X^T B X)^{1/2} - \alpha}{E^T X + (X^T D X)^{1/2} + \beta}$$

$$\begin{aligned} \text{Subject to} \quad & AX \leq b \\ & X \geq 0 \end{aligned} \qquad (28.56)$$

where A is a mxn matrix, C, E, X are n-vectors, B and D are nxn positive semi-definite matrices and α, β are sealars.

It is assumed that $E^T X + (X^T D X)^{1/2} + \beta > 0$ for all X satisfying the constraints.

Sinha and Aylawadi [427] derived necessary and sufficient optimality conditions for a class of nonlinear fractional programming problems where the objective functions and one or more of the constraint functions are nondifferentiable. Also, see [428]

Generalized Fractional Programs

In many applications, we get a fractional program where the problem is to maximize a finite sum of ratios on a convex set, that is the problem is

$$\text{Maximize} \quad \sum_{i=1}^{m} \frac{f_i(X)}{g_i(X)} \qquad (29.57)$$

$$\text{Subject to} \quad X \in S,$$

where S is a convex subset of R^n.

Such a problem is referred to as a generalized fractional program.

The generalized fractional program has applications in inventory models, economics, statistics and management science. Almogy and Livus [10] considered a multi-stage stochastic shipping problem and formulated its deterministic equivalent as a sum of ratios problem. Hodgson and Lowe [231] developed a model for simultaneously minimizing the set-up cost, inventory holding cost and material handling cost. This gives rise to a fractional program (28.57). Also, see [406] and references therein.

In 1970, Wadhwa [491] gave a method of solution for the problem of maximizing the sum of two ratios of linear functions subject to linear constraints. The method consists in considering a closely related parametric linear programming problem, a solution of which leads to a solution of the original problem.

Consider the problem

$$\text{Maximize} \quad F(X) = \frac{C_1^T X + \alpha_1}{D_1^T X + \beta_1} + \frac{C_2^T X + \alpha_2}{D_2^T X + \beta_2} \tag{28.58}$$

$$\text{Subject to} \quad AX \le b$$

$$X \ge 0$$

where $X \in R^n$, A is an mxn matrix, b an m-vector and C_1, C_2, D_1, D_2 are n-vectors and α_1, α_2, β_1, β_2 are scalors.

It is assumed that the constraints set

$$S = \{X \mid AX \le b, X \ge 0\}$$

is nonempty and bounded and that for every $X \in S$,

$$D_1^T X + \beta_1 > 0, \; D_2^T X + \beta_2 > 0.$$

The outline of the proposed method of solution is as follows. A programming problem closely related to the original problem is first considered, where if one restricts the value of one particular variable to a fixed value θ and then treat θ as a parameter, the problem essentially becomes a linear programming problem with a parameter θ in the technology matrix. The optimal solution of this problem is then a function of θ and the objective function f of the related problem can be expressed in terms of this function and θ. The maximum of f as a function of one variable θ is formed from which the solution of the original problem is easily obtained.

Consider the problem

$$\text{Maximize} \quad f(X, t, u, v) = C_1^T X + \alpha_1 t + v$$

$$\begin{aligned} \text{Subject to} \quad & AX - bt \le 0 \\ & D_1^T X + \beta_1 t = 1 \\ & D_2^T X + \beta_2 t - u = 0 \\ & -C_2^T X + \alpha_2 t + uv = 0 \\ & X, t, u \ge 0 \end{aligned} \tag{28.59}$$

Theorem 28.8. Every (X, t, u, v) feasible to (28.59) has t, u positive.

Proof : Suppose $(\bar{X}, \bar{t} = 0, \bar{u}, \bar{v})$ is feasible to (28.59). Then $\bar{X} \neq 0$, otherwise the second constraint in (28.59) will not be satisfied. Let $\hat{X} \in S$. Then $A\bar{X} \le 0, \bar{X} \ge 0$,

$$X^* = \hat{X} + \lambda \bar{X} \in S, \text{ for all } \lambda \ge 0$$

which implies that S is unbounded contradicting the assumption. Hence $\bar{t} > 0$.

Again, since $(\bar{X}/\bar{t}) \in S, D_2^T\bar{X} + \beta_2\bar{t} > 0$, $\bar{u}$ must be positive.

Theorem 28.9. If x_0 is an optimal solution of the problem (28.58), then there exist $t_0 > 0$ and u_0, v_0 so that (t_0x_0, t_0, u_0, v_0) is an optimal solution of the problem (28.59).

Proof : Sionce $D_1^T X_0 + \beta_1 > 0$, $D_2^T X_0 + \beta_2 > 0$, there exist $t_0 = \dfrac{1}{D_1^T X_0 \beta} > 0$ and u_0, v_0 so that (t_0X_0, t_0, u_0, v_0) is feasible for the problem (28.59). Let (X_1, t_1, u_1, v_1) be any feasible solution for (28.59). By Theorem 28.8, $t_1 > 0$ and hence X_1/t_1 is feasible for the problem (28.58). We therefore, have

$$\text{Max } F(X) = \frac{C_1^T X_0 + \alpha_1}{D_1^T X_0 + \beta_1} + \frac{C_2^T X_0 + \alpha_2}{D_2^T X_0 + \beta_2} \geq \frac{C_1^T X_1/t_1 + \alpha_1}{D_1^T X_1/t_1 + \beta_1} + \frac{C_2^T X_1/t_1 + \alpha_2}{D_2^T X_1/t_1 + \beta_2}$$

or
$$\frac{t_0\left(C_1^T X_0 + \alpha_1\right)}{t_0\left(D_1^T X_0 + \beta_1\right)} + \frac{t_0\left(C_2^T X_0 + \alpha_2\right)}{t_0\left(D_2^T X_0 + \beta_2\right)} \geq \frac{C_1^T X_1 + \alpha_1 t_1}{D_1^T X_1 + \beta_1 t_1} + \frac{C_2^T X_1 + \alpha_2 t_1}{D_2^T X_1 + \beta_2 t_1}$$

or
$$t_0 (C_1^T X_0 + \alpha_1) + \frac{u_0 v_0}{u_0} \geq C_1^T X_1 + \alpha_1 t_1 + \frac{u_1 v_1}{u_1}$$

or
$$t_0 (C_1^T X_0 + \alpha_1) + v_0 \geq C_1^T X_1 + \alpha_1 t_1 + v_1$$

or
$$f(t_0X_0, t_0, u_0, v_0) \geq f(X_1, t_1, u_1, v_1)$$

which proves the theorem.

Theorem 28.10. If (X_1, t_1, u_1, v_1) is an optimal solution of the problem (28.59), then X_1/t_1 is optimal for the problem (28.58)

Proof : Suppose X^* is any feasible solution of the problem (28.58).

Then, by assumption $D_1^T X^* + \beta_1 > 0$, $D_2^T X^* + \beta_2 > 0$ and let

$u^* = D_2^T X^* + \beta_2$, $v^* = C_2^T X^* + \alpha_2$ and $\hat{t} = \dfrac{1}{D_1^T X^* + \beta_1}$. We now define

$\hat{X} = \hat{t}X^*, \hat{u} = \hat{t}u^*, \hat{v} = \dfrac{v^*}{u^*}$, so that $(\hat{X}, \hat{t}, \hat{u}, \hat{v})$ is feasible for the problem (28.59)

Hence $\quad f(X_1, t_1, u_1, v_1) \geq f(\hat{X}, \hat{t}, \hat{u}, \hat{v})$

i.e.
$$C_1^T X_1 + \alpha_1 t_1 + v_1 \geq C_1^T \hat{X} + \alpha_1 \hat{t} + \hat{v}. \tag{28.60}$$

By Theorem 28.8, $t_1 > 0$ and hence X_1/t_1 is feasible for (28.58) and $F(X_1/t_1)$ $$= \frac{C_1^T(X_1/t_1+\alpha_1}{D_1^T(X_1/t_1)+\beta_1} + \frac{C_2^T(X_1/t_1+\alpha_2}{D_2^T(X_1/t_1)+\beta_2}$$

$$= C_1^T X_1 + \alpha_1 t_1 + \frac{C_2^T X_1 + \alpha_2 t_1}{D_2^T X_1 + \beta_2 t_1}$$

$$= C_1^T X_1 + \alpha_1 t_1 + v_1$$

$$\geq C_1^T \hat{X} + \alpha_1 \hat{t} + \hat{v}, \qquad \text{by (28.60)}$$

$$= \hat{t}(C_1^T X^* + \alpha_1) + \frac{v^*}{u^*}$$

$$= \frac{C_1^T X^* + \alpha_1}{D_1^T X^* + \beta_1} + \frac{C_2^T X^* + \alpha_2}{D_2^T X^* + \beta_2}$$

Hence X_1/t_1 is optimal for the problem (28.58).

In problem (28.59), the value of v is now reatricted to a fixed value θ and then θ is treated as a parameter in the linear programming problem

$$\begin{aligned}
\text{Maximize} \quad & z = C_1^T X + \alpha_1 t \\
\text{Subject to} \quad & Ax - bt \leq 0 \\
& D_1^T X + \beta_1 t = 1 \\
& D_2^T X + \beta_2 t - u = 0 \\
& -C_2^T X - \alpha_2 t + \theta u = 0 \\
& X, t, u \geq 0
\end{aligned} \tag{28.61}$$

It is clear that the parameter θ varies over the range $(\theta_2; \theta_m)$ where

$$\theta_L = \text{Min}\left[F_2(X) = \frac{C_2^T X + \alpha_2}{D_2^T X + \beta_2} \middle| x \in S \right] \tag{28.62}$$

and $$\theta_M = \text{Max}\left[F_2(X) = \frac{C_2^T X + \alpha_2}{D_2^T X + \beta_2} \middle| x \in S \right] \tag{28.63}$$

The problems (28.62) and (28.63) are linear fractional programming problems and can be solved by the simplex method as illustrated in [72].

To begin with θ is given the value θ_M and the following problem is solved

$$\text{Maximize} \quad z - C_1^T X - \alpha_1 t = 0$$

$$\text{Subject to} \quad AX - bt \leq 0 \tag{28.64}$$
$$D_1^T X + \beta_1 t = 1$$
$$-D_2^T X + \beta_2 t + u = 0$$
$$-C_2^T X - \alpha_2 t + \theta u = 0$$
$$X, t, u \geq 0$$

where the solution of (28.63) provides an initial basic feasible solution to the problem.

Note that, since by Theorem 28.8 every feasible solution of (28.64) must have t and u positive, they must always remain in the basis.

The range of θ : (θ_1, θ_M) for which the optimal basis remains optimal is determined and continuing the process of parametric technique, the optimal basis for each subset of the internal (θ_2, θ_m) is obtained. The optimum z for each subset of the interval is a function of a single variable θ and can be expressed as

$$z_i = g_i(\theta), \text{ for } \theta_{i+1} \leq \theta \leq \theta_i,\ i = 0, 1, ...k. \tag{28.65}$$

where $\theta_0 = \theta_M$ and $\theta_{k+1} = \theta_L$

The objective function of the problem (28.59) is then given by

$$f_i(\theta) = g_i(\theta), + \theta, \text{ for } \theta_{i+1} \leq \theta \leq \theta_i,\ i = 0, 1, ...k. \tag{28.66}$$

which are functions of a single variable θ. For each interval, the maximum of f is obtained either by simple inspection or by the method of differential calculus. Suppose that $\theta = \theta_i^*$ maximizes $f_i(\theta)$ in the interval (θ_{i+1}, θ_i). The optimal value of the objective function of the problem (28.59) is then given by

$$\text{Max } f = \underset{i}{\text{Max}} \, [f_i\,(\theta_i^*)] \tag{28.67}$$

If this maximum is obtained for $i = i_0$, then the optimal basis of the problem (28.64) for $\theta_{i_0+1} \leq \theta \leq \theta_{i_0}$ will also be optimal for (28.59) and an optimal solution X_0, t_0, u_0 and $v_0 = \theta_{i_0}{}^*$ of the problem is easily obtained. By Theorem 28.10, X_0/t_0 is optimal solution of the original problem (28.58).

Almogy and Levin [11] considered fractional programs of maximizing a sum of linear or concave–convex fractional functions on closed and bounded polyhedral sets and have shown that, under certain assumptions, problems of this type can be transformed into equivalent ones of maximizing multi-parameter linear or concave functions subject to additional feasibility constraints.

Cambini, Martein and Schaible [63] considered the problem of maximizing the sum of m concave–convex fractional functions on a convex set and had shown that this problem is equivalent to the one whose objective function is the sum of m linear fractional functions defined on a suitable convex set. Successively, using the Charnes–Cooper transformation, the objective function is transformed into the sum of one linear function and (m – 1) linear fractional functional functions. As a special case, the problem of maximizing the sum of two linear fractional functions subject to linear constraints is considered. Theoretical properties are studied and an algorithm converging in a finite number of iterations is proposed.

In recent years, a number of other solution methods have also been proposed [49, 136, 153, 176, 281, 293]

Jokseh [259] considered the fractional programming problems: Maximize (or Minimize)

$$\text{(i)} \quad F(X) = f_1(X) + \frac{1}{f_2(X)}$$

$$\text{or} \quad \text{(ii)} \quad F(x) = \frac{1}{f_1(X)} + \frac{1}{f_2(X)}$$

$$\text{subject to} \quad AX \leq b \qquad (28.68)$$
$$X \geq 0$$

where $f_1(X)$, $f_2(X)$ are linear functions of $X \in R^n$, A is an mxn matrix and b, an m-vector.

Jokseh reduced the problems to parametric linear programs by treating the value of one of the linear forms in the objective function as a parameter, He also derived conditions for extrema and discussed the possibilities for local and global extrema.

For the problem

$$\text{Maximize} \quad F(X) = c^TX + \frac{d^TX}{h^TX} \qquad (28.69)$$

$$\text{Subject to} \quad X \in S$$

where S is a convex subset of R^n c, d, h are n-vector and $h^TX > 0$.

Schaible [399] investigated F(X) in terms of quasi-concavity and quasi-convexity to get some insight into the nature of local optima of the problem. Under some conditions the objective function F(X) can be written as a quotient of a concave and convex functions, which can be related to a convex program [395a] by an extension of Charnes–Cooper's variable transformation [72]. Then duality relations are obtainable for (28.69).

Multi-objective fractional programs

Not much work has been done on algorithms for multi-objective linear fractional programming problems. Consider the problem

$$\text{Maximize} \quad \{F_1(X), F_2(X), \ldots F_k(X)\} \qquad (28.70)$$
$$\text{Subject to} \quad X \in S = \{X \in R^n \mid AX = b, X \geq 0, b \in R^m\}$$

where Fi is a linear fractional $\frac{C_i^TX + \alpha_i}{D_i^TX + \beta_i}$, for i = 1, 2,...k and for all $X \in S$, the denominators are positive.

The only algorithm known, for this problem is given by Kornbluth and Steuer [287]. It computes all weak efficient vertices provided that S is bounded and the set of all weak efficient points is determined, A point $\bar{X} \in S$ is said to be weak efficient if and only if there dos not exist another $X \in S$, such that $F_i(X) > F_i(\bar{X})$ for all i.

Schaible [404] discussed some results in multi-objective fractional programming with regard to the connectedness of the set of efficient solutions. For optimality conditions and duality of multi-objective fractional programs, see [421].

28.3.5. Indefinite Quadratic Programming

In this section, we consider the problem of maximizing a quadratic function that can be factored into two linear functions subject to linear constraints. Such a problem may arise in case there is a competitive market with respect to two competitors for a given product whose total demand is constant, the sale of the product depends linearly on the market prices and the problem is to maximize profit for one of the competitors [449]. Necessary and sufficient conditions for expressing a quadratic function as a product of two linear functions plus an additive constant have been derived by Schaible [393].

Consider the problem

$$\begin{array}{ll} \text{Maximize} & F(X) = (C^TX + \alpha)(D^TX + \beta) \\ \text{Subject to} & AX \leq b \\ & X \geq 0 \end{array} \tag{28.71}$$

where A is an mxn matrix, b an m-vector and X, C, D are n-vectors and α, β are scalars. It is assumed that the constraint set $S = \{X \mid AX \leq b, X \geq 0\}$ is nonempty and bounded.

Several authors investigated the above problem and proposed different methods to find its solution.

In 1996, Swarup reduced the problem to a parametric linear programming problem [451] and assuming that both $C^TX + \alpha$ and $D^TX + \beta$ are positive for all $X \in S$, Swarup [452] developed a simplex-type algorithm to solve the problem. Swarup [450] had also shown that a solution of the problem can be obtained from a solution of the convex programming problem

$$\begin{array}{ll} \text{Minimize} & \dfrac{t^2}{C^TX + \alpha t} \\ \text{Subject to} & AX - bt \leq 0 \\ & D^TX + \beta t = 1 \\ & X, t \geq 0 \end{array} \tag{28.72}$$

With the assumption that for every $X \in S$, $D^TX + \beta > 0$ and that $\max \{F(X) \mid X \in S\}$ is positive, Sinha and Lal [433] showed that a solution of the problem (28.71) can be obtained from a solution of a convex programming problem where the objective function is linear and the constraints are linear and one convex quadratic function.

Sinha and Wadhwa [436] considered a generalization of the problem (28.71)

$$\begin{aligned} \text{Maximize} \quad & F(X) = [f(X) + \alpha]\,[g(X) + \beta] \\ \text{Subject to} \quad & AX \le b \\ & X \ge 0 \end{aligned} \tag{28.73}$$

where A, b, α and β are the same as in (28.71) as also is the assumption that the set $S = \{X | AX \le b, X \ge 0\}$ is nonempty and bounded. It is assumed that f(X) and g(X) are concave homogenous functions of degree one and further at an optimal solution of (28.73) $f(X) + \alpha$ and $g(X) + \beta$ are both positive.

It had been shown that if $(\bar{X}, \bar{t})$ is an optimal solution of the problem

$$\begin{aligned} \text{Maximize} \quad & f(X) + \alpha t \\ \text{Subject to} \quad & AX - bt \le 0 \\ & g(X) + \beta t - t^2 \ge 0 \\ & X, t \ge 0 \end{aligned} \tag{28.74}$$

then $(\bar{X}/\bar{t})$ is an optimal solution for the problem (28.73)

28.4. Exercises

1. A textile company produces two types of materials A: a strong upholstery material and B: a regular dress material. The material A is produced according to direct orders from furniture manufacturers and the material B is distributed to retail fabric stores. The average production rates for the material A and B are identical, 1000 metres per hour. By running two shifts the operational capacity is 80 hours per week.

   The material department reports that the maximum estimated sales for the following week is 70,000 metres of material A and 45,000 metres of material B. According to the accounting department the profit form a metre of material A is \$ 2.50 and from a metre of material B is \$ 1.50.

   The management of the company believes that a good employer-employee relationship is an important factor for business success. Hence, the management decides that a stable employment level is a primary goal for the firm. Therefore, whenever there is demand exceeding normal production capacity, the management simply expands production capacity by providing overtime. However, the management feels that overtime operation of the plant of more than 10 hours per week should be avoided because of the accelerating costs. The management has the following goals in order of their importance.

   (i) Avoid any under-utilization of production capacity.
   (ii) Limit the overtime operation of the plant to 10 hours per week.

(iii) Achieve the sales goals of 70,000 metres of material A and 45,000 metres of material B.
(iv) Minimize the overtime operation of the plant as much as possible.

Formulate and solve this problem as a goal-programming problem.

2. ABC furnitures produce three products: tables, desks and chairs. All furniture are produced in the central plant. Production of a desk requires 3 hours in the plant, a table takes 2 hours and a chair requires only 1 hour. The normal plant capacity is 40 hours a week. According to marketing department, the maximum number of desks, tables and chairs that can be sold are 10, 10 and 12, respectively. The president of the firm has set the following goals arranged according to their importance.

   (i) Avoid any under utilization of production capacity.
   (ii) Meet the order of XYZ store for seven desks and five chairs.
   (iii) Avoid the overtime operation of the plant beyond 10 hours.
   (iv) Achieve the sales goals of 10 desks, 10 tables and 12 chairs.
   (v) Minimize the overtime operation as much as possible.

   Formulate and solve the given problem as a goal programming problem.

3. A company manufactures two products A and B each of which requires processing on two machines M_1 and M_2. The normal machine time available on both the machines M_1 and M_2 is 40 hours per week. The machine requirements in hours for a unit of each product are given in the table below

Product	Machine time	
	M_1	M_2
A	1	1.5
B	2	1.5

   The profit per unit of product A is \$80 and the profit per unit of product B is \$100.

   The company sets the following goals ranked in order of priority

   (i) Overtime is not to be allowed.
   (ii) Meet a production quota of 100 units of each product per week.
   (iii) Maximize profit.

   Formulate and solve the problem as a goal programming problem.

4. An electronic firm produces two types of television sets: color and black-and-white. According to the past experience, production of either a color or a black-and-white set requires an average of one hour in the plant. The plant has a normal production capacity of 40 hours a week. The marketing department reports that because of limited market, the maximum numbers

color and black-and-white sets that can be sold in a week are 24 and 30, respectively. The net profit from sale of a colour set is \$80, whereas it is \$40 from a black-and-white set.

The manager of the firm has set the following goals arranged in the order of importance

P1: Avoid any under utilization of normal production capacity.
P2: Sell as many television sets as possible. Since the net profit from the sale of a color television set is twice the amount from a black-and-white set, achieve twice as much sales for color sets as for black-and-white sets.
P3: Minimize the overtime operation of the plant as much as possible.

Formulate the given problem as a goal-programming problem and solve it.

5. Using the sequential linear goal programming method, solve the problem

Lex. minimize $a = [(d_1^+ + d_2^+), d_3^+, d_4^-]$

Subject to
$$\begin{aligned} &x_1 + x_2 + d_1^- - d_1^+ = 10 \\ &x_1 + 2x_2 + d_2^- - d_2^+ = 12 \\ &4x_1 + x_2 + d_3^- - d_3^+ = 4 \\ &x_2 + d_4^- - d_4^+ = 7 \\ &x_j \geq 0, j = 1, 2 \\ &d_i^+, d_i^- \geq 0, i = 1, 2 \end{aligned}$$

6. Using Ecker–Kauda method find an efficient point for the multi-objective linear program.

Maximize $z_1 = 3x_1 - x_2$

Maximize $z_2 = x_1 + 2x_2$

Subject to
$$\begin{aligned} &2x_1 - x_2 \leq 2 \\ &x_1 \leq 2 \\ &x_1, x_2 \geq 0. \end{aligned}$$

7. [544] Find all efficient extreme points and all maximal efficient faces for the multi-objective linear program.

Maximize
$$Z = \begin{pmatrix} 4 & 1 & 2 \\ 1 & 3 & -1 \\ -1 & 1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Subject to
$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \leq \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}$$

$$x_1, x_2, x_3 \geq 0$$

8. Find all efficient extreme points and maximal efficient faces of the problem.

$$\text{Maximize} \quad Z = \begin{pmatrix} 2 & 1 & 2 & -1 & -2 \\ 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & -2 & 0 & 2 \\ -2 & 2 & 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$$

$$\text{Subject to} \quad \begin{pmatrix} 0 & 4 & 0 & 3 & 3 \\ 2 & 4 & 0 & 4 & 1 \\ 4 & 0 & 2 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \le \begin{pmatrix} 27 \\ 40 \\ 38 \\ 24 \\ 27 \end{pmatrix}$$

$$x_j \ge 0, \quad j = 1, 2, 3, 4, 5.$$

9. Solve the following fractional programming problems

i) Maximize $\dfrac{3x_1 + 2x_2}{x_1 + x_2 + 7}$

Subject to
$3x_1 + 4x_2 \le 12$
$5x_1 + 3x_2 \le 15$
$x_1, x_2 \ge 0$

ii) Maximize $\dfrac{3x_1 + 2x_2}{7x_1 + 5x_2 + 4}$

Subject to
$x_1 + 3x_2 \le 4$
$x_1 + x_2 \le 1$
$x_1, x_2 \ge 0$

iii) Minimize $\dfrac{-5x_1 - 6x_2}{2x_2 + 7}$

Subject to
$2x_1 + 3x_2 \le 6$
$2x_1 + x_2 \le 3$
$x_1, x_2 \ge 0$

iv) Maximize $\dfrac{2x_1 + x_2}{2x_1 + 3x_2 + 1}$

Subject to
$-x_1 + x_2 \le 1$
$x_1 + x_2 \le 2$
$x_1, x_2 \ge 0$

10. Let $F(X) = \dfrac{C^TX+\alpha}{D^TX+\beta}$ and let s be a convex set such that $D^TX + \beta > 0$ over S. Then show that F is both pseudoconvex and pseudoconcave over S.

11. [125] solve the following nonlinear fractional program

$$\text{Maximize} \quad \frac{-3x_1^2 - 2x_2^2 + 4x_1 + 8x_2 - 8}{x_1^2 + x_2^2 - 6x_2 + 8}$$

$$\begin{aligned} \text{Subject to} \quad & x_1 + 3x_2 \le 5 \\ & x_1, x_2 \ge 0 \end{aligned}$$

12. [435]. Consider the problem

$$\text{Minimize } F(X) = C^TX + \alpha + \frac{1}{D^TX+\beta}$$

$$\begin{aligned} \text{Subject to} \quad & AX \le b \\ & X \ge 0, \end{aligned}$$

where $X \in R^n$, C, D are n-vectors, b an m-vector and the constraint set is nonempty and bounded. Further, for every feasible solution x, $C^TX + \alpha \ge 0$, $D^TX + \beta > 0$

Show that a solution of the problem can be obtained from a solution of

$$\text{Minimize} \quad G(X,t) = \frac{C^TX + \alpha t + 1}{t}$$

$$\begin{aligned} \text{Subject to} \quad & AX - bt \le 0 \\ & D^TX - \beta t \ge t^2 \\ & X, t \ge 0 \end{aligned}$$

13. Let fi(X), i = 1, 2,...m are differentiable concave functions on a convex set $S \subset R^n$ and are positive. Prove that $\underset{i=1}{\overset{m}{\pi}}$ fi(X) is pseudoconcave on S.

14. [436] Show that a solution of the problem (28.73) can be obtained from a solution of the problem

$$\begin{aligned} \text{Maximize} \quad & f(x) + \alpha t \\ \text{Subject to} \quad & Ax - bt \le 0 \\ & g(x) + \beta t \ge t^2 \\ & x, t \ge 0 \end{aligned}$$

15. Solve the problem

$$\begin{aligned} \text{Minimize} \quad & (2x_1 + 3x_2 + 2)(5 - x_2) \\ \text{Subject to} \quad & x_1 + x_2 \le 1 \\ & 4x_1 + x_2 \ge 2 \\ & x_1, x_2 \ge 0 \end{aligned}$$

CHAPTER 29

Dynamic Programming

29.1. Introduction

Dynamic programming is a mathematical technique concerned with the optimization of multi-stage decision processes. The technique was developed in the early 1950s by Richard Bellman [44], who also coined its name "Dynamic Programming". The name might suggest that dynamic programming refers to problems in which changes overtime were important. However, the technique can be applied to problems in which time is no way relevant.

In this technique, the problem is divided into small subproblems(stages) which are then solved successively and thus forming a sequence of decisions which leads to an optimal solution of the problem.

Unlike linear programming, there is no standard mathematical formulation of the dynamic programming problem. Rather, dynamic programming is a general approach to solving optimization problems. Each problem is viewed as a new one and one has to develop some insight to recognize when a problem can be solved by dynamic programming technique and how it could be done. This ability can possibly be best developed by an exposure to a wide variety of dynamic programming applications. For this purpose, we present several examples and show how the dynamic programming technique can be used to find their solutions.

However, There are some common features of all dynamic programming problems, which act as a guide to develop the dynamic programming model.

29.2. Basic Features of Dynamic Programming Problems and the Principle of Optimality

The basic features which characterize dynamic programming problems are as follows:

(a) The problem can be divided (decomposed) into subproblems which are called stages.

(b) At each stage, the system is characterized by a small set of parameters called the state variables.

(c) At each stage, there is a choice of a number of decisions.

(d) The effect of a decision at each stage is to transform the current state into a state of the system at the next stage.
(e) Given the current state, an optimal decision for the remaining stages is independent of the decision taken in previous stages.
(f) The purpose of the process is to optimize a predefined function of the state variables called the objective function or the criterion function.

Thus, the common characteristic of all dynamic programming models is expressing the decision problem by a functional equation obtained by an application of the principle of optimality.

Principle of Optimality [Bellman]

An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

The principle states that the optimal policy starting in a given state depends only on that state and not upon how one got to the state. This is the Markovian property in dynamic programming.

29.3. The Functional Equation

Since a multi-stage decision process can be characterized by the initial state of the system and the length of the process, a functional equation may be developed as follows.

Consider an N-stage process and let s be the set of state parameters.

Define for $n \leq N$, $f_n(s, x_n)$ = the return from an n-stage process given that there are s states and a decision x_n is used.

$R_n(s, x_n)$ = the return from the first stage of an n-stage process with state s, using decision x_n.

s^1 = the new state resulting from decision x_n.

and $f^*_n(s)$ = the total return from an n-stage process where the system is in state s and an optimal decision is used.

Then $\quad f^*_n(s) = \text{Max } f_n(s, x_n) = f_n(s, x_n^*)$

where x_n^* denote the value of x_n, which maximizes $f_n(s, x_n)$.

Thus, we have

$$f_1^*(s) = \underset{x_1 \in S}{\text{Max}}\ R_1(s, x_1) \tag{29.1}$$

$$f_n^*(s) = \underset{x_n \in S}{\text{Max}}\ [R_n(s, x_n) + f^*_{n-1}(s^1)] \tag{29.2}$$

$$n = 2, 3, \ldots N$$

The form of the functional equation may however differ depending on the nature of the problem.

To illustrate the dynamic programming technique, let us first consider the well known cargo loading problem.

29.4. Cargo Loading Problem

Consider a well-known problem of loading a vessel (or a Knapsack or flyaway-kit) with different items so that the total value of its content is maximum subject to the constraint that the total weight must not exceed a specified limit.

Mathematically, the problem can be stated as,

$$\begin{aligned} \text{Maximize} \quad & v_1x_1 + v_2x_2 + \dots + v_Nx_N \\ \text{Subject to} \quad & w_1x_1 + w_2x_2 + \dots + w_Nx_N \le W \\ & x_i, \ (i = 1, 2 \dots N) \text{ are nonnegative integers,} \end{aligned} \tag{29.3}$$

where v_i and w_i are the value and the weight per unit of the ith item respectively,

x_i is the number of units of item i, $\quad i = 1, 2, \dots N$ and

W is the maximum allowable weight.

For dynamic programming formulation, let us consider the items as stages and the state of the system be defined as the weight capacity available.

Let $f_n(s_n, x_n)$ be the value of the load for the stages 1, 2,...n when the system is in state s_n and a decision x_n is used.

$f^*_n(s_n)$ = Total value of the load, when the system is in state s and an optimal decision is used.

Then $\quad f^*_n(s_n) = \underset{x_n}{\text{Max}} \ f_n(s_n, x_n) = f_n(s_n, x_n^*)$

Where x_n^* is the value of x_n, which maximizes $f_n(s_n, x_n)$

Thus for the first stage,

$$f^*_1(s_1) = \underset{x_1}{\text{Max}} \ f_1(s_1, x_1)$$

$$= \underset{x_1 \in \left(0,1,\dots\left[\frac{s_1}{w_1}\right]\right)}{\text{Max}} (v_1x_1) \tag{29.4}$$

and $\quad f^*_n(s_n) = \underset{x_1 \in \left(0,1,\dots\left[\frac{s_1}{w_1}\right]\right)}{\text{Max}} [v_nx_n + f^*_{n-1}(s_n - w_nx_n)]$

$$n = 2, 3, \dots N \tag{29.5}$$

where $\left[\dfrac{s_n}{w_n}\right]$ is the largest integer in $\left(\dfrac{s_n}{w_n}\right)$

Consider the following simple numerical problem with the data as given in Table 29.1

Table 29.1

Item i	*Weight/Unit* w	*Value/Unit* v
1	3	5
2	4	8
3	2	4

and W = 7

The computations for the three stages of the problem are given in the following tables.

Stage1. Since $s_1 = \{0, 1, 2, \ldots 7\}$, the largest value of x_1 is $\left[\frac{s_1}{w_1}\right] = \left[\frac{7}{3}\right] = 2.$

Table 29.2

	$f_1(s_1, x_1) = 5x_1$			Optimal solution	
s_1 \ x_1	0	1	2	$f_1^*(s_1)$	x_1^*
0	0			0	0
1	0			0	0
2	0			0	0
3	0	5		5	1
4	0	5		5	1
5	0	5		5	1
6	0	5	10	10	2
7	0	5	10	10	2

Stage 2. For the two-stage process the largest value of x_2 is $\left[\frac{7}{4}\right] = 1.$

Table 29.3

	$f_2(s_2, x_2) = 8x_2 + f^*_1(s_2 - 4x_2)$		Optimal solution	
s_2 \ x_2	0	1	$f_2^*(s_2)$	x_2^*
0	0 + 0 = 0		0	0
1	0 + 0 = 0		0	0
2	0 + 0 = 0		0	0
3	0 + 5 = 5		5	0
4	0 + 5 = 5	8 + 0 = 8	8	1
5	0 + 5 = 5	8 + 0 = 8	8	1
6	0 + 10 = 10	8 + 0 = 8	8	1
7	0 + 10 = 10	8 + 5 = 13	13	1

Stage 3. Finally for the three-stage process, the largest value of x_3 is $\left[\frac{7}{2}\right] = 3$.

Table 29.4

	$f_3(s_3, x_3) = 4x_3 + f_2(s_3 - 2x_3)$				Optimal solution	
s_3 \ x_3	0	1	2	3	$f_3^*(s_3)$	x_3^*
0	0 + 0 = 0				0	0
1	0 + 0 = 0				0	0
2	0 + 0 = 0	4 + 0 = 4			4	1
3	0 + 5 = 5	4 + 0 = 4			5	0
4	0 + 8 = 8	4 + 0 = 4	8 + 0 = 8		8	0, 2
5	0 + 8 = 8	4 + 5 = 9	8 + 0 = 8		9	1
6	0 + 8 = 8	4 + 8 = 12	8 + 0 = 8	12 + 0 = 12	12	1, 3
7	0 + 13 = 13	4 + 8 = 12	8 + 5 = 13	12 + 0 = 12	13	0, 2

Optimal solution corresponding to W = 7, can now easily be read out from the tables calculated above.

If we take $x_3^* = 0$, then there are 7 choices left for stage 2 and then from Table 29.3 we find $x_2^* = 1$. Hence only (7-1x4) = 3 choices are left for stage 1 and using Table 29.2 we get $x_1^* = 1$.

Thus an optimal solution is given by

$$x_1^* = 1,\ x_2^* = 1 \text{ and } x_3^* = 0$$

and the total value of the load = $1 \times 5 + 1 \times 8 + 0 \times 4 = 13$.

An alternative optimal solution can be obtained if we take $x_3^* = 2$ and proceed in the same way as above. The alternative optimal solution is then given by,

$x_1^* = 1$, $x_2^* = 0$, $x_3^* = 2$ with the same total value 13, as was expected.

29.5. Forward and Backward Computations, Sensitivity Analysis

The computation procedure of dynamic programming problem differs depending on whether the computation starts with the initial stage or the final stage. If the computation starts at the first stage and moves towards the final stage, it is known as the forward computation. If, on the other hand, the computation begins at the final stage reaching ultimately the initial stage, it is called the backward computation. Both the procedures give the same result but it is usually convenient to solve the problem by backward computation when the initial state is specified while the forward computation is more efficient when the final state is given.

Sensitivity Analysis

The tabular computations of dynamic programming problems also provide considerable information to study sensitivity of the solution to variations of state input or to know the effect of making the planning horizon longer on the total return.

The backward calculation is well suited for sensitivity analysis to see how the total return varies with the input state since at the final table, the problem of whole range of inputs can be easily solved.

On the other hand, if we want to know how the length of the planning horizon influences the optimal decision, we use the forward calculation since additional stages can be added in the process of computation.

29.6. Shortest Route Problem

Consider the problem of finding a shortest route through a net work, which arises, is a wide variety of applications. Suppose that a salesman wants to find a shortest route from his starting point (station 1) to his destination (station 10) from the road map as given in Figure 29.1.

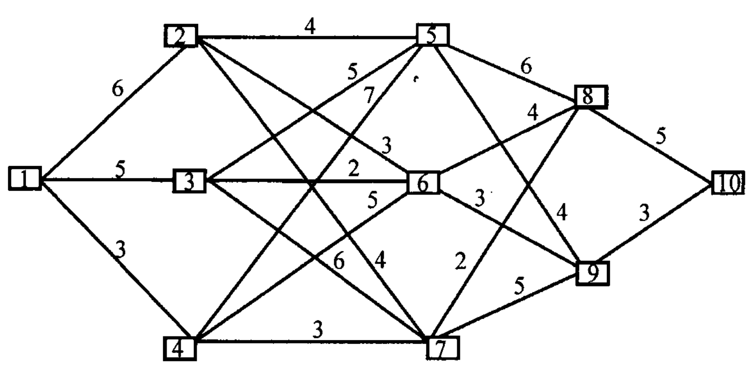


Figure 29.1.

There are eight intermediate stations, all of which however, are not connected with one another. The distances between connecting stations are indicated in the road map.

It is noted that the salesman has to travel through four stages regardless of the particular routine and each station in a stage represents a state.

Let d_{ij} denote the distance between state i and state j and x_n be the decision variable to move to the next station, when there are n more stages to go. (n = 1, 2, 3, 4)

Let f_n (s, x_n) be the total distance of the best over-all policy for the last n stages when the salesman is in state s and x_n is the decision taken.

$$f_n^*(s) = \underset{x_n}{\text{Min}}\ f_n(s, x_n) = f_n(s, x_n^*)$$

where x_n^* is the value of x_n which minimizes f_n (s, x_n).

Then, $$f_n^*(s) = \underset{x_n}{\text{Min}}\ [d_{sxn} + f^*_{n-1(xn)}] \tag{29.6}$$

Computational Procedure

For convenience the distance d_{ij} from state i to state j is reproduced in the table below.

Table 29.5

	2	3	4
1	6	5	3

	5	6	7
2	4	3	4
3	5	2	6
4	7	5	3

	8	9
5	6	4
6	4	3
7	2	5

	10
8	5
9	3

Stage 1. In this problem, the computation is carried out from the end of the process. For this one-stage problem we then have

s	$f_1^*(s)$	x_1^*
8	5	10
9	3	10

For other stages, the results are obtained by recurrence relation.

Stage 2.

s \ x_2	$f_2(s, x_2) = dsx_2 + f_1^*(x_2)$: 8	9	Optimal solution: $f_2^*(s)$	x_2^*
5	6 + 5 = 11	4 + 3 = 7	7	9
6	4 + 5 = 9	3 + 3 = 6	6	9
7	2 + 5 = 7	5 + 3 = 8	7	8

Stage 3.

s \ x_3	$f_3(s, x_3) = ds,x_3 + f_2^*(x_3)$: 5	6	7	Optimal solution: $f_3^*(s)$	x_3^*
2	4 + 7 = 11	3 + 6 = 9	4 + 7 = 11	9	6
3	5 + 7 = 12	2 + 6 = 8	6 + 7 = 13	8	6
4	7 + 7 = 14	5 + 6 = 11	3 + 7 = 10	10	7

Stage 4.

s \ x_4	$f_4(s, x_4) = dsx_4 + f_3^*(x_4)$: 2	3	4	Optimal solution: $f_4^*(s)$	x_4^*
1	6 + 9 =15	5 + 8 = 13	3 + 10 = 13	13	3, 4

The computation now terminates and the optimal solution can be read out from the tables above.

The result of the stage-4 table indicates that an optimal decision for the salesman is to go from state 1 to either state 3 or state 4. If he chooses to go to state 3, then the 3-stage table shows that for s = 3, an optimal decision is to go to state 6. Continuing to the 2-stage table we find that when the salesman enters state 6, an optimal decision is to go to state 9 and from state 9, he finally goes to state 10.

Hence an optimal route is

$$1 \to 3 \to 6 \to 9 \to 10.$$

If the salesman chooses to go to state 4 initially an alternative optimal route is obtained as

$$1 \to 4 \to 7 \to 8 \to 10$$

The minimum distance that the salesman has to travel is 13.

29.7. Investment Planning

An entrepreneur is considering to invest his capital to four activities and expects to get returns as given in the table below. The total budget is limited to 5 units of money and only an integer number of units can be allocated. The entrepreneur wishes to find a plan of allocation so that return is maximum.

Table 29.6. Return from Investments

		Activity			
		1	2	3	4
	0	0	0	0	0
	1	7	5	3	3
	2	9	7	4	5
Capital	3	9	8	5	8
	4	9	8	5	9
	5	9	8	5	9

Computations

For dynamic programming formulation, let us consider the four activities as the four stages and let the decision variable x_n be the number of units of capital allocated to the nth stage from the end. Thus x_n is the number of units of capital allocated to activity $5 - n$. $(n = 1, 2, 3, 4)$.

Since the choice of the decision variable at any stage is the number of units of capital available, they constitute the state of the system.

Let $r_{5-n}(x_n)$ = the return from the $(5 - n)$ th activity, that is, from the first stage of the n-stage process from the end, $n = 1, 2, 3, 4$.

$f_n(s, x_n)$ = The return from the last n stages when the state available is s and x_n is the decision taken.

Then $\quad f_n^*(s) = \underset{x_n=0,1,\dots s}{\text{Max}} f_n(s, x_n)$

$$x_n = 0, 1, \dots s$$

$$= \underset{x_n=0,1,\dots s}{\text{Max}} \left[r_{5-n}(x_n) + f_{n-1}^*(s - x_n)\right] \tag{29.7}$$

$$n = 2, 3, 4.$$

and $\quad f_1^*(s) = \underset{x_n=0,1,2,\dots s}{\text{Max}} r_4(x_1) \qquad (29.8)$

We now proceed with our calculation beginning with the first stage from the end (Activity 4) and move backward to the stage 4 (Activity 1).

Stage 1.

s	$f_1^*(s)$	x_1^*
0	0	0
1	3	1
2	5	2
3	8	3
4	9	4
5	9	4, 5

Stage 2.

	$f_2(s, x_2) = r_3(x_2)+f_1^*(s-x_2)$						Optimal solution	
$s \backslash x_2$	0	1	2	3	4	5	$f_2^*(s)$	x_2^*
0	0						0	0
1	3	3					3	0, 1
2	5	6	4				6	1
3	8	8	7	5			8	0, 1
4	9	11	9	8	5		11	1
5	9	12	12	10	8	5	12	1, 2

Stage 3.

	$f_3(s, x_3) = r_2(x_3)+f_2^*(s-x_3)$						Optimal solution	
$s \backslash x_3$	0	1	2	3	4	5	$f_3^*(s)$	x_3^*
0	0						0	0
1	3	5					5	1
2	6	8	7				8	1
3	8	11	10	8			11	1
4	11	13	13	11	8		13	1, 2
5	12	16	15	14	11	8	16	1

Stage 4.

	$f_4(s, x_4) = r_1(x_4)+f_3^*(s-x_4)$						Optimal solution	
$s \backslash x_4$	0	1	2	3	4	5	$f_4^*(s)$	x_4^*
5	16	20	20	17	14	9	20	1, 2

The, solution of the problem can now be read out from the tables above. Optimal allocations of the capital to the four activities are as follows:

Activity 1 x_4^*	Activity 2 x_3^*	Activity 3 x_2^*	Activity 4 x_1^*
1	1	1	2
1	2	1	1
2	1	1	1

Total return is 20.

29.8. Inventory Problem

Suppose that a Company producing a single item wants to determine a production schedule to meet the fluctuating demand over the next n periods, so that the total cost incurred is minimum.

It is assumed that the production is instantaneous (i.e. the time to produce the item is negligible). The demand varies from one period to another but otherwise known (by an accurate forcast) and the shortages are not allowed. However, there is a cost of holding on the inventory at the end of the period. Further, it is assumed that the inventory level is zero at the end of the period n.

Let us define for the period i (i = 1, 2,...n)

x_i = Quantity of the item produced.

d_i = demand in integer.

z_i = inventory at the beginning of the period (entering Inventory).

h_i = holding cost per unit of inventory at the end of the period.

k_i = set up cost

$p_i(x_i)$ = production cost of x_i in the period.

Now, the total cost incurred in each period depends on the set up cost, production cost and the holding cost for the inventory at the end of the period.

Let

$$c_i(x_i) = \delta_i k_i + p_i(x_i)$$

Where

$$\delta_i = \begin{cases} 0, & \text{if } x_i = 0 \\ 1, & \text{if } x_i > 0 \end{cases}$$

The objective function can then be written as

$$\text{Minimize} \quad \sum_{i=1}^{n} [c_i(x_i) + h_i(z_{i+1})]$$

where z_{i+1} is the amount of inventory carried forward from period i to period i+1 and hence

$$z_{i+1} = z_i + x_i - d_i, \quad \text{for } i = 1, 2, \ldots n$$

and $\quad z_{n+1} = 0$

The dynamic programming formulation for forward computation for this problem can be developed as follows.

Let each period be considered as a stage and the states of the system at any stage i be the amount of inventory at the end of the period i (i.e. z_{i+1}).

Let $f_i(z_{i+1})$ be the minimum inventory cost for the first i periods 1 to i, given that z_{i+1} is the inventory at the end of the ith period. Since it may be more economical to produce more in one period and store the excess to meet the demand in future, the values of z_{i+1} is limited by

$$0 \le z_{i+1} \le d_{i+1} + d_{i+2} + d_{i+3} + \ldots d_n.$$

The recurrence relations are then given by,

$$f_1(z_2) = \underset{0 \le x_1 \le d_1 + z_2}{\text{Min}} [c_1(x_1) + h_1 z_2] \tag{29.9}$$

$$f_i = (z_{i+1}) = \underset{0 \le x_i \le d_i + z_{i+1}}{\text{Min}} [c_i(x_i) + h_i z_{i+1} + f_{i-1}(z_{i+1} + d_i - x_i)] \; i = 2, 3, \ldots n \tag{29.10}$$

We now consider a numerical problem to illustrate the computational procedure described above.

Consider a three-period inventory problem with demand varying from one period to another but otherwise known and the holding cost is based on the inventory at the end of the period.

The data for the problem are given in the following table.

Period i	Set up cost k	Holding cost in dollar h	Demand d
1	3	1	4
2	6	3	3
3	4	2	5

The initial (entering) inventory z_1 at period 1 is 1 and the inventory at the end of period 3 is equal to zero. The production cost at period i is given by

$$p_i(x_i) = \begin{cases} 15x_i, & \text{if } 0 \le x_i \le 4 \\ 60 + 30(x_i - 4), & \text{if } x_i \ge 5 \end{cases}$$

This means that the cost of production per unit is 15 for the first four units and 30 for any numbr of units in excess of that.

Computations

Stage 1. Since $z_1 = 1$, the smallest value of x_1 is $d_1 - z_1 = 4 - 1 = 3$

and the largest $x_1 = d_1 + d_2 + d_3 - 1 = 4+3+5-1 = 11$

and $0 \le z_2 \le 3 + 5 = 8.$, $k = 3$

		$f_1(x_1 : z_2) = c_1(x_1) + h_1 z_2$									Optimal solution	
	x_1	3	4	5	6	7	8	9	10	11		
z_2	$c_1(x_1)$ \\ $h_1(z_2)$	48	63	90	120	150	180	210	240	270	$f_1(z_2)$	x_1^*
0	0	48									48	3
1	1		64								64	4
2	2			92							92	5
3	3				123						123	6
4	4					154					154	7
5	5						185				185	8
6	6							216			216	9
7	7								247		247	10
8	8									278	278	11

Stage 2.

$d_2 = 3, 0 \le x_2 \le 8, 0 \le z_3 \le 5, k = 6$

		$f_2(x_2 : z_3) = c_2(x_2) + h_2 z_3 + f_1(z_3 + d_2 - x_2)$									Optimal solution	
	x_2	0	1	2	3	4	5	6	7	8		
z_3	$h_2 z_3$ \\ $c_2(x_2)$	0	21	36	51	66	96	126	156	186	$f_2(z_3)$	x_2^*
0	0	123	113	100	99						99	3
1	3	157	147	131	118	117					117	4
2	6	191	181	165	149	136	150				136	4
3	9	225	215	199	183	167	169	183	216		167	4
4	12	259	249	233	217	201	200	202	235		200	5
5	15	293	283	267	251	235	234	233	235	240	233	6

Stage 3. $d_2 = 5, 0 \le x_3 \le 5, z_4 = 0, k = 4$

		$f_3(x_3:z_4) = c_3(x_3)=+h_3z_4+f_2(z_4+d_3-x_3)$						Optimal solution	
	x_3	0	1	2	3	4	5		
z_4	$c_3(x_3)$ / h_3z_4	0	19	34	49	64	94	$f_3(z_4)$	x_3*
0	0	233	219	201	185	181	193	181	4

The solution is obtained from the tables above as

Period i	1	2	3
x_i^*	3	4	4

at a total cost of $181.

29.9. Reliability Problem

Consider the problem of designing a complex equipment where reliability (that is probability of no failure) is a most important requirement.

Suppose that the components are arranged in series so that a failure of one component will cause the failure of the whole system. To increase the reliability of the system, one might instal duplicate units in parallel on each component (subsystem) with a switching device so that a component is automatically replaced by new one when the old one fails.

Suppose that the whole system has N subsystems with duplicate components in parallel (stages) which are arranged in series and that at least one component must be used at each substation.

Let $1 + m_j$ = the number of components used at the jth stage, where m_j is the number of stand-by units at stage j ($m_j = 0, 1, 2,\ldots; j = 1, 2,\ldots N$)

and $\phi_j(m_j)$ = the probability of successful operation of the jth stage when $1+m_j$ components are used at the jth stage.

Assuming that the probabilities at different stages are independent, we have the system reliability of the N-stage device.

$$R_N = \prod_{j=1}^{N} \phi_j(m_j)$$

Let c_j = the cost of a single component at the jth stage.

and C be the total capital available.

The objective is to determine the value of m_j which will maximize the total reliability of the system without exceeding the available capital.

The problem thus becomes

$$\text{Maximize} \quad R_N = \prod_{j=1}^{N} \phi_j(m_j)$$

$$\text{Subject to} \quad \sum_{j=1}^{N} c_j m_j \le C$$

$$m_j = 0, 1, 2, \ldots$$

Let $f_n(x_n)$ represent the optimal reliability for the stages 1 through n when the state (capital) at stage n is x_n where $0 \le x_n \le C$. The recurrence relations may then be obtained as

$$f_1(x_1) = \underset{m_1}{\operatorname{Max}} \phi_1(m_1) \tag{29.11}$$

$$0 \le c_1 m_1 \le x_1$$

and

$$f_n(x_n) = \underset{\substack{m_n \\ 0 \le c_n m_n \le x_n}}{\operatorname{Max}} \left[\phi_n(m_n) - f_{n-1}(x_n - c_n m_n)\right]$$

$$n = 2, 3, \ldots N \tag{29.12}$$

From the recurrence relations, computations for different stages can be carried out from which the optimal solution of the problem can be obtained.

29.10. Cases where Decision Variables are Continuous

In all the examples considered above, the decision process was of finite length and the stage, state and decision variables were represented by integer numbers and it was observed that dynamic programming technique can be suitably applied to find an optimal solution of the problems. Let us now show how the dynamic programming technique can be applied to problems, where the decision variables are not restricted to be integers.

Consider the simple allocation problem where we are to maximize the return function

$$R(x_1, x_2, \ldots x_N) = \sum_{i=1}^{N} g_i(x_i)$$

$$\text{Subject to} \quad \sum_{i=1}^{N} x_i \le b \tag{29.13}$$

$$x_i \ge 0, \quad i = 1, 2, \ldots N$$

where each $g_i(x_i)$ is assumed to be continuous for all $x_i \ge 0$.

Let, as in the previous, the activities be considered as stages and s = b be the state of the system. Let s_n, denote the state at the stage n, so that

$$0 \le s_n \le b, \quad n = 1, 2, \ldots N.$$

Define $f_N(s) = \underset{(x_i)}{\operatorname{Max}} R(x_1, x_2, \ldots x_N)$

and thus $f_n(s_n)$ represents the maximum return from the n-stage process when s_n is the state at the stage n. The recurrence relations are then,

$$f_n(s_n) = \max_{0 \le x_n \le s_n} [g_n(x_n) + f_{n-1}(s_n - x_n)] \quad , n \ge 2 \tag{29.14}$$

and $$f_1(s_1) = \max_{0 \le x_1 \le s_1} g_1(x_1) \tag{29.15}$$

It is clear that $f_n(0) = 0$, provided that $g_i(0) = 0$ for each i. , n = 1, 2,...N.

Now, to find an optimal policy, we are to determine the sequence f_n (x), n = 1, 2,...N, for all x in the interval [0, b]. It is clearly not possible to find all the values of a function in an interval. However, there are several search procedures (see [515]), which under certain conditions can be applied for finding an optimum of a function, particularly if the function is unimodal such as strictly convex or concave function. In the general case, when the function does not possess any special structure we use some interpolation scheme which permits us to recreate a general value from a few carefully chosen values.

To achieve this, each of the functions $f_n(s_n)$, n = 1, 2,...N are evaluated and tabulated only at each of the finite grid points.

$$s_n = 0, \Delta, 2\Delta, \ldots r\Delta = b$$

The values of $f_n(s_n)$, for s_n distinct from these grid points are then obtained by interpolation. The type of interpolation to be used depends on the accuracy desired and on the time required to achieve this accuracy.

If $\quad k\Delta \le s_n \le (k+1)\Delta$,

the simplest approximate value of $f_n(s_n)$ is obtained by setting,

$$f_n(s_n) = f_n(k\Delta)$$

The next simplest approximation is obtained by the linear interpolation formula

$$f_n(s_n) = f_n(k\Delta) + (s_n - k\Delta)[f_n(n+1)\Delta - f_n(k\Delta)]/\Delta$$

However, more accurate higher order interpolation formulas may be used, if so desired.

Moreover, the decision variables x_n also are allowed to range over the same set of grid points as above.

The computation for finding an optimal policy then proceeds as follows.

$f_1(s_1)$ is computed from the relation

$$f_1(s_1) = \max_{0 \le x_1 \le s_1} g_1(x_1)$$

where s_1 takes the values 0, Δ, 2Δ,...rΔ.

The set of values $\{f_1(k\Delta)\}$, thus obtained is stored (or tabulated) in the memory of the computer along with the corresponding maximizing x_1-values

$\{x_1(k\Delta)\}$, k = 0, 1, 2,...r.

We now compute

$$f_2(s_2) = \max_{0 \le x_2 \le s_2} [g_2(x_2) + f_1(s_2 - x_2)]$$

where s_2 assumes only the values $0, \Delta, 2\Delta, \ldots r\Delta$. Since no enumerative process can yield maximization over a continuous range of values, we replace the interval $[0, s_2]$ by a discrete set of values and compute

$$f_2(k\Delta) = \max_{x_2 = k\Delta} [g_2(x_2) + f_1(k\Delta - x_2)], \qquad k = 0, 1, 2, \ldots r.$$

To begin the maximization process, we first evaluate, $g_2(0) + f_1(k\Delta)$ and $g_2(\Delta) + f_1((k-1)\Delta)$ and retain the largest of these two values. The value $g_2(2\Delta) + f_1((k-2)\Delta)$ is then computed and compared with the previously obtained larger value and the larger of these two values is retained. This process is continued until all the values of $g(x_2) + f_1(k\Delta - x_2)$, for $x_2 = 0, \Delta, 2\Delta, \ldots k\Delta$ have been compared. The process yields $f_2(k\Delta)$ and also the maximizing point (points) $x_2(k\Delta)$. Following this procedure, the values of $f_2(s_2)$ are computed and stored for each s_2 belonging to the set $\{0, \Delta, 2\Delta, \ldots r\Delta\}$.

The procedure may then be continued for the N-stage process and the result may be tabulated.

Table 29.7

s	$f_1(s)$	$x_1(s)$	$f_2(s)$	$x_2(s)$	....	$f_N(s)$	$x_N(s)$
0	$f_1(0)$	$x_1(0)$	$f_2(0)$	$x_2(0)$	....	$f_N(0)$	$x_N(0)$
Δ	$f_1(\Delta)$	$x_1(\Delta)$	$f_2(\Delta)$	$x_2(\Delta)$	....	$f_N(\Delta)$	$x_N(\Delta)$
2Δ	$f_1(2\Delta)$	$x_1(2\Delta)$	$f_2(2\Delta)$	$x_2(2\Delta)$	....	$f_N(2\Delta)$	$x_N(2\Delta)$
$\vdots$							
$\vdots$							
$\vdots$							
$r\Delta$	$f_1(r\Delta)$	$x_1(r\Delta)$	$f_2(r\Delta)$	$x_2(r\Delta)$	....	$f_N(r\Delta)$	$x_N(r\Delta)$

The solution to the N-stage process can now be read out from the table 29.7. Given a particular value of s_N, the state at the stage N, the value of $x_N(s_N)$ is noted from the column of $x_N(s)$ which is the maximizing value of x_N and is denoted by x^*_N. Once x^*_N I determined, we have a problem of (N–1) stage process with the state of the system $s_{N-1} = x_N - x^*_N$. The maximizing value x^*_{N-1} for the (N–1) stage process can then easily be obtained from the table. Continuing this procedure we find the optimal solution, $x^*_N, x^*_{N-1} \ldots x^*_2, x^*_1$ in that order.

As already indicated, the decision process may have several optimal policies. It is advisable to retain all the optimal policies as they may sometime turn out to be quite helpful in selecting the final decision to be implemented. Moreover, near optimal policies should also be recorded, as they may be important in providing simple approximations for more complex situations.

For more details, see Bellman[44] and Bellman and Dreyfus[46].

29.11. The Problem of Dimensionality

In the problems presented so far, the states of the system were represented by a single variable only and at each stage, only a single decision variable was to be determined. We now consider more complicated problems arising from various realistic situations, in which the state variables may be more than one (a state vector) and at each stage, values of two or more decision variables must be determined. The dynamic programming problem is then said to be multidimensional. The recurrence relations for the problem can be developed in the same way as was done in the previous sections but it can be seen that the number of computations increases exponentially with the increase in the dimension of the problem. Since the amount of information to be stored is enormous, it may tax the computer memory and increase the computation time or even may be beyond the range of available computers. This presents a great limitation to the use of dynamic programming and is aptly called the "Curse of dimensionality", by Richard Bellman.

To overcome these difficulties, several methods have been suggested [44, 45, 46], the most powerful of which is perhaps the method of Lagrange Multipliers.

29.11.1. Allocation Problems with Two Types of Resources and Two Decision Variables

Consider an allocation problem with two different types of resources which are to be allocated to a number of independent activities. The problem of our concern is to determine an allocation of the resources to the activities so that the return function is maximum.

Consider the problem

$$\text{Maximize} \quad R(x_1, x_2 \ldots x_N; y_1, y_2 \ldots y_N) = \sum_{i=1}^{N} g_i(x_i, y_i)$$

$$\text{Subject to} \quad \sum_{i=1}^{N} a_{1i}\, x_i \le b_1 \tag{29.16}$$

$$\sum_{i=1}^{N} a_{2i}\, y_i \le b_2$$

$$x_i \ge 0,\ y_i \ge 0,\ i = 1, 2, \ldots N$$

where a_{1i}, a_{2i} have positive values for each i and b_1, b_2 are the quantities of the two type of resources to be allocated to the N activities. The activities are regarded as stages and the levels of activity (x_i, y_i) represent the decision variables at stage i (i = 1, 2,...n). The states of the system are defined as the amount of two resources available to be allocated to the current stage and the succeeding stages. Thus, the state s_n at stage n is the vector $s_n = (s_{1n}, s_{2n})^T$.

It is assumed that each function $g_i(x_i, y_i)$ is continuous for all $x_i \ge 0$, $y_i \ge 0$, (i = 1, 2,...N). This implies that the state functions $f_n(s_{1n}, s_{2n})$ are continuous of

s_{1n} and s_{2n}.

Following the general approach to dynamic programming we define

$$f_i^*(s_{1n}, s_{2n}) = \underset{\substack{x_1, x_2 \ldots x_n \\ y_1, y_2 \ldots y_n}}{\text{Max}} R_n(x_1, x_2, \ldots x_n; y_1, y_2, \ldots y_n)$$

where maximization is taken over x_i, y_i satisfying

$$\sum_{i=1}^{n} a_{1i} x_i \leq s_{1n}$$

$$\sum_{i=1}^{n} a_{2i} y_i \leq s_{2n}$$

$$x_i \geq 0, y_i \geq 0, i = 1, 2, \ldots n.$$

The recurrence relations are then given by

$$f_i^*(s_{1i}, s_{2i}) = \underset{\substack{0 \leq a_{1i} x_i \leq s_{1i} \\ 0 \leq a_{2i} y_i \leq s_{2i}}}{\text{Max}} g_i(x_i, y_i) \tag{29.17}$$

$$f_n^*(s_{1n}, s_{2n}) = \underset{\substack{0 \leq a_{1n} xn \leq s_{1n} \\ 0 \leq a_{2n} yn \leq s_{2n}}}{\text{Max}} [g_n(x_n, y_n) + f_{n-1}(s_{1n} - a_{1n}x_n, s_{2n} - a_{2n}y_n)$$

$$n = 2, 3, \ldots N \tag{29.18}$$

Since the decision variables are nonnegative and not restricted to be integers, we follow a simple extension of the approach discussed in section 29.10. To determine the sequence of functions $\{f_n(s_{1n}, s_{2n})\}$ in the region defined by $(0 \leq s_{n1} \leq b_1; 0 \leq s_{n2} \leq b_2)$, the functions are now evaluated at a set of lattice points, say the points $s_{n1} = k\Delta$, $s_{n2} = l\Delta$, k, l = 0, 1, 2,..., which are required to be stored in the computer memory. The computational effort involved however is enormous which rapidly increases with the increase in the number of state variables. The amount of information to be stored may be so enormous that it may be beyond the range of available computers.

29.11.2. Allocation Problems with Two Constraints and One Decision Variable

Consider the problem of allocating two resources to N activities subject to two constraints involving one decision variable. Such a problem appears in many contexts. For example, in cargo-loading problem, we may have both weight and volume restrictions.

Mathematically, the problem can be stated as

$$\text{Maximize } R(x_1, x_2 \ldots xn) = \sum_{i=1}^{N} g_i x_i$$

$$\text{Subject to} \qquad \sum_{i=1}^{N} a_{1i}x_i \le b_1 \tag{29.19}$$

$$\sum_{i=1}^{N} a_{2i}x_i \le b_2$$

$$x_i \ge 0,\ i = 1, 2, \ldots N.$$

where a_{1i}, a_{2i} ($i = 1, 2, \ldots N$) have positive values. It is assumed that each function $g_i(x_i)$ is continuous for all $x_i \ge 0$.

As in section 29.11.1, the N activities are regarded as stages and the levels of activity x_i represent the decision variables at stage i ($i = 1, 2, \ldots N$). The states are defined as the amount of resources to be allocated to the current and the remaining stages.

Let the optimal return from the n-stage process be defined by

$$f_n^*(s_{1n}, s_{2n}) = \underset{x_1, x_2, \ldots x_n}{\text{Max}} R_n(x_1, x_2, \ldots x_n)$$

where the maximization is taken over x_i satisfying

$$\sum_{i=1}^{n} a_{1i}x_i \le s_{1n}$$

$$\sum_{i=1}^{n} a_{2i}x_i \le s_{2n}$$

$$x_i \ge 0,\ i = 1, 2, \ldots n$$

We then have the recurrence relations

$$f_1(s_{11}, s_{21}) = \underset{\substack{0 \le a_{11} x_1 \le s_{11} \\ 0 \le a_{21}x_1 \le s_{21}}}{\text{Max}} g_1(x_1)$$

and

$$f_n^*(s_{1n}, s_{2n}) = \underset{\substack{0 \le a_{1n}x_n \le s_{1n} \\ 0 \le a_{2n}x_n \le s_{2n}}}{\text{Max}}[g_n(x_n) + f_{n-1}(s_{1n} - a_{1n}x_n, s_{2n} - a_{2n}x_n)$$

$$n = 2, 3, \ldots N. \tag{29.21}$$

As in the previous case the computational effort to solve the problem is enormous and it may not be possible to make use of available computers for our purpose.

From sections 29.11.1 and 29.11.2, we note that as the number of state variables increases, the total number of computations increase enormously and due to huge memory requirements, it becomes essentially impossible to use present-day computers to solve the problems by dynamic programming. The most obvious way to reduce requirements is to reduce storage the number of state variables. In the next section we will discuss methods that can be used to reduce the dimension of the problem. However, problems involving only two decision variables can usually be solved on a large computer.

As an example,

consider the following linear programming problem,

$$\begin{aligned} \text{Maximize} \quad & z = 5x_1 + 3x_2 \\ \text{Subject to} \quad & 2x_1 + 3x_2 \le 36 \\ & x_1 \le 12 \\ & x_2 \le 8 \\ & x_1 \ge 0,\ x_2 \ge 0 \end{aligned}$$

The two activities are considered as the two stages and since there are three resources, the state s is represented by a three-component vector. Thus, $s = (s_1, s_2, s_3)$, where s_i is the amount of resource i remaining to be allocated (i = 1, 2, 3). The nonnegative decision variables x_1, x_2 are continuous and so they possess infinite number of possible values within the feasible space. This creates complications in the computation as discussed in section 29.10.

However, the present problem is small enough so that it can still be solved without much difficulty.

Let $\quad f_2^*(s_1, s_2, s_3) = \text{Max } z = \underset{(x_1,x_2)}{\text{Max}} (5x_1 + 3x_2)$

Using the backward recursion, we have

$$f_1^*(s_1,s_2,s_3) = \underset{3x_2 \le s_1}{\text{Max}}\, 3x_2 = 3\text{Min}\left(\frac{s_1}{3,} s_3\right)$$

$$x_2 \le s_3$$
$$x_2 \ge 0$$

and the optimal value of x_2 is

$$x^*_2 = \min\left(\frac{s_1}{3,} s_3\right)$$

Now, $\quad f_1^*(s_1,s_2,s_3) = \underset{2x_1 \le 36}{\text{Max}}[5x_1 + f_1^*(s_1 - 2x_1, s_2 - x_1, s_3)]$

$$x_1 \le 12$$
$$x_1 \ge 0$$

$$= \text{Max}\ [5x_1 + f_1^*(36-2x_1,\ 12-x_1,\ 8)]$$

$$2x_1 \le 36$$
$$x_1 \le 12$$
$$x_1 \ge 0$$

$$= \underset{0 \le x_1 \le 12}{\text{Max}}\left[5x_1 + 3\text{Min}\left(\frac{36-2x_1}{3}, 8\right)\right]$$

Now, $\text{Min}\left[\frac{36-2x_1}{3}, 8\right] = \begin{cases} 8, & \text{for } 0 \le x_1 \le 6 \\ 12 - \frac{2}{3}x_1 & \text{for } 6 \le x_1 \le 12 \end{cases}$

and hence

$$5x_1 + 3\,\text{Min}\left(\frac{36-2x_1}{3}, 8\right) = \begin{cases} 5x_1 + 24, & \text{for } 0 \le x_1 \le 6 \\ 3x_1 + 36 & \text{for } 6 \le x_1 \le 12 \end{cases}$$

which achieve its maximum when $x_1 = 12$.

It then follows that $x_1^* = 12$ and $x_2^* = \text{Min}\left(\frac{36-2x_1^*}{3}, s_3\right) = \text{Min}(4, 8) = 4$

Thus the optimal solution of the problem is

$x_1 = 12, x_2 = 4$

and $\quad$ Max $z = 72$

29.12. Reduction in Dimensionality

We have already noted that the greatest obstacle to the use of dynamic programming is the size of the state vector. If the number of state variables is more than two, a large number of values must be computed and stored which may be beyond the range of available computers. To overcome this difficulty, several methods have been suggested. One of the most powerful of these methods is the use of Lagrange multipliers that reduces the number of state variables to a manageable size in many cases. The procedure is based on Everett's method [152] of using Lagrange multipliers to solve constrained optimization problems.

Everett has shown that Lagrange multipliers can be used to solve a general class of problems.

$$\begin{aligned} &\text{Max} && f(X) \\ &\text{Subject to} && g_i(X) \le b_i,\ i = 1, 2, \ldots m. \\ & && X \in S \subset R^n. \end{aligned} \qquad (29.22)$$

Theorem 29.1

If X_0 is an optimal solution of the problem,

$$\text{Maximize } L(X, \lambda) = f(X) - \sum_{i=1}^{m} \lambda_i g_i(X)$$

$$\text{Subject to} \qquad X \in s \qquad (29.23)$$

for a set of real nonnegative Lagrange multipliers $(\lambda_1, \lambda_2, \ldots \lambda_m) = \lambda^T$, then X_0 is an optimal solution of the problem.

Maximize $f(X)$

Subject to $g_i(X) \leq g_i(X_0)$, $i = 1, 2, \ldots m$.

$X \in S$ (29.24)

Proof: Since X_0 maximizes $L(X, \lambda)$ over S, we have,

$$f(X_0) - \sum_{i=1}^{m} \lambda_i g_i(X_0) \geq f(X) - \sum_{i=1}^{m} \lambda_i g_i(X)$$

for all $X \in s$.

$$f(X_0) - f(X) > \sum_{i=1}^{m} \lambda_i \; [g_i(X_0) - g_i(X)] \geq 0$$

for all $X \in S$, satisfying $g_i(X) \leq g_i(X_0)$, since all $\lambda_i \geq 0$.

Hence X_0 is an optimal solution of (29.24).

Now, if $g_i(X_0) = b_i$, $i = 1, 2, \ldots m$, X_0 is also an optimal solution, an optimal solution of the original problem.

It should be noted that the problem considered is quite general since the functions involved and the set S are completely arbitrary. No assumptions such as continuity, differentiability or convexity are made for the functions and there are no restrictions on the set S. S, for example may be discrete or continuous.

Thus the Lagrange multipliers method is applicable to a wider class of problems than the class of problems that can be formulated as dynamic programs.

In the above procedure, if with the present set of $\lambda_1, \lambda_2, \ldots \lambda_m$, $g_i(x_0) \neq b_i$, $i = 1, 2, \ldots m$, then another set of λ_i's must be chosen and the new Lagrange function has to be maximized. This process is continued till we find a set of λ_i's, yielding $g_i(x_0) = b_i$, for all i.

The following result acts as a guide in the search of λ-values that will yield $g_i(x^0) = b_i$, for all i.

Theorem 29.2. Let $\lambda = (\lambda_1, \lambda_2, \ldots \lambda_m) \geq 0$ and $\mu = (\mu_1, \mu_2, \ldots \mu_m) \geq 0$ be any two sets of Lagrange multipliers such that

$$\lambda_i = \mu_i, \quad i = 1, 2, \ldots m; \; i \neq k \tag{29.25}$$

$$\lambda_k > \mu_k$$

and $X_{0\lambda}$, $X_{0\mu}$ are the corresponding optimal solution of the problem. (29.23) Then $g_k(X_{0\lambda})$ is monotonically decreasing function of λ_k.

Proof: Since $X_{0\lambda}$ maximizes $L(X, \lambda)$, we have

$$f(X_{0\lambda}) - \sum_{i=1}^{m} \lambda_i g_i(X_{0\lambda}) \geq f(X_{0\mu}) - \sum_{i=1}^{m} \lambda_i g_i(X_{0\mu})$$

and since $X_{0\mu}$ maximizes $L(X, \mu)$, we have

$$f(X_{0\mu}) - \sum_{i=1}^{m} \mu_i g_i(X_{0\mu}) \geq f(X_{0\lambda}) - \sum_{i=1}^{m} \mu_i g_i(X_{0\lambda})$$

Adding these inequalities and rearranging, we get

$$\sum (\lambda_i - \mu_i)\left[g_i(X_{0\lambda}) - g_i(X_{0\mu}\right] \leq 0$$

Hence by (29.25)

$$g_k(X_{0\lambda}) - g(X_{0\mu}) \leq 0$$

Theorom 29.2 implies that if we want to decrease the value of $g_k(x)$ at an optimal solution of (29.23), then λ_k should be increased, keeping the other λ's fixed. Thus interpolation and extrapolation can be used to determine the desired set of λ's. However, there is no guarantee that there will exist a set of Lagrange multipliers such that one can find an optimal solution of (29.23) that satisfies $g_i(x) = b_i$ for all i. (Exercise 18) Everett has shown that if the variables are continuous and the function to be maximized is concave, then the Lagrange multipliers method will work.

Since the restriction on the memory capacity of computers is such that it is preferable to carry out a large number of one-dimensional problems rather than one multidimensional problem, Lagrange's dimensionality reducing procedure discussed above often permits us to treat problems which would otherwise escape us. To preserve the advantage of dynamic programming, the Lagrange and dynamic programming can be synthesized by treating some of the constraints with Lagrange multipliers and the remainder with state variables. The choice of the number of Lagrange multipliers to be introduced depends upon the individual problem, the type of computer available and the time available for computation.

To illustrate the reduction of state variables using Lagrange multipliers method, we consider the following example

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}(x_1^2 + x_2^2) \\ \text{Subject to} \quad & x_1 + x_2 = 8 \\ & x_1 + 2x_2 \leq 10 \\ & x_1, x_2 \geq 10 \end{aligned} \tag{29.26}$$

This is a two-stage problem with two-state variables.

Following the procedure of section 29.12 we transform the problem into an equivalent problem with one constraint of the problem with a Lagrange multiplier, we have

$$\text{Minimize} \quad F_1(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + \lambda(x_1 + 2x_2) \tag{29.27}$$

$$\text{Subject to} \quad x_1 + x_2 = 8 \tag{29.28}$$

$$x_1, x_2 \geq 0$$

We solve the problem for a fixed value of λ by the procedure described in section 29.10 and find the value of $x_1 + 2x_2$ at this point. The process is then repeated for successive values of λ until $x_1 + 2x_2 = 10$.

Let $\lambda = 0$. The problem then reduces to

$$\text{Minimize} \quad F_0(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$$

$$\text{Subject to} \quad x_1 + x_2 = 8$$

$$x_1, x_2 \geq 0$$

The optimal solution of this problem is found it to be $x_1^* = x_2^* = 4$.

Clearly, $x_1^* + 2x_2^* = 12 > 10$

Using Theorem 29.2, we find that λ should be increased.

Let us take $\lambda = 5$. The optimal solution of the problem is then given by $x_1^* = 6.5$, $x_2^* = 1.5$. Then $x_1^* + 2x_2^* = 9.5 < 10$. The desired value of λ, therefore should lie in the interval (0, 5). By linear interpolation, the new value of λ is estimated to be equal to 4.

With $\lambda = 4$, the optimal solution of the problem is given to be $x_1^* = 6$, $x_2^* = 2$, that yields $x_1^* + 2x_2^* = 10$. Hence $x_1^* = 6$, $x_2^* = 2$ is the optimal solution to the given problem.

Thus, by treating one of the constraints with a Lagrange multiplier one of the state variables is eliminated and the computations and storage of optimal data is reduced considerably.

29.13. Stochastic Dynamic Programming

All the decision processes considered in previous sections, had the property that the outcome of any decision was uniquely determined by the choice of this decision. Such processes are called deterministic. There are however many multi-stage processes arising from realistic situations which do not have this property. There are nondeterministic processes in which for each choice of a decision, the outcome may be a random variable having a probability distribution. We call such processes as stochastic.

29.13.1 As an illustration, let us first consider a rephrased version of Bellman's gold mining problem. [234]

Suppose that we are concerned with using a bomber to inflict maximum damage on some enemy. The bomber has two possible enemy targets, A and B to attack. A raid on target A results either in a fraction r_1 of the enemy's resources at A being destroyed or in the bomber being shot down before inflicting any damage, the probability of the bomber surviving a mission to A being p_1. Target B is similarly associated with a fraction r_2 and a probability p_2. The enemy's resources initially are x at A and y at B. The problem now is to determine a policy of attack that will maximize the total expected damage to the enemy resources.

Let us define

$f_N(x,y)$ = expected damage if the optimal policy is followed when a maximum number of N raids is possible and the system starts with resources x at A and y at B.

Then,

$$f_1(x, y) = \text{Max}\,[p_1 r_1 x,\ p_2 r_2 y]. \tag{29.29}$$

Now, if the N-raid policy starts with an attack on A, then by the principle of optimality, the total expected damage is

$$f_A(x, y) = p_1[r_1 x + f_{N-1}[(1-r_1)x,\ y)] \tag{29.30}$$

and if the target B is attacked first, the total expected damage is

$$f_B(x, y) = p_2[r_2 y + f_{N-1}[x,\ (1-r_2)y)] \tag{29.31}$$

Since we wish to maximize the total expected damage from N-raids, the basic recurrence relation is

$$\begin{aligned} f_n(x, y) &= \text{Max}[f_A(x, y),\ f_B(x, y)] \\ &= \text{Max}\,[p_1\,\{r_1 x\, f_{N-1}\,((1-r_1)\,x,\ y)\},\ p_2\{r_2 y + f_{N-1}(x,\ (1-r_2)y)\}] \end{aligned} \tag{29.32}$$

29.13.2. Stochastic Inventory Problems

Let us consider a simple problem of stocking and supplying a single item to meet in an unknown demand, whose probability distribution is only known. It is assumed that orders are made at each of a finite set of equally spaced times and immediately fulfilled. If the supply is less than the demand observed, a further supply is made to satisfy the demand as far as possible. Any excess of demand over supply then incurs a penalty cost.

Suppose that the following functions are known.

$\phi(s)\,ds$ = the probability that the demand will lie between s and s + ds.

c (z) = the cost of ordering z items to increase the stock level.

k (z) = the cost of ordering z items to meet an excess, z of demand over supply, the penalty cost.

It is further assumed that the functions are independent of time.

Let x denote the stock level at the initial stage and our aim is to determine an ordering policy which minimizes the expected cost of carrying out an N-stage process.

Let us define

$f_n(x)$ = the optimal cost for an n-stage process starting with an initial stock x and using an optimal ordering policy.

Suppose that at the first stage, a quantity y-x is ordered to bring the stock up to the level y.

Then

$f_1(x)$ = minimum expected cost at the first stage,

$$= \underset{y \geq x}{\text{Min}} \left[c(y-x) + \int_y^\infty k(s-y)\phi(s)ds \right] \tag{29.34}$$

Following the principle of optimality, we then have the functional equation

$$f_n(x) = \underset{y \geq x}{\text{Min}} \left[c(y-x) + \int_y^\infty k(s-y)\phi(s)ds \right.$$

$$\left. + f_{n-1}(0) \int_y^\infty \phi(s)ds + \int_0^y f_{n-1}(y-s)\phi(s)ds \right] \tag{29.35}$$

$$n = 2, 3, \ldots N.,$$

considering the different possible cases of an excess of demand over supply and supply over demand.

Dynamic programming techniques may also be used in many other cases of inventory problems which arise under various assumptions regarding cost function (ordering and penalty), distribution function or involving lag in time of supply. For details the reader is referred to Bellman [44].

29.14. Infinite Stage Process

If the number of stages in a multi-stage decision process approaches infinity, it becomes an infinite stage process. When there are a very large member of stages remaining and there is regularity in the stage returns and transformations, we might expect the optimal decision to be independent of the particular stage number. Then in place of sequence of equations (29.2), the single equation.

$$f^*(s) = \underset{x \in s}{\text{Max}} \left[R(s,x) + f^*(s^1) \right] \tag{29.36}$$

might serve as a good approximation.

Now, the question arises whether the equation possesses a finite solution and if so, is the solution unique? Bellman[44] has shown that under certain assumptions a unique solution to (29.36) does exist and discusses this problem in great detail.

Although, for some simple problems the steady state solutions can be obtained easily, in general, solving the infinite stage optimization equation is difficult. Bellman has shown how methods of successive approximation can be used to solve this type of equation.

An infinite stage process also arises when the stages correspond to time periods. Here, the horizon is finite but the time periods are very small. In the limit, as the size of the time periods approaches zero, we assume that decisions are made continuously. Thus, for any finite horizon there will be an infinite number of decisions.

For example, consider the problem of a missile fitting a target in a specified (finite) time internal. Theoretically, the target has to be observed and commands to the missile for changing its direction and speed have to be given continuously. Thus, an infinite number of decisions have to be made in a finite time interval. Since a stage has been defined as a point where decisions are made, this problem will be a continuous infinite stage problem.

Thus, the model of a continuous multi-stage process can be written as

$$\text{Max} \int_{t_1}^{t_2} f(t, x, y)\,dt$$

$$\text{Subject to} \quad \frac{dx}{dt} = g(t, x, y), \quad t_1 \le t \le t_2 \tag{29.37}$$

$$x_1 = x(t_1) = k$$

where x is the state variable and y, the decision variable.

The determination of a function to optimize an integral is a problem in the calculus of variations. However, the analytical solutions, using calculus of variations, cannot be obtained except for very simple problems. The dynamic programming approach, on the otherhand, provides a very efficient numerical approximation procedure for solving continuous multi-stage problem.

29.15. Exercises

1. Use dynamic programming to show that $\sum_{i=1}^{N} p_i \log p_i$, subject to $\sum_{i=1}^{N} p_i = 1$, $p_i \ge 0$, $i = 1, 2, \ldots N$ is minimum when $p_1 = p_2 = \ldots = p_N = \frac{1}{N}$.
2. Consider the cargo-loading problem presented in Section 29.4. Suppose that in addition to the weight limitation w, there is also the volume restriction Q, where q_i is the volume per unit of item i. Obtain the dynamic programming formulation of the problem.
3. Solve the following linear programming problem using dynamic programming

   Maxmize $7x_1 + 8x_2$

   Subject to $2x_1 + 5x_2 \le 15$

   $x_1 + 2x_2 \le 8$

   x_1, x_2 nonnegative integers.
4. Solve the following linear program by the dynamic programming technique.

   Maxmize $3x_1 + 4x_2$

   Subject to $x_1 + 6x_2 \le 6$

   $2x_1 + x_2 \le 4$

$$x_1, x_2 \geq 0.$$

5. Solve the following linear program by the dynamic programming technique.

$$\begin{aligned} \text{Maxmize} \quad & 4x_1 + 2x_2 \\ \text{Subject to} \quad & 3x_1 + x_2 \leq 42 \\ & x_1 + x_2 \leq 21 \\ & 3x_1 + 2x_2 \leq 48 \\ & x_1, x_2 \geq 0. \end{aligned}$$

6. Formulate the functional equations of dynamic programming for the problem

$$\begin{aligned} \text{Minimize} \quad & \sum_{i=1}^{N} x_i^p, \; p > 0 \\ \text{Subject to} \quad & \sum_{i=1}^{N} x_i \geq b, \; b > 0 \\ & x_i \geq 0, \; i = 1, 2, \ldots N. \end{aligned}$$

7. Solve the following problems by the dynamic programming technique.

(a)
$$\begin{aligned} \text{Minimize} \quad & x_1^2 + x_2^2 + x_3^2 \\ \text{Subject to} \quad & x_1 + x_2 + x_3 \geq 15, \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

(b)
$$\begin{aligned} \text{Maximize} \quad & x_1^2 + 2x_2^2 + 4x_3 \\ \text{Subject to} \quad & x_1 + 2x_2 + x_3 \leq 8 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

8. Use dynamic programming to solve the problem

$$\begin{aligned} \text{Maximize} \quad & \prod_{i=1}^{N} x_i \\ \text{Subject to} \quad & \sum_{i=1}^{N} x_i = c \\ & x_i \geq 0, \; i = 1, 2, \ldots N. \end{aligned}$$

9. Formulate the functional equations of dynamic programming for the problem

$$\begin{aligned} \text{Minimize} \quad & \sum_{i=1}^{N} x_i^p, \; p > 0 \\ \text{Subject to} \quad & \prod_{i=1}^{N} x_i = c \\ & x_i \geq 1, \; i = 1, 2, \ldots N. \end{aligned}$$

10. Solve the following problem by the dynamic programming technique.

$$\text{Maximize} \quad \sum_{i=1}^{4}(4x_i - ix_i^2)$$

$$\text{Subject to} \quad \sum_{i=1}^{4} x_i = 10$$

$$x_i \geq 0,\ i = 1, 2, 3, 4$$

11. There are n machines available and each of them can do two jobs. If x of them do the first job, they produce goods worth 3x and if x of them do the second job, lthey produce goods worth 2.5x. The machines are subject to attrition, so that after doing the first job only 1/3x out of x remain available for further work and if they were doing the second job, the available number is 2/3x. The process is repeated with the remaining machines for two more stages. Find the number of machines to be allocated to each job at each stage in order to maximize profit.

12. A manufacturing firm stocks certain basic material every month for a smooth functioning of its production schedule. The purchase price p_n and the demand forecast d_n for the next six months by the management are given below

Month (n)	:	1	2	3	4	5	6
Purchase price (p_n)	:	11	18	13	17	20	10
Demand (d_n)	:	8	5	3	2	7	4

The basic material is purchased at the beginning of each month.
Due to the limited space, the wareouse cannot hold more than 9 units of the basic material. When the initial stock is 2 units and the final stock is required to be zero, find by the use of dynamic programming, an ordering policy for the next 6 months so as to minimize the total purchase cost.

13. A company has 6 salesmen and 3 market areas A, B, C. It is desired to determine the number of salesmen to allocate to each market area to maximize profit. The following table gives the profit from each market area as a function of the number of salesmen allocated.

	Number of salesmen						
Market area	0	1	2	3	4	5	6
A	38	41	48	58	66	72	83
B	40	42	50	60	66	75	82
C	60	64	68	78	90	102	109

Use the dynamic programming technique to solve the above problem.

14. A man is engaged in buying and selling identical items. He operates from a warehouse that can hold 500 items. Each month he can sell any quantity that he chooses up to the stock at the begining of the month. Each month, he can buy as much as he wishes for delivery at the end of the month so long as his stock does not exceed 500 items. For the next four months he has the following forecasts of cost and sale prices.

Month	:	1	2	3	4
Cost	:	27	24	26	28
Sale price	:	28	25	25	27

If he currently has a stock of 200 items, what quantities should he sell and buy in the next four months in order to maximize his profit? Find the solution using dynamic programming.

15. Consider the transportation problem with m origins and n destinations. Let a_i be the amount available at origin i, i = 1, 2,...m and let d_j be the amount demanded at destination j, j = 1, 2,...n, where $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} d_j$.
Assuming that the cost of transporting x_{ij} units from origin i todestination j is $g_{ij}(x_{ij})$, formulate the problem as a dynamic programming model.

16. Consider a transportation problem with two origins and five destinations. The availabilities at the origins, the demands at the destinations and the transportation costs as a function of the number of units x_{ij} transported are given in the following table.

		Destination 1	2	3	4	5	Available
Origin	1	$3x_{11}^2$	$4x_{12}$	$2x_{13}^2$	$5x_{14}^{1/2}$	$3x_{15}$	10
	2	$4x_{21}$	$2x_{22}^2$	$5x_{23}^{1/2}$	$3x_{24}$	$2x_{25}^2$	15
Demand		7	3	5	8	2	

Find an optimal solution of the problem by using the dynamic programming technique.

17. Use dynamic programming to find the shortest route for travelling from city A to city E whose distance matrix (d_{ij}) is given as

$$(d_{ij}) = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{pmatrix} 0 & 22 & 7 & \infty & \infty \\ 22 & 0 & 12 & 10 & 22 \\ 7 & 12 & 0 & \infty & 42 \\ \infty & 10 & \infty & 0 & 8 \\ \infty & 22 & 42 & 8 & 0 \end{pmatrix} \end{array}$$

18. Show that Lagrange multiplier method fails to solve the problem.

Maximize $3x_1 + 2x_2$

Subject to $x_1 - x_2 = 0$

$x_1 + x_2 \leq 3,$

$x_1, x_2 \geq 0$, x_1, x_2 integers.

19. Using dynamic programming technique, solve the following reliability problem. For notations see section 29.9.

	j = 1		j = 2		j = 3	
m_j	$\phi_1(m_1)$	c_1m_1	$\phi_2(m_2)$	c_2m_2	$\phi_3(m_3)$	c_3m_3
1	.5	2	.7	3	.6	1
2	.7	4	.8	5	.8	2
3	.9	5	.9	6	.9	3

20. To conduct a research project on a certain engineering problem, three research teams A, B and C are engaged who are trying three different approaches to solve the problem. The probability that the respective teams will not succeed is estimated as 0.80, 0.60 and 0.40 respectively, so that the probability that all three teams will fail is (0.80) (0.60) (0.40) = 0.192. To minimize this probability, it has been decided to assign two more top scientists among the three teams.

The following table gives the estimated probability that the respective teams will fail when 0, 1, 2 additional scientists are added to that team.

		Team		
		A	B	C
Number of new scientists	0	0.80	0.60	0.40
	1	0.50	0.40	0.20
	2	0.30	0.20	0.15

How should the two additional scientists be allocated to the teams?

Bibliography

1. Abadie, J (ed), Nonlinear Programming, North Holland Publishing company, Amsterdam, 1967.
2. Abadie, J, "On the Kuhn–Tucker Theorem," in [1], 1967.
3. Abadie, J (ed), Integer and Nonlinear Programming, North Holland Publishing Company, Amsterdam, 1970.
4. Abadie, J and J Carpenter, "Generalization of the Wolfe Reduced Gradient Method to the case of Nonlinear constraints," in optimization, R Fletcher (ed), Academic Press, London, 1969.
5. Abadie, J and AC Williams, "Dual and Parametric Methods in Decompostion," in Recent Advances in Mathematical Programming, RL Graves and P Wolfe (eds), McGraw-Hill Book Company, New York, 1963.
6. Aggarwal, SP and Kanti Swarup, "Fractional Functionals Programming with a Quadratic Constraints," Operations Research Vol. 14, pp 950-956, 1966.
7. Aggarwal, SP and IC Sharma," Maximization of the Transimission Rate of a Discrete Constant Channel," Unternehmensforschung Vol. 11, pp 152-155, 1970.
8. Aggarwal, SP and PC Saxena, "Duality Theorems for Nonlinear Fractional Programs," ZAMM Vol. 55, pp 523-524, 1975.

8a. Aggarwal, SP and PC Saxena. "A class of Fractional Programming Problems", New Zealand, Operational Research, Vol. 7, pp. 79-90, 1979.

9. Allen, RGD, Mathematical Economies, McMillan & Company, London, 1957.
10. Almogy, Y and O Levin, "Parametric Analysis of a Multistage Stochastic Shipping Problem", in Operational Research, J Lawrence (ed), Tavistock Publications, London, 1970.
11. Almogy, Y and O Levin, "A Class of Fractional Programming Problems," Operations Research, Vol. 19, pp 57-67, 1971.
12. Anstreicher, KM, "Analysis of a Modified Karmarkar Algorithm for Linear Programming," Working Paper Series B, No 84, Yale School of Organization and Management, Box 1A, New Haven, CT06520, 1985.

13. Anstreicher, KM, "A Strenthened Acceptance Criterion for Approximate Projection in Karmarlar's Algorithm" Operations Research Letters Vol. 5, pp 211-214, 1986.

14. Apostol, TM, Mathematical Analysis, Addison-Wesley, Reading, Mass, 1957.

15. Arrow, KJ and AC Enthoven, "Quasi-Concave Programming," Econometrica Vol. 29, pp 779-800, 1961.

16. Arrow, KJ, L Hurwicz, and Huzawa (eds), Studies in Linear and Non Linear Programming, Stanford University Press, Stanford, 1958.

17. Arrow, KJ, L Hurwicz, and Huzawa, "Constraint Qualifications in Maximization Problems", Naval Research Logistics Quarterly Vol. 8, pp 175-191, 1961.

18. Avriel M, Nonlinear Programming: Analysis and Methods, Prentice Hall, Englewood Cliffs, NJ 1976.

19. Avriel M and I Zang, "Generalized Convex Functions with Applications to Nonlinear Programming" in Mathematical Programs for Activity Analysis, P Van Moeseki (ed) 1974.

19a. Aylawadi, D.R., "Duality for Homogeneous Fractional Programming with Nonlinear Constraints", J. Math, Sciences, Vol. 12-13, pp 29-32, 1978.

20. Balas E, "Solution of Large Scale Transportation Problems Through Aggregation," Operations Research Vol. 13, pp 82-93, 1965.

21. Balas E, "The Dual Method for the Generalized Transportation Problem" Management Science, Vol. 12, pp 555-568, 1966.

22. Balas E, "Nonconvex Quadratic Programming via Generalized Polars," SIAM J Applied Mathematics Vol. 28, pp 335-349, 1975

23. Balinski ML, "Fixed Cost Transportation Problems," Naval Research Logistics Quarterly, Vol. 8, pp 41-54, 1961.

24. Balinski ML, "An Algorithm for Finding All Vertices of Convex Polyhedral Sets," SIAM J Applied Mathematics Vol. 9 pp 72-88, 1961.

25. Balinski ML and P Wolfe, (eds), Non Differentiable Optimization, Mathematical Programming study, No 2, American Elsevier, New York, 1975.

26. Barankin EW and R Dorfman, "On Quadratic Programming," University of California Publications in Statistics Vol. 2, pp 285-318, University of California Press, Berkeley, California.

27. Bartle RG, The Elements of Real Analysis, John Wiley & Sons, New York, 1976.

28. Bazaraa MS, "A Theorem of the Alternative with Applications to Convex Programming; Optimality, Duality and Stability," J Mathematical Analysis and Applications, Vol. 41, pp 701-715, 1973.

29. Bazaraa MS and JJ Goode, "Necessary Optimality Criteria in Mathematical Programming in the Presence of Differentiability," J Mathematical Analysis and Applications Vol. 40 pp 609-621, 1972.

30. Bazaraa MS and JJ Goode, "On Symmetric Duality in Nonlinear Programming," Operations Research, Vol. 21, pp 1-9, 1973.
31. Bazaraa MS, JJ Goode and CM Shetty, "Optimality Criteria without Differentiablity," Operations Research, Vol. 19, pp 77-86, 1971.
32. Bazaraa MS, JJ Goode and CM Shetty, "A Unified Nonlinear Duality Formulation" Operations Research, Vol. 19, pp 1097-1100, 1971.
33. Bazaraa MS, JJ Goode and CM Shetty, "Constraint Qualifications Revisited," Management Science, Vol. 18, pp 567-573, 1972.
34. Bazaraa MS, JJ Jarvis and HD Sharali, Linear programming and Network Flows, John Wiley and Sons, New York, 1990.
35. Bazaraa MS, HD Sharali and CM Shetty, Non Linear Programming: Theory and Algorithms, John Wiley and Sons, New York, 1993.
36. Beale EML, "An Alternative Method for Linear Programming", Proceedings of the Cambridge Philosophical Society, Vol. 50, pp 513-523, 1954.
37. Beale EML, "On Minimizing a Convex Function Subject to Linear Inequalitites," J Royal Stastical Society, Ser B, Vol. 17, pp 173-184, 1955.
38. Beale EML, "Cycling in the Dual Simplex Algorithms," Naval Research Logistics Quaterly, Vol. 2, pp 269-276, 1955.
39. Beale EML, "On Quadratic Programming. Naval Research Logistics Quaterly Vol. 6, pp 227-244, 1959.
40. Beale EML, "The Simplex Method Using Pseudo-basic Variables for Structured Linear Programming Problems", in Recent Advances in Mathematical Programming, RL Graves and P Wolfe (eds), McGraw-Hill Book Company, NY, 1963.
41. Beale EML, "Numerical Methods," in Nonlinear Programming, J Abadic (ed), 1967.
42. Bector CR, "Programming Problems with convex Fractional Functions," Operations Research, Vol. 16, pp 383-391, 1968.
43. Bector CR, "Duality in Nonlinear Fractional Programming," Zietschrift fiir Operations Research, Vol. 17, pp 183-193, 1973.
44. Bellman R, Dynamic Programming, Princeton University Press, Princeton, NJ, 1957.
45. Bellman R, Adaptive Control Processes: A Guided Tour, Princeton University Press, Princeton, New Jersey, 1961
46. Bellman R, and SE Dreyfus, Applied Dynamic Programming, Princeton University Press, Princeton, NJ, 1962.
47. Benders JF, "Partitioning Procedures for Solving Mixed Variables Programming Problems," Numerische Mathematik, Vol. 4, pp 238-252, 1962.
48. Benson HP, "Finding an Initial Efficient Point for a Linear Multiple Objective Program," J Operational Research Society, Vol. 32, pp 495-498, 1981.

49. Benson HP, "Global Optimization Algorithm for the Nonlinear Sum of Ratios Problem", J Optimization Theory and Applications, Vol. 112, pp 1-29, 2002.

50. Bereanu B, "Distribution Problems and Minimum-Risk Solutions in Stochastic Programming," in Colloquium on Applications of Mathematics in Economics, A Prekopa (ed), Academiai Kiado, Budapest, pp 37-42, 1963.

51. Bereanu B, "On Stochatic Linear Programming," Rev Math Purest Appl, Acad Rep Populaire Roumaine, Vol. 8, pp 683-697, 1963.

52. Berge C, Topological Spaces, MacMillan Company, New York, 1963.

53. Berge C and A Ghouila Houri, Programming, Games and Transportation Networks, John Wiley and Sons, New York, 1965.

54. Birkhoff G and S Maclane, A Survey of Modern Algebra, MacMillan Company, New York, 1953.

55. Bitran GR, "Theory and Algorithms for Linear Multiple Objective Programs with Zero-One variables", Mathematical Programming, Vol. 17, pp 362-390, 1979.

56. Bitran GR and TL Magnanti, "Duality and Sensitivity Analysis for Fractional Programs," Operations Research, Vol. 24, pp 657-699, 1976.

57. Bitran GR and AG Novaes, "Linear Programming with a Fractional Objective Function," Operations Research, Vol. 21, pp 22-29, 1973.

58. Bland RG, D Goldfarb and MJ Todd, "The Ellipsoid Method: A Survey," Operations Research, Vol. 29, pp 1039-1091, 1981.

59. Blum E, and W Oettli, "Direct Proof of the Existence Theorem for Quadratic Programming," Operations Research, Vol. 20, pp 165-167, 1972.

60. Boot JCG, Quadratic Programming, North Holland Publishing Co., Amsterdam, 1964.

61. Bradley SP and SC Frey, Jr, "Fractional Programming with Homogenous Functions," Operations Research Vol. 22, pp 350-357, 1974.

62. Buck RC, Mathematical Analysis, McGrawHill Book Co New York, 1965.

63. Cambini A, L Martein and S Schaible, "On Maximizing a Sum of Ratios", J Information and Optimization Sciences, Vol. 10, pp 65-79, 1989.

64. Carathèodory C, Uber den variabilitatsbereich der koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen, Mathematsche Annälen, Vol. 64, pp 95-115, 1907.

65. Chadha SS, "A Dual Fractional Program", ZAMM, Vol. 51, pp 560-561, 1971.

66. Chandra S and TR Gulati, "A Duality Theorem for a Nondifferentiable Fractional Programmig Problem", Management Science, Vol. 23, pp 32-37, 1976.

67. Charnes A, "Optimality and Degeneracy in Linear Programming," Econometrica, Vol. 20, pp 160-170, 1952.

68. Charnes A and WW Cooper, "The Stepping Stone Method of Explaining Linear Programming Calculations in Transportation Problems," Management Science, Vol. 1, No 1, 1954.
69. Charnes A and WW Cooper, "Nonlinear Power of Adjacent Extreme Point Methods of Linear Programming", Econometrica, Vol. 25, pp 132-153, 1957.
70. Charnes A and WW Cooper, "Chance Constrained Programming," Management Science, Vol. 6, pp 73-79, 1959.
71. Charnes A, and WW Cooper, Management Models and Industrial Applications of Linear Programming, Vols. I and II, John Wiley & Sons, New York, 1961.
72. Charnes A, and WW Cooper, "Programming with Linear Fractional Functionals," Naval Research Logistics Quaterly Vol. 9, pp 181-186, 1962.
73. Charnes A and WW Cooper, "Deterministic Equivalents for Optimizing and Satisficing under Chance Constraints," Operations Research, Vol. 11, pp 18-39, 1963.
74. Charnes A and WW Cooper, "Goal Programming and Multiple Objective Optimization – Part 1", Europeon Journal of Operational Research, Vol. 1, pp 39-54, 1977.
75. Charnes A WW Cooper and RO Ferguson, "Optimal Estimation of Executive Compensation by Linear Programming," Management Science, Vol. 1, pp 138-151, 1955.
76. Charnes A, WW Cooper and KO Kortanek, "A Duality Theory for Convex Programs with Convex Constraints," Bull American Mathematical Society, Vol. 68, pp 605-608, 1962.
77. Charnes A, MJL Kirby and WM Raike, "Solution Theorems in Probabilistic Programming – A Linear Programming Approach," J Mathematical Analysis and Applications, Vol. 20, pp 565-582, 1967.
78. Charnes A and CE Lemke, "The Bounded Variable Problem" ONR Research Memorandum No 10, Graduate School of Industrial Administration Carnegie Institute of Technology, Pittsburgh, Pa, 1954.
79. Charnes A and CE Lemke, "Minimisation of Non linear Separable convex Functionals," Naval Research Logistics Quaterly, Vol. 1, No 4, 1954.
80. Charnes A, T Song and M Wolfe, "An Explicit Solution Sequence and Convergence of Karmarkar's Algorithm," Research Report CCS 501, College of Business Administration, The University of Texas at Austin, Austin, 1984
81. Contini B, "A Stochastic Approach to Goal Programming," Operations Research, Vol. 16, pp 576-586, 1968.
82. Cottle RW, "A Theorem of Fritz John in Mathematical programming," The Rand Corporation Research Memorandum RM-3858-PR, 1963.
83. Cottle RW, "Symmetric Dual Quadratic Programs", Quaterly of Applied Mathematics, Vol. 21, pp 237-243, 1963.

84. Cottle RW and GB Dantzig, "Complementary Pivot Theory of Mathematical Programming," Linear Algebra and Applications, Vol. 1, pp 103-125, 1968.
85. Cottle RW and GB Dantzig, "A Generalization of the Linear Complementarity Problem," Journal of Combinational theory, Vol. 8, pp 79-90, 1970.
86. Cottle RW and CE Lemke (eds), Non linear Programming, American Mathematical Society, Providence, RI 1976.
87. Courant R, Differential and Integral Calculus, Vol. II, Interscience Publishers, New York, 1947.
88. Courant R, and D Hilbert, Methods of Mathematical Physics, Vol. I, Interscience Publishers, New York, 1953.
89. Courtillot M, "On Varying All the Parameters in a Linear Programming Problem and Sequential Solution of a Linear Programming Problem", Operations Research, Vol. 10, No 4, 1962.
90. Cramer H, Mathematical Methods of Statistics, Princeton University Press, Princeton, NJ, 1946.
91. Craven, BD, "A generalization of Lagrange Multipliers," Bull Australian Mathematical Society, Vol. 3, pp 353-362, 1970.
92. Craven BD, "Fractional Programming—A Survey," Opsearch, Vol. 25, pp 165-175.
93. Craven, BD, Control and Optimization, Chapman and Hall, London, 1995
94. Craven BD and BMond, "The Dual of a Fractional Linear Program," Journal of Mathematical Analysis and Applications, Vol. 42, pp 507-512, 1973.
95. Craven BD and B Mond, "Duality for Homogenous Fractional Programming," Cahiers Du Centre D'etudes de Recherche Operationnelle, Vol. 18, pp 413-417, 1976.
96. Dantzig GB, "Maximization of a Linear Function of Variables subject to Linear Inequalities," Chap 21 in [283], 1951.
97. Dantzig GB, "Application of the Simplex Method to a Transportation Problem," Chap 23 in [283], 1951.
98. Dantzig GB, "Computational Algorithm of the Revised Simplex Method" RM 1266, The Rand Corporation, 1953.
99. Dantzig GB, "The Dual Simplex Algorithm, RM 1270, The Rand Corporation, 1954.
100. Dantzig GB, "Composite Simplex–Dual Simplex Algorithm, I," RM 1274, The Rand Corporation, 1954.
101. Dantzig GB, "Upper Bounds, Secondary Constraints and Block Triangularity in Linear Programming," Econometrica, Vol. 23, pp 174-183, 1955.
102. Dantzig GB, "Linear Programming under Uncertainty," Management Science, Vol. 1, pp 197-206, 1955.

103. Dantzig, GB, "Recent Advances in Linear Programming," Management Science, Vol. 2, pp 131-144, 1956.
104. Dantzig GB, " Discrete Variable Extremum Problems," Operations Research, Vol. 5, pp 266-277, 1957.
105. Dantzig GB, "On the Significance of Solving Linear Programming Problems with Some Integar Variables, Economerica, Vol. 28, pp 30-44, 1960.
106. Dantzig GB, "On the Shortest Route through a Network," Management Science, Vol. 6, pp 187-190, 1960.
107. Dantzig GB, "General Convex Objective Forms" in Mathematical Methods in the Social Sciences, K Arrow, S Karlin, and P Suppes(eds), Stanford University Press, Stanford, 1960.
108. Dantzig GB, "Quadratic Programming, A Variant of the Wolfe–Markowitz Algortihms," ORC 61-2, Operations Research Center, University of California, Berkeley, Calif, 1961.
109. Dantzig GB, Linear Pogramming and Extensions, Princeton University Press, Princeton, NJ, 1963.
110. Dantzig GB, "Linear Control Processes and Mathematical Programming," SIAM J control, Vol. 4, pp 56-60, 1966.
111. Dantzig GB, E Eisenberg, and RW Cottle, "Symmetric Dual Nonlinear Programs," Pacific J Mathematics, Vol. 15, pp 809-812, 1965.
112. Dantzig GB, LR Ford, Jr, and DR Fulkerson, "A Primal Dual Algorithm for Linear Programs," in Linear Inequalities and Related systems, Annals of Mathematics Study, No 38, HW Kuhn and AW Tucker (eds), 1956.
113. Dantzig GB, SM Johson and WB White, "A Linear Programming Approach to the Chemical Equilibrium Problem," Management Science, Vol. 5, pp 38-43, 1958.
114. Dantzig GB and A Madansky, "On the Solution of Two-Stage Linear Programs under Uncertainty," Proceedings Fourth Berkeley Symposium on Mathematical Statistics and Probability, J Neyman (ed) Vol. I, 1961.
115. Dantzig GB and W Orchard-Hays, Alternate Algorithm for the Revised Simplex Method Using Product Form for the Inverse, RAND Report RM-1268, The RAND Corporation, 1953.
116. Dantzig GB and A Orden, "Duality Theorems", RAND Report RM-1265, The RAND Corporation, 1953.
117. Dantzig GB, A Orden and P Wolfe, "Generalized Simplex Method for Minimizing a Linear Form under Linear Inequality Restraints," Pacific J Mathematics, Vol. 5, pp 183-195, 1955.
118. Dantzig GB and RM Vanslyke," Generalized Upper Bounded Techniques for Linear Programming," ORC 64-17, and ORC 64-18, Operations Research Center, University of California, Berkeley, Calif, 1964.

119. Dantzig GB and P Wolfe, "Decomposition Principle for Linear Programs," Operations Research Vol. 8, pp 101-111, 1960.
120. Dantzig GB and P Wolfe, "The Decomposition Algortihm for Linear Programming," Econometrica, Vol. 29, pp 767-778, 1961.
121. Dantzig GB and G Infanger, "Multi-Stage Stochastic Linear Programs for Portfolio Optimization", in Proceedings of the Third RAMP Symposium, Tokyo, 1991.
122. Dauer JP and RJ Krueger, "An Iterative Approach to Goal Programming," Operational Research Quaterly Vol. 28, pp 671-681, 1977.
123. Davies D, "Some Practical Methods of Optimization," in Abadie [3], 1970.
124. Dennis JB, Mathematical Programming and Electrical Networks, John Wiley & Sons, New York, 1959.
125. Dinkelbach W, "On Nonlinear Fractional programming," Management Science, Vol. 13, pp 492-498, 1967.
126. Dixon, LCW, Nonlinear Optimization, The English University Press, London, 1972.
127. Dorfman R, Application of Linear Programming to the Theory of the Firm, University of California Press, Berkeley, 1951
128. Dorfman R, PA Samuelson and RM Solowo, Linear Programming and Economic Analysis, McGraw Hill Book Company, New York, 1958.
129. Dorn WS, "Duality in Quadratic Programming," Quarterly of Applied Mathematics, Vol. 18, pp 155-162, 1960.
130. Dorn, WS, "On Lagrange Multipliers and Inequalities," Operations Research, Vol. 9, pp 95-104, 1961.
131. Dorn WS, "Self-dual Quadratic Programs," SIAMJ, Vol. 9, pp 51-54, 1961.
132. Dorn WS, "Linear Fractional Programming", IBM Research Report, RC – 830, 1962.
133. Dorn WS, "Nonlinear Programming – A survey," Management Science, Vol. 9, pp 171-208, 1963.
134. Dreyfus S, "Computational Aspects of Dynamic Programming," Operations Research, Vol. 5, pp 409-415, 1957.
135. Duffin RJ "Infinite Programs" in [292], 1956.
136. Dur M, R Horst and NV Thoai, "Solving Sum of Ratios Fractional Programs using Efficient Points", Optimizaton, Vol. 49, 2001.
137. Dwyer PS, "The Solution of the Hitchcock Transportation Problem with the Method of Reduced Matrices," Engineering Research Institute Report, University of Michigan, Ann Arbor, Michigan, 1955.
138. Eaves BC, "On the Basic Theorem of Complementarity," Mathematical Programming, Vol. 1, pp 68-75, 1971.

139. Eaves, BC, "The Linear Complementarity Problem," Management Science, Vol. 17, pp 612-634, 1971.

140. Ecker JG and NS Hegner, "On Computing an Initial Efficient Extreme Point," Journal of the Operational Research Society, Vol. 29, pp 1005-1007, 1978.

141. Ecker JG, NS Hegner, and IA Kouada, "Generating All Maximal Efficient Faces for Multiple Objective Linear Programs," Journal of Optimization Theory and Applications, Vol. 30, pp 353-381, 1980.

142. Ecker JG and IA Kouada, "Finding Efficient points for Linear Multiple Objective Programs," Mathematical Programming, Vol. 8, pp 375-377, 1975.

143. Ecker JG and IA Kouada, "Finding All Efficient points for Linear Multiple Objective Programs," Mathematical Programming, Vol. 14, pp 249-261, 1978.

144. Egervây E, "On Combinational Properties of Matrices (1931)" translated by HW Kuhn, Paper no. 4, George Washington University Logistics Research Project, 1954.

145. Eisenberg E, "Duality in Homogenous Programming," Proc American Mathematical Society, Vol. 12, pp 783-787, 1961.

146. Eisenberg E, "Supports of a Convex Function," Bull American Mathematical Society, Vol. 68, pp 192-195, 1962.

147. Eisenberg E, "On Cone Functions" in Recent Advances in Mathematical Programming, RL Graves and P Wolfe (eds), 1963.

148. Eisenberg E, "A Gradient Inequality for a Class of Non Differentiable Functions," Operations Research, Vol. 14, pp 157-163, 1966.

149. Evans JP, "On Constraint Qualifications in Nonlinear Programming," Naval Research Logistics Quaterly, Vol. 17, pp 281-286, 1970.

150. Evans JP and FJ Gould, "A Nonlinear Duality Theorem Without Convexity," Econometrica Vol. 40, pp 487-496, 1972.

151. Evans JP and RE Steuer, "A Revised Simplex Method for Linear Multiple Objective Programs," Mathematical Programming, Vol. 5, pp 54-72, 1973.

152. Everett H, "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources," Operations Research, Vol. 11, pp 399-417, 1963.

153. Falk JE and SW Palocsay, "Optimizing the Sum of Linear Fractional functions", in Recent Advances in Global Optimization, CS Floudas and PM Pardalos (eds), Kluwer Academic Publishers, Dordrecht, 1992.

154. Fan K, I Glicksburg and AJ Hoffman, "Systems of Inequalities Involving Convex Functions,' Proceed American Mathematical Society, Vol. 8, pp 617-622, 1957.

155. Farkas J, "Uber die Theorie der einfachen Ungleichungen," Journal für die Reine and Angewandte Mathematik, Vol. 124, pp 1-24, 1902.

156. Feller W, An Introduction to Probality Theory and its Applications, Vol 1, Second Edition, John Wiley & Sons, New York, 1957.

157. Fenchel W, "Convex Cones, Sets and Functions," Lecture Notes, Department of Mathematics, Princeton University, 1953.

158. Ferguson, Allen R, and GB Dantzig, "The Allocation of Aircraft to Routes—An Example of Linear Programming under Uncertain Demand," Management science, Vol. 3, pp 45-73, 1956.

159. Ferland JA, "Mathematical Programming Problems with Quasi-Convex Objective Functions," Mathematical Programming, Vol. 3, pp 296-301, 1972.

160. Fiacco AV and GP McCormick, Nonlinear Programming: Sequential Unconstrained Minimization Techniques, John Wiley & Sons, New York, 1968

161. Finkbeiner DT, II, Introduction to Matrices and Linear Transformations, WH Freeman and Company, London, 1966.

162. Fleming, WH, Functions of Several Variables, McGraw-Hill Book Company, New York, 1965.

163. Fletcher R(ed), Optimization, Academic Press, London, 1969.

164. Fletcher R, "A General Quadratic Programming Algorithm," J Institute of Mathematics and Its Applications, Vol. 7, pp 76-91, 1971.

165. Fletcher R, Practical Methods of Optimization, Second Edition, John Wiley and Sons, New York, 1987

166. Fletcher R, "A General Quadratic Programming Algorithm," J Institute of Mathematics and Its Applications, Vol. 7, pp 76-91, 1971.

167. Flood MM, "On the Hitchcock Distribution Problem", Pacific Journal of Mathematics, Vol. 3, No. 2, 1953.

168. Flood MM, "The Traveling Salesman Problem," Operations Research, Vol. 4, pp 61-75, 1956.

169. Ford LR and DR Fulkerson, "Maximal Flow Through a Network," Canadian Journal of Mathematics, Vol. 8, pp 399-404, 1956.

170. Ford LR and DR Fulkerson, "Solving the Transportation Problem," Management Science, Vol. 3, No. 1, 1956.

171. Ford LR and DR Fulkerson, "A Primal-Dual Algorithm for the Capacitated Hitchcock Problem," Naval Research Logistics Quarterly, Vol. 4, pp 47-54, 1957.

172. Ford LR and DR Fulkerson, Flows in Networks, Princeton University Press, Princeton, NJ, 1962.

173. Frank M and P Wolfe, "Am Algorithm for Quadratic Programming," Naval Research Logistics Quarterly, Vol. 3, pp 95-110, 1956.

174. Frechet M, Recherches Thèoriques Modernes sur la Thèorie des Probabilities, Paris, 1937.

175. Freund RJ, The Introduction of Risk into a Programming Model", Econometrica, Vol. 24, pp 253-263, 1956.

176. Fraund RW and F Jarre, "Solving the Sum of Ratios Problem by an Interior Point Method", J of Global Optimization, Vol. 19, 2001.

177. Frisch RAK, "The multiplex method for Linear and Quadratic Programming," Memorandum of Social Institute, University of Oslo, Norway, 1957.

178. Frisch RAK, "Quadratic Programming by the Multiplex Method in the General Case where the Quadratic Form may be Singular," Memorandum of Economic Institute, University of Oslo, Norway, 1960.

179. Gacs P and L Lovasz, "Khachian's Algorithm for Linear Programming," Mathematical Programming Study 14, North-Holland, Amsterdam, pp 61-68, 1981.

180. Gal T, "A General Method for Determining the set of All Efficient Solutions to a Linear Vector Maximum Problem," European Journal of Operational Research Vol. 1, pp 307-322, 1977.

181. Gale D, "Convex Polyhedral Cones and Linear Inequalitites," in [283], 1951.

182. Gale D, "Neighboring Vertices on a Convex Polyhedron," in [292], 1956.

183. Gale D, The Theory of Linear Economic Model, McGraw-Hill Book Company, New York, 1960.

184. Gale D, "On the Number of Faces of a Convex Polytope," Technical Report No 1, Department of Methematics, Brown University, 1962.

185. Gale D, HW Kuhn and AW Tucker, "Linear Programming and the Theory of Games" in Activity Analysis of Production and Allocation, TC Koopmans (ed), John Wiley & Sons, New York, 1951.

186. Garvin WW, Introduction to Linear Programming, McGraw-Hill Book Co., New York, 1960.

187. Gass SI, Linear Programming: Methods and Applications, 4th ed, McGraw-Hill Book Co., 1975.

188. Gass SI, "The Dualplex Method for Large-scale Linear Programs," ORC Report 66-15, Operations Research Center, University of California, Berkeley, Calif, 1966.

189. Gass SI and TL Saaty, "The Computational Algorithm for the Parametric Objective Function," Naval Research Logistics Quarterly Vol. 2, pp 39-45, 1955.

190. Gass SI and TL Saaty, "Parametric Objective Function, Part II: Generalization," Operations Research, Vol. 3, pp 395-401, 1955.

191. Gassner Betty J, "Cycling in the Transportation Problem," Naval Research Logistics Quarterly, Vol. 11, No 1, 1964.

192. Gay DM, "A Variant of Karmarkar's Linear Programming Algorithms for Problems in Standard Form," Mathematical Programming, Vol. 37, pp 81-90, 1987.

193. Geoffrion AM, "Strictly Concave Parametric Programming, I, II" Management Science, Vol. 13, pp 244-253, 1966 and Vol. 13, pp 359-370, 1967.

193a Geoffrion, A.M., "Solving Dicriterion Mathematical Programs," Operations Research, Vol. 15, pp 39-54, 1967.

194. Geoffrion AM, "Stochastic Programming with Aspiration or Fractile Criteria," Management Science, Vol. 13, pp 672-679, 1967.

195. Geoffrion AM, "Proper Efficiency and the Theory of Vector Maximization," J Mathematical Analysis and Applications, Vol. 22, pp 618-630, 1968.

196. Geoffrion AM, "Primal Resource—Directive Approaches for Optimizing Nonlinear Decomposable Systems," Operations Research, Vol. 18, pp 375-403, 1970.

197. Geoffrion AM, "Duality in Nonlinear programming: A Simplified Applications-Oriented Development," SIAM Review, Vol. 13, pp 1-37, 1971.

198. Gill PE, W Murray, MA Saunders, JA Tomlin and MH Wright, "On Projected Newton Barrier Methods for Linear Programming and an Equivalence to Karmarkar's Projective Method", Mathematical Programming, Vol. 36, pp 183-209, 1986.

199. Gilmore PC and RE Gomary, "A Linear Programming Approach to the Cutting Stock Problem – Part I," Operations Research, Vol. 9, pp 849-859, 1961.

200. Gilmore PC and RE Gomory, "A Linear Programming Approach to the Cutting Stock Problem – Part 2" Operations Research, Vol. 11, pp 863-867, 1963.

201. Goldfarb D, "Extension of Davidson's Variable Metric Method to Maximization Under Linear Inequality and Equality Constraints," SIAM J Applied Mathematics, Vol. 17, pp 739-764, 1969.

202. Goldfarb D and S Mehrotra, "A Relaxed Version of Karmarkar's Method," Mathematical programming, Vol. 40, pp 289-316, 1988.

203. Goldman AJ and AW Tucker, "Theory of Linear Programming, in [292], pp 53-98, 1956.

204. Goldman AJ and AW Tucker, "Polyhedral Convex Cones," in [292], pp 19-39, 1956.

205. Gol'stein EG, "Dual Problems of Convex and Fractionally-Convex Programming in Functional Spaces," Soviet Math-Doklady, Vol. 8, pp 212-216, 1967.

206. Gomory RE, "An Algorithm for Integer Solutions to Linear Programs", Princeton-IBM, Mathematics Research Project, Tech Report No. 1, 1958 Also, in Recent Advances in Mathematical Programming RL Groves and P Wolfe (eds), McGraw-Hill Book Co, New York, 1963.

207. Gomory RE, "An Algorithm for the Mixed Integer Problem", Research Memorandum RM – 2597, The Rand Corporation, Santa Monica, Calif, 1960
208. Gomory RE, "All-Integer Programming Algorithm," IBM Research Center, RC 189, New York, 1960.
209. Gordan P, Uber die Auflosungen linearer Gleichungen mit reelen coefficienten," Math Annalen, Vol. 6, pp 23-28, 1873.
210. Graves RL, "Parametric Linear Programming", in [211], pp 201-210, 1963.
211. Graves RL and PWolfe (eds), Recent Advances in Mathematical Programming, McGraw-Hill Book Company, New York, 1963.
212. Greenberg HJ and WP Pierskalla, "A Review of Quasi-Convex Functions," Operations Research, Vol. 19, pp 1553-1570, 1971.
213. Griffith RE and RA Stewart, "A Nonlinear Programming Technique for the Optimization of Continuous Processing Systems," Management Science, Vol. 17, pp 379-392, 1961.
214. Gupta SK and CR Bector, "Nature of Quotients, Products and Rational Powers of Convex (Concave) – Like Functions," Mathematics Student, Vol. 36, pp 63-67, 1968.
215. Hadley G, Linear Algebra, Addison-Wesley, Reading, Mass, 1961.
216. Hadley G, Linear Programming, Addison-Wesley, Reading, Mass, 1962.
217. Hadley G, Nonlinear and Dynamic Programming, Addison-Wesley, Reading, Mass, 1964.
218. Halmos PR, Finite – Dimensional Vector Spaces, D Van Nostrand Company, Princeton, NJ, 1958.
219. Hannan EL, "Nondominance in Goal Programming," INFOR, Vol. 18, pp 300-309, 1978.
220. Hannan EL,"On Fuzzy Goal Programming", Decision Sciences, Vol. 12, pp 522-531, 1981.
221. Hanson MA, "A Duality Theorem in Nonlinear Programming with Nonlinear constraints," Australian Journal of Statistics, Vol. 3, pp 64-72, 1961.
222. Hanson MA, "An Algorithm for Convex Programming," Australian Journal of Statistics, Vol. 5, pp 14-19, 1963.
223. Hanson MA, "Duality and Self-Duality in Mathematical Programming," SIAM J Applied Mathematics, Vol. 12, pp 446-449, 1964.
224. Hanson MA, "On Sufficiency of the Kuhn-Tucker Conditions," J Mathematical analysis and Applications, Vol. 80, pp 545-550, 1981.
225. Hanson MA, and BMond, "Further Generalizations of Convexity in Mathematical Programming," J Information and Optimization Sciences, Vol. 3, pp 25-32, 1982.
226. Hartley HO, "Nonlinear Programming by the Simplex Method," Econometrica, Vol. 29, pp 223-237, 1961.

227. Hildreth C, "A Quadratic Programming Procedure," Naval Research Logistics Quarterly, Vol. 4, pp 79-85, 1957.

228. Hiller FS and GJ Lieberman, Introduction to Operations Research, Holden-Day, Inc, San Francisco, 1968.

229. Hirche J, "A Note on Programming Problems with Linear-plus-Linear-Fractional Objective Functions," European Journal of Operational Research, Vol. 89, pp 212-214, 1996.

230. Hitchcock FL, "The Distribution of a Product from Several Sources to Numerous Localities," J Mathematical Physics, Vol. 20, pp 224-230, 1941.

231. Hodgson TJ and TJ Lowe, "Production Lot Sizing with Material Handing Cost Considerations", IEE Transactions, Vol. 14, pp 44-51, 1982.

232. Hoffman AJ, "Cycling in the Simplex Algorithm," National Bureau of Standards, Report No. 2974, 1953.

233. Hoffman AJ, "Some Recent Applications of the Theory of Linear Inequalities to Extremal Combinatorial Analysis," in Proc Sympos Appl Math, Vol. 10, R Bellman and M Hall, Jr (eds), 1960.

234. Houlden BT (ed), Some Techniques of Operational Research, The English Universities Press Ltd, London, 1962.

235. Houthakker, HS, "On the Numerical Solution of Transportation Problem," Operations Research, Vol. 3, No. 2, 1955.

236. Houthakker HS, "The Capacity Method of Quadratic Programming," Econometrica, Vol. 28, pp 62-87, 1960.

237. Howard RA, Dynamic Programming and Markov Processes, John Wiley and Sons, New York, 1960.

238. Huard P, "Dual Programs," in [211], 1963.

239. Huard P, "Resolution of Mathematical Programming with Nonlinear Constraints by the Method of Centres," in [1] 1967.

240. Ibaraki T, "Solving Mathematical Programming Problem with Fractional Objective Function," in Generalized Concavity in Optimization and Economics, S Schaible and WT Ziemba (eds), Academic Press, New York, 1981.

241. Ibaraki T, "Parametric Approaches to Fractional Programs," Technical Report, Kyoto University, 1962.

242. Ignizio JP, Goal Programming and Extensions, Lexington, Mass, DC Heath and Co, 1976.

243. Ignizio JP, "A Review of Goal Programming: A Tool for Multi-objective Analysis," J Operational Research, Vol. 29, pp 1109-1119, 1978.

244. Ignizio JP, "The Determination of a Subset of Efficient Solutions via Goal Programming," Computers and Operations Research, Vol. 8, pp 9-16, 1981.

245. Ignizio JP, Linear Programming in Single & Multiple-Objective Systems, Prentice-Hall, NJ, 1982.

246. Ignizio JP, "A Note on Computational Methods in Lexicographic Linear Goal Programming", J Operational Research, Vol. 34, pp 539-542, 1983.

247. Ignizio JP and JH Perlis, "Sequential Linear Goal Programming: Implementation via MPSX," Computers and Operations Research, Vol. 6, pp 141-145, 1979.

248. Ijiri Y, Management Goals and Accounting for Control, North Holland, Amsterdam, 1965.

249. Isbell JR and WH Marlow, "Attrition Games," Naval Research Logistics Quarterly, Vol. 3, pp 71-93, 1956.

250. Isermann H, "Proper Efficiency and the Linear Vector Maximum Problem," Operations Research, Vol. 22, pp 189-191, 1974.

251. Iserman H, "The Enumeration of the Set of All Efficient Solutions for a Linear Multiple Objective Program," Operational Research Quarterly, Vol. 28, pp 711-725, 1977.

252. Isermann H, "Lexicographic Goal Programming: The Linear Case," in Multiobjective and Stochastic Optimization, GA Lewandowski and AP Wierzbicki (eds), Int Inst for Applied Systems Analysis, Luxemberg, Austria.

253. Jagannathan R, "On Some Properties of Programming Problems in Parametric Form Pertaining to Fractional Programming," Management Science, Vol. 12, pp 609-615, 1966.

254. Jagannathan R, "Duality for Nonlinear Fractional Programs," Zietschrift fur Operations Research, Vol. 17, pp 1-3, 1973.

255. Jagannathan R, "A Sequential Algorithm for a Class of Programming Problems with Nonlinear Constraints," Management Science, Vol. 21, pp 13-21, 1974.

256. Jagannathan R and S Schaible, "Duality in Generalized Fractional Programming via Farkas' Lemma," J Optimization theory and applications, Vol. 41, pp 417-424, 1983.

257. John F, "Extremum Problems with Inequalities as Subsidiary Conditions," Studies and Essays: Courant Anniversary Volume, KO Friedrichs, OE Neugebauer, and JJStoker (eds), Interscience Publishers, New York, 1948.

258. Johnsen E, Studies in Multiobjective Decission Models, Studenlitteratur, Economic Research Center, Lund, Sweden, 1968.

259. Joksch HC, "Programming with Fractional Linear Objective Functions," Naval Research Logistics Quarterly, Vol. 11, pp 197-204, 1964.

260. Jones PC and ES Marwil, "A Dimensional Reduction Variant of Khachiyan's Algorithm for Linear Programming Problems," EG & G, Idaho, POB 1625, Idaho Falls, Idaho 83415, 1980.

261. Jones PC and ES Marwil, "Solving Linear Complementarity Problems with Khachiyan's Algorithm," EG & G Idaho, POB 1625, Idaho Falls, Idaho 83415, 1980.

262. Kall P, Stochastic Linear Programming, Springer-Verlag, New York, 1976.

263. Kambo NS, Mathematical Programming Techniques, Affiliated East West Press, New Delhi, 1991.

264. Kantorovich LV, "Mathematical Methods in the Organization and Planning of Production," Publication House of the Leningrad State University, 1939 Translated in Management Science, Vol. 6, pp 366-422, 1958.

265. Kantorovich LV, "On The Translocation of Masses," Compt Rend Aead Sci, USSR, Vol. 37, pp 199-201, 1942. Translated in Management Science, Vol. 5, No. 1, 1958.

266. Karamardian S, "Strictly Quasi-Convex (Concave) Functions and Duality in Mathematical Programming," J Mathematical Analysis and Applications, Vol. 20, pp 344-358, 1967.

267. Karamardian S, "Generalized Complementarity Problem," J Optimization Theory and Applications, Vol. 8, pp 161-168, 1971.

268. Karlin S, Mathematical Methods and Theory in Games, Programming and Economics, Vol. 1 and 2, Addison-Wesley, Reading, Mass, 1959.

269. Karmarkar N, "A New Polynomial-Time Algorithm for Linear Programming," Combinatorica, Vol. 4, pp 373-395, 1984.

270. Karmarkar N, JC Lagarias, L Slutsman, and P Wang, "Power Series Variants of Karmarkar-Type Algorithms, AT & T Technical Journal, Vol. 68, pp 20-36, 1989.

271. Kaska, J., "Duality in Linear Fractional Programming", Ewn. Mat. Obzor, Vol. 5, pp 442-453, 1969.

271a. Kaska J and M Pisek, "Quadratic Linear Fractional Programming," Ewn. Mat Obzor, Vol. 2, pp 169-173, 1966.

272. Kelley JE, "Parametric Programming and the Primal-Dual Algorithm," Operations Research, Vol. 7, 1959.

273. Kelley JE, "The Cutting Plane Method for Solving Convex Programs," SIAM Journal Vol. 8, pp 703-712, 1960.

274. Khachiyan LG, "A Polynomial Algorithm for Linear Programming," Doklady Akad Nauk USSR, 244, 1093-1096. Translated in Soviet Math. Doklady, Vol. 20, pp 191-194, 1979.

275. Klee V and GJ Minty, "How Good is the Simplex Algorithm?," in Inequalities III, O Shisha (ed), Academic Press, New York, pp 159-175, 1972

276. Klein D and EL Hannan, "An Algorithm for the Multiple Objective Integer Programming Problem," European Journal of Operational Researh, Vol. 9, pp 378-385, 1982.

277. Klein M, "Inspection-Maintenance-Replacement Schedule under Markovian deterioration" Management Scieence, Vol. 9, pp, 25-32, 1962.
278. Kojima M, "Determining Basic Variables of Optimal Solutions in Karmarkar's New LP Algorithm," Algorithmica, Vol. 1, pp 499-516, 1986.
279. Kojima M, N Megiddo and SMizuno, "A Primal-Dual Infeasible Interior-Point Algorithm for Linear Programming," Mathematical Programming, Vol. 61, pp 263-280, 1993.
280. Konig D, "Grphen and Matrices," Mat Fiz Lapok, 38, p 116, 1931.
281. Konno H and H Yamashita, "Minimization of the Sum and the Product of Several Linear Fractional Functions", Naval Research Logistics, Vol. 46, pp 583-596, 1999.
282. Koopmans TC, "Optimum Utilization of the Transportation System," Econometrica, Vol. 17, (Supplement), 1949.
283. Koopmans TC (ed), Activity Analysis of Production and Allocation, Cowles Commission Monograph 13, John Wiley & Sons, New York, 1951
284. Kornbluth JSH, "A Survey of Goal Programming," Omega, Vol. 1, pp 193-205, 1973.
285. Kornbluth JSH and GR Salkin, "A Note on the Economic Interpretation of the Dual Variables in Linear Fractional Programming," ZAMM, Vol. 52, pp 175-178, 1972,
286. Kornbluth JSH and RE Steuer, "Goal Programming with Linear Fractional Criteria," European Journal of Operational Research, Vol. 8, pp 58-65, 1981.
287. Kornbluth JSH and RE Steuer, "Multiple Objective Linear Fractional Programming," Management Science, Vol. 27, pp 1024-1039, 1981.
288. Kortanek KO and JP Evans, "Pseudo-Concave Programming and Lagrange Regularity," Operations Research, Vol. 15, pp 883-891, 1967.
289. Kovacs LB, "Gradient Projection Method for Quasi-Concave Programming," in Colloquium on Applications of Mathematics to Economics, A Prekopa (ed), Budapest, 1963.
290. Kuhn HW, "The Hungarian Method for the Assignment Problem," Naval Research Logistics Quarterly, Vol. 2, pp 83-97, 1955.
291. Kuhn HW and AW Tucker, "Nonlinear programming," in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, JNeyman (ed), University of California Press, Berkley, Calif, pp 481-492, 1950.
292. Kuhn HW and AW Tucker (ed), Linear Inequalities and Related Systems, Annals of Mathematics Study, No 38, Princeton University Press, Princeton, NJ, 1956.
293. Kuno T, "A Branch and Bound Algorithm for Maximizing the Sum of Several Linear Ratios", J Global Optimization, Vol. 22, pp 155-174, 2002.

294. Kunzi HP, W Krelle and W Oettli, Nonlinear Programming, Bleisdell, Waltham, Mass 1966.

295. Kydland F, "Duality in Fractional Programming," Naval Research Logistics Quarterly, Vol. 19, pp 691-697, 1972.

296. Lasdon LS, "Duality and Decomposition in Mathematical Programming," IEEE Transactions on Systems Science and Cybernetics, Vol. 4, pp 86-100, 1968.

297. Lasdon LS, Optimization Theory for Large Systems, Macmillan, New York, 1970.

298. Lee SM, Goal Programming for Decision Analysis, Auerbach Publishers, Philadelphia, PA, 1972.

299. Lee SM and RL Morris, "Integer Goal Programming Methods," TIMS Studies in the Management Sciences, Vol. 6, pp 273-289, 1977.

300. Lemke CE, "The Dual Method of Solving the Linear Programming Problem," Naval Research Logistics Quarterly, Vol. 1, pp 36-47, 1954.

301. Lemke CE, "A Method of Solution for Quadratic Programs," Management Science, Vol. 8, pp 442-455, 1962.

301a. Lemke CE, "On Complementary Pivot Theory", in Mathematics of the Decision Sciences, G.B. Danting and A.F. Veinoff (eds), 1968.

302. Lemke CE, "Recent Results on Complementarity Problems," in Nonlinear Programming, JB Rosen, OL Mangasarian, and K Ritter (eds), Academic Press, New York, 1970.

303. Leontief WW, The Structure of the American Economy, 1919-1939, Oxford University Press, New York, 1951.

304. Lintner J, "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," Rev Econ and Stat, Vol. 47, pp 13-37, 1965.

305. Luenberger DG, "Quasi-Convex Programming," SIAM Journal Applied Mathematics, Vol. 16, pp 1090-1095, 1968.

306. Luenberger DG, Introduction to Linear and Nonlinear Programming, Second Edition, Addison-Wesley, Reading, Mass 1984.

307. Lustig IJ, Feasibility Issues in a Primal-Dual Interior Point Method for Linear Programming," Mathematical Programming, Vol. 49, pp 145-162, 1991.

308. Madansky A, "Some Results and Problems in Stochastic Linear Programming," The Rand Corporation, Paper P – 1596, 1959.

309. Madansky A, "Inequalities for Srochastic Linear Programming Problems," Management Science, Vol. 6, No. 2, 1960.

310. Madansky A, "Methods of Solution of Linear Programs under Uncertainty," Operations Research, Vol. 10, pp 463-471, 1962.

311. Madansky A, "Dual Variables in Two Stage Linear Programming under Uncertainty," J Math Analysis and Applications, Vol. 6, pp 98-108, 1963.

312. Madansky A, "Linear Programming under Uncertainty," in Recent Advances in Mathematical Programming, RL Graves & P Wolfe (eds), McGraw-Hill, New York, 1963.

313. Mahajan DG and MN Vartak, "Generalization of Some Duality Theorems in Nonlinear Programming," Mathematial Programming, Vol. 12, pp 293-317, 1977.

314. Mangasarian OL, "Duality in Nonlinear Programming," Quarterly of Applied Mathematics, Vol. 20, pp 300-302, 1962.

315. Mangasarian OL, "Nonlinear Programming Problems with Stochastic Objective Functions," Management Science, Vol. 10, pp 353-359, 1964.

316. Mangasarian OL, "Pseudo-Convex Functions," SIAM J – Control, Vol. 3, pp 281-290, 1965.

317. Mangasarian OL, "Nonlinear Fractional Programming," J Operations Research Society of Japan, Vol. 12, pp 1-10, 1969.

318. Mangasarian OL, Nonlinear Programming, McGraw-Hill Book Co New York, 1969.

319. Mangasarian OL, "Optimality and Duality in Nonlinear Programming" in Proceedings of Princeton Symposium on Mathematical Programming, HW Kuhn (ed), 1970.

320. Mangasarian OL, "Convexity, Pseudo-Convexity and Quasi-Convexity of Composite Functions," Cahiers Centre Etudes Recherche Operationnelle, Vol. 12, pp 114-122, 1970.

321. Mangasarian OL, "Linear Complimentarity Problems Solveable by a Single Linear Program," Mathematical Programming, Vol. 10, pp 265-270, 1976.

322. Mangasarian OL and S Fromovitz, "The Fritz John Necessary Optimality Conditions in the Presence of Equality and Inequality Constraints," J Math Analysis and Applications, Vol. 17, pp 37-47, 1967.

323. Mangasarian OL and J Ponstein, "Minimax and Duality in Nonlinear Programming," J Math Analysis and Applications, Vol. 11, pp.504-518, 1965.

324. Manne AS, "Linear Programming and Sequential Decisions", Management Science, Vol. 6, pp 259-267, 1960.

325. Manne AS, "On the Job-Shop Scheduling Problem," Operations Research, Vol. 8, pp 219-223, 1960.

326. Markowitz HM, "Portfolio Selection", Journal of Finance, Vol. 7, pp 77-91, 1952.

327. Markowitz HM, "The Optimization of a Quadratic Function Subject to Linear Constraints, Naval Research Logistics Quarterly, Vol. 3, pp 111-133, 1956.

328. Markowitz HM, Portfolio Selection: Efficient Diversification of Investments, John Wiley and Sons, New York, 1959.

329. Markowitz HM and AS Manne, "On the Solution of Discrete Programming Problems," Econometrica, Vol. 25, pp 84-110, 1957.

330. Martos B, "Hyperbolic Programming," Naval Research Logistics Quarterly, Vol. 11, pp 135-155, 1964.

331. Martos B, "The Direct Power of Adjacent Vertex Programming Methods," Management Science, Vol. 12, pp 241-252, 1965.

332. Martos B, "Quasi-Convexity and Quasi-Monotonicity in Nonlinear Programming," Studia Sci Math Hungarica, Vol. 2, pp 265-273, 1967.

333. Martos B, "Quadratic Programming with a Quasiconvex Objective Function," Operations Research, Vol. 19, pp 87-97, 1971.

334. Martos B, Nonlinear Programming: Theory and Methods, American Elsevier, New York, 1975.

335. McCormick GP, "Second Order Conditions for Constrained Minima," SIAM J Applied Mathematics, Vol. 15, pp 641-652, 1967.

336. McCormick GP, "Anti-Zig-Zagging by Bending," Management Science, Vol. 15, pp 315-320, 1969.

337. McCormick GP, "The Variable Reduction, Method for Nonlinear Programming," Management Science, Vol. 17, pp 146-160, 1970.

338. McCormick GP, "A Second Order Method for the Linearly Constrained Nonlinear Programming Problems," in Nonlinear Programming, JB Rosen, OL Mangasarian, K Ritter (eds), 1970.

339. Mehndiratta SL, "General Symmetric Dual Programs," Operations Research, Vol. 14, pp 164-172, 1966.

340. Mehndiratta SL, "Symmetry and Self-duality in Nonlinear Programming," Numerische Mathematik, Vol. 10, pp 103-109, 1967.

341. Mehndiratta SL, "Self-Duality in Mathematical Programming," SIAM J Applied Mathematics, Vol. 15, pp 1156-1157, 1967.

342. Mehndiratta SL, "A Generalization of a Theorem of Sinha on Supports of a Convex Function," Australian Journal of Statistics, Vol. 11, pp 1-6, 1969.

343. Meister B and W Oettli, "On the Capacity of a Discrete Constant Channel," Information and control, Vol. II, pp 341-351, 1967.

344. Miller CE, "The Simplex Method for Local Separable Programming," in Recent Advances in Mathematical Programming, RL Graves and P Wolfe (eds), 1963.

345. Mishra AD, "Optimality Criteria for a Class of Nondifferentiable Programs," Opsearch, Vol. 12, pp 91-106, 1976.

346. Mjelde KM, "Allocation of Resources according to a Fractional Objective," European J Operational Research, Vol. 2, pp 116-124, 1978.

347. Mjelde KM, Methods of the Allocation of Limited Resources, John Wiley & Sons, 1983.

348. Mond B, "A Symmetric Dual Theorem for Nonlinear Programs," Quarterly of Applied Mathematics, Vol. 23, pp 265-268, 1965.

349. Mond B, "On a Duality Theorem for a Nonlinear Programming Problem," Operations Research, Vol. 21, pp 369-370, 1973.
350. Mond B, "A Class of Nondifferentiable Mathematical Programming Problems," J Math Analysis and Applications, Vol. 46, pp 169-174, 1974.
351. Mond B, "A class of Nondifferentiable Fractional Programming Problems," ZAMM, Vol. 58, pp 337-341, 1978.
352. Mond B and RW Cottle, "Self-Duality in Mathematical Programming," SIAM J Applied Mathematics, Vol. 14, pp 420-423, 1966.
353. Mond B and BD Craven, "A Note on Mathematical Programming with Fractional Objective Function," Naval Research Logistics Quarterly, Vol. 20, pp 577-581, 1973.
354. Mond B and M Schechter, "On a Constraint Qualification in a Nondifferentiable Programming Problem," Naval Research Logistics Quarterly, Vol. 23, pp 611-613, 1976.
355. Monroe AE, Early Economic Thought; Selections from Economic Literature Prior to Adam Smith, Harvard University Press, Cambridge, Massachusetts, 1924.
356. Motzkin TS, Beitrage zur theorie der Linearen Ungleichungen, Dissertation, University of Basel, Jerusalem, 1936.
357. Murty KG, Linear and Combinational Programming, John Wiley & Sons, New York, 1976.
358. Mylander WC, "Finite Algorithms for Solving Quasiconvex Quadratic Programs," Operations Research, Vol. 20, pp 167-173, 1972.
359. Nemhauser GL, Introduction to Dynamic Programming, John Wiley & Sons, New York, 1967.
360. Ohlson JA and WT Ziemba, "Portfolio Selection in a Lognormal Market when the Investor has a Power Utility Function," J Financial Quantitive Anal Vol. 11, pp 57-71, 1976.
361. Orden A, "A Procedure for Handling Degeneracy in the Transportation Problem", Minieograph DCS/Comptroller, Headquarter US Air Force, Washington, DC, 1951.
362. Orden A, "The Transhipment Problem," Management Science Vol. 2, pp 276-285, 1956.
363. Panne C van de, Methods for Linear and Quadratic Programming, North Holland, Amsterdam, 1974..
364. Panne C van de, and A Whinston, "Simplical Methods for Quadratic Programming" Naval Research Logistics Quarterly, Vol. 11, pp 273-302, 1964.
365. Panne C van de, and A Whinston, "The Simplex and the Dual Method for Quadratic Programming," Operational Research Quarterly, Vol. 15, pp 355-388, 1964.

366. Parikh SC, "Equivalent Stochastic Linear programs," SIAM J Applied Mathematics, Vol. 18, pp 1-5, 1970.

367. Pearson ES and HO Hartley, Biometrica Tables for Statisticians, Vol. I, University of Cambridge Press, Cambridge, England, 1958.

368. Philip J, "Algorithms for the Vector Maximization Problem," Mathematical Programming, Vol. 2, pp 207-229, 1972.

369. Ponstein J, "Seven Kinds of Convexity," SIAM Review 9, pp 115-119, 1967.

370. Powell MJD, "A Survey of Numerical Methods for Unconstrained Optimization," SIAM Review 12, pp 79-97, 1970.

371. Powell MJD, "Recent Advances in Unconstrained Optimization," Mathematical Programming, Vol. 1, pp 26-57, 1971.

372. Prèpoka A, Stochastic Programming, Kluwer Publishers, Dordrecht, 1995.

373. Rani O and RN Kaul, "Duality Theorems for a Class of Nonconvex Programming Problems," J Optimization theory and Applications, Vol. 11, pp 305-308, 1973.

374. Rech P, "Decomposition and Interconnected Systems in Mathematical programming", ORC Report 65-31 Operations Research Center, University of California, Berkley, Calif, 1965.

375. Reinfeld NV and WR Vogel, Mathematical Programming, Prentice-Hall, Englewood, Cliffs, NJ, 1958.

376. Riley V and SI Gass, Linear Programming and Associated Techniques: A comprehensive bibliography on linear, nonlinear and dynamic programming, John Hopkins Press, Baltimore, Md 1958.

377. Ritter K, "Ein Verfahren zur Lösung parameterabhängiger nichtlinearer Maximum Problems," Unternehmensforschung, Band 6, pp 149-166, 1962.

378. Ritter K, "A Decomposition Method for Linear Programming Problems with Coupling Constraints and Variables," Mathematics Research Center, US Army, MRC Report No 739, University of Wisconsin, Madison, Wis, 1967.

379. Rockafellar RT, "Duality Theorems for Convex Functions," Bulletin of the American Math Society 70, pp 189-192, 1964.

380. Rockafellar RT, "Duality in Nonlinear programming", in Mathematics of the Decision Sciences, GB Dantzig and A Veinott (eds), American Mathematical Society, Providence, RI, 1969.

381. Rockafellar RT, Convex Analysis, Princeton University Press, Princeton, NJ, 1970.

382. Rosen JB, "The Gradient Projection Method for Nonlinear programming, Part I, Linear Constraints," SIAM J Applied Mathematics, Vol. 8, pp 181-217, 1960.

383. Rosen JB, "The Gradient Projection Method for Nonlinear Programming, Part II Nonlinear Constraints," SIAM J Applied Mathematics, Vol. 9, pp 514-532, 1961.

384. Rosen JB, "Convex Partition Programming" in Recent advances in Mathematical Programming, RL Graves and P Wolfe (eds), McGraw Hill Book Co, New York, 1963.
385. Rosen JB, "Primal Partition Programming for Block Diagonal Matrices," Numerical Mathematics, Vol. 6, pp 250-260, 1964.
386. Rudin W, Principles of Mathematical Analysis, Second Edition, McGraw-Hill Book Co., New York, 1964.
387. Saaty TL, "The Number of Vertices of a Ployhedron," American Monthly Vol. 62, 1995.
388. Saaty TL, "Coefficient Perturbation of a Constrained Extremum," Operations Research, Vol. 7, 1959.
389. Saaty TL and J Bram, Nonlinear Mathematics, McGraw-Hill Book Co, New York, 1964.
390. Saaty TL, and SI Gass, "The Parametric Objective Function, Part 1," Operations Research, Vol. 2, 1954.
391. Samuelson PA, "The Fundamental Approximation Theorem of Portfolio Analysis in terms of Means, Variances and Higher Moments", Rev Econom Studies, pp 537-542, 1974.
392. Savage SL, "Some Theoretical Implications of Local Optimization," Mathematical Programming, Vol. 10, pp 354-366, 1976.
393. Schaible S, "On Factored Quadratic Functions," Zeitschrift für Operations Research, Vol. 17, pp 179-181, 1973.
394. Schaible S, "Maximization of Quasi-Concave Quotients and Products of Finitely Many Functionals," Cahiers du Centre d'Etudes de Recherche Operationnelle, Vol. 16, pp 45-53, 1974.
395. Schaible S, "Parameter-free Convex Equivalent and Dual Programs of Fractional Programming Problems", Zeitschrift fur Operations Research, Band 18, pp 187-196, 1974.
396. Schaible S, "Fractional Programming I, Duality," Management Science, Vol. 22, pp 858-867, 1976.
397. Schaible S, Duality in Fractional Programming: a Unified Approach, Operations Research, Vol. 24, pp 452-461, 1976.
398. Schaible S, "Minimization of Ratios," J Optimization Theory and Applications, Vol. 19, pp 347-352, 1976.
399. Schaible S, "A Note on the Sum of a Linear and Linear-Fractional Function," Naval Research Logistics Quarterly, Vol. 24, pp 691-693, 1977.
400. Schaible S, "Fractional Programming: Applications and Algorithms," European J Operational Research, Vol. 7, pp 111-120, 1981.
401. Schaible S, "Bibliography in fractional programming," Zeitschrift für Operations Research, Vol. 26, pp 211-241, 1982.

402. Schaible S, “Bieriteria Quasi-Concave Programs,” Cahiers du centre d’Etudes de Recherche Operationnelle, Vol. 25, pp 93-101, 1983.

403. Schaible S, “Fractional programming”, Zeitschrift für Operations Research, Vol. 27, pp 39-54, 1983.

404. Schaible S, “Fractional Programming: Some Recent Developments”, J Information and Optimization Sciences, Vol. 10, pp 1-14, 1989.

405. Schaible S and T Ibraki, “Fractional Programming,” European J Operational Research, Vol. 12, pp 325-338, 1983.

406. Schaible S and J Shi, “Fractional programming: the Sum-of-Ratios Case,” Management Science, Vol. 48, pp 1-15, 2002.

407. Schniederjans MJ, Linear Goal Programming, Petrocelli Books, Princeton, NJ, 1984.

408. Schniederjans MJ, Goal Programming Methodology and Applications, Kluwer Publishers, Boston, 1995.

409. Schroeder RG, “Linear Programming Solutions to Ratio Games,” Operations Research, Vol. 18, pp 300-305, 1970.

410. Sengupta JK, “Stochastic Linear Programming with Chance-Constraints,” Internat Econom Rev, Vol. 11, pp 287-304, 1970.

411. Sengupta JK, Stochastic Programming: Methods and Applications, North-Holland, Amsterdam, 1972.

412. Sengupta JK and GTintner, “A Review of Stochastic Linear Programming,” Rev Int Statist Inst, Vol. 39, pp 197-233, 1971.

413. Sharma IC and K Swarup, “On Duality in Linear Fractional Functionals Programming,” Zeitschrift für Operations Research, Vol. 16, pp 91-100, 1972.

414. Shetty CM, “Solving Linear Programming Problems with Variable Parameters,” J Industrial Engineering, Vol. 10, No 6, 1959.

415. Shor NZ, “Cut-off Method with Space Extension in Convex Programming Problems,” Kibernetika, Vol. 1, pp 94-95, 1977. (English translation in Cybernetics, Vol. 13, pp 94-96, 1977).

416. Simmonard M, Linear Programming (translated by WS Jewell) Prentice-Hall, Englewood Cliffs, NJ 1966.

417. Simon HA “Dynamic Programming under Uncertainity with a Quadratic Function,” Econometrica, Vol. 24, pp 74-81, 1956.

418. Simmons DM, Nonlinear Programming for Operations Research, Prentice-Hall, Englewood Cliffs, NJ, 1975.

419. Simmons GF, Introduction to Topology and Modern Analysis, McGraw-Hill Book Co, New York, 1963.

420. Singh, C, “Nondifferentiable Fractional Programming with Hanson-Mond Classes of Functions,” J Optimization Theory and Applications, Vol. 49, pp 431-447, 1986.

421. Singh C, "A Class of Multiple-Criteria Fractional Programming Problems," J Mathematical Analysis and Applications, Vol. 115, 1986.

422. Sinha SM, "Stochastic Programming," PhD Thesis, Universiy of California, Berkley, Calif, January, 1963, Summary in Recent Advances in Mathematical Programming (Title: Programming with Standard Errors in the Constraints and the Objective), RL Graves and P Wolfe (eds), pp 121-122, McGraw-Hill Book Co, 1963.

423. Sinha SM, "Stochastic Programming," Research Report, ORC 63-22, Operations Research Center, University of California, Berkeley, Calif, August, 1963.

424. Sinha SM, "An Extension of a Theorem on Support of a Convex Function," Management Science, Vol. 12, pp 380-384, 1966.

425. Sinha SM, "A Duality Theorem for Nonlinear Programming," Management Science, Vol. 12, pp 385-390, 1966.

426. Sinha SM and DR Aylawadi, "Optimality Criteria for a Class of Nondifferentiable Mathematical Programs," Zamm, Vol. 62, 1982.

427. Sinha SM and DR Aylawadi, "Optimality Conditions for a Class of Nondifferentiable Mathematical Programming Problems," Opsearch, Vol. 19, pp 225-237, 1982.

428. Sinha SM and DR Aylawadi, "Optimality Conditions for a Class of Nondifferentiable Fractional Programming Problems," Indian J Pure and Applied Mathematics, Vol. 14, pp, 167-174, 1983.

429. Sinha SM and CP Bajaj, "The Maximum Capacity Route Through a Set of Specified Nodes," Cahiers du centre d'Etudes de Recherche Operationnelle, Vol. 11, pp 133-138, 1969.

430. Sinha SM and CP Bajaj, "The Maximum Capacity Route Through sets of Specified Nodes," Opsearch, Vol. 7, pp 96-114, 1970.

431. Sinha SM and OP Jain, "On Indefinite Programming Problems," Zeitschrift für Operations Research, Band 18, pp 41-45, 1974.

432. Sinha SM and PC Jha, "Interactive Goal Programming," Int J Management and Systems, Vol. 6, pp 129-138, 1990.

433. Sinha SM and SS Lal, "On Indefinite Quadratic Programming," J Mathematical Sciences, Vol. 3, pp 71-75, 1968.

434. Sinha SM and K Swarup, "Mathematical Programming: A Survey," J Mathematical Sciences, Vol. 2, pp 125-146, 1967.

435. Sinha, SM and GC Tuteja, "On Fractional Programming", Opsearch, Vol. 36, pp 418-424, 1999.

436. Sinha SM and V Wadhwa, "Programming with a Special Class of Nonlinear Functionals," Unternehmensforchung, Band 14, pp 215-219, 1970.

437. Slater M, "Lagrange Multipliers Revisted: A Contribution to Nonlinear Programming," Cowles Commission Discussion Paper, Mathematics 403, 1950.

438. Steuer RE, "Multiple Objective Linear Programming with Interval Criterion Weights," Management Science, Vol. 23, pp 305-316, 1976.

439. Steuer RE, "An Interactive Multiple Objective Linear Programming Procedure," TIMS Studies in the Management Sciences, Vol. 6, pp 225-239, 1977.

440. Steuer RE, "Goal Programming Sensitivity Analysis Using Interval Penalty Weights," Mathematical Programming, Vol. 17, pp 16-31, 1979.

441. Steuer RE, Multiple Criteria Optimization: Theory, Computation and Application, John Wiley & Sons, New York, 1985

442. Steuer RE and EU Choo, "An Interactive Weighted Techebycheff Procedure for Multiple Objective Programming," Mathematical Programming, Vol. 26, pp 326-344, 1983.

443. Stiemke E, "Uber positive losungen homogener linearer Gleichungen," Mathematische Annalen, Vol. 76, pp 340-342, 1915.

444. Stigler GJ, "The Cost of Subsistence," J Farm Econ, Vol. 27, pp 303-314, 1945.

445. Stoer J, "Duality in Nonlinear Programming and the Minimax Theorem," Numerische Mathematik, Vol. 5, pp 371-379, 1963.

446. Swarup K, "Linear Fractional Functionals Programming," Operations Research, Vol. 13, pp 1029-1036, 1965.

447. Swarup K, "Programming with Quadratic Fractional Functionals," Opsearch, Vol. 2, pp 23-30, 1965.

448. Swarup K, "Some Aspects of Linear Fractional Functionals Programming," Australian Journal of Statistics, Vol. 7, pp 90-104, 1965.

449. Swarup K, "Contributions Pertaining to Indefinite Quadratic Programming," Opsearch, Vol. 3, pp 207-211, 1966.

450. Swarup K, "Programming with Indefinite Quadratic Function with Linear Constraints," Cahiers du Centre d'Etudes de Recherche Operationnelle, Vol. 8, pp 132-136, 1966.

451. Swarup K, "Indefinite Quadratic Programming," Cahiers du Centre d'Etudes de Recherche Operationnelle, Vol. 8, pp 217-222, 1966.

452. Swarup K, "Quadratic Programming", Cahiers du Centre d'Etudes de Recherche Operationnelle, Vol. 8, pp 223-233, 1966

453. Swarup K, "Transportation Technique in Linear Fractional Programming," J Royal Naval Sci Ser, Vol. 21, pp 256-260, 1966.

454. Swarup K, "Indefinite Quadratic Programming with a Quadratic Constraint," Economick-Mathematicky Obzor, Vol. 4, pp 69-75, 1968.

455. Swarup K, "Duality in Fractional Programming," Unternehmensforschung, Vol. 12, pp 106-112, 1968.

456. Symonds GH, Linear Programming: The Solution of Refinery Problems, Esso Standard Oil Company, New York, 1955.

457. Taha HA, "Concave Minimization over a Convex Polyhedron," Naval Research Logistics Quarterly, Vol. 20, pp 533-548, 1973.
458. Tamiz M, DF Jones and El-Darzi, "A Review of Goal Programming and its Applications," Annals of Operations Research, Vol. 58, pp 39-53, 1993.
459. Tanabe K, "An Algorithm for the Constrained Maximization in Nonlinear Programming," J Operations Research Society of Japan, Vol. 17, pp 184-201, 1974.
460. Tardos E, "A Strongly Polynominal Algorithm to Solve Combinational Linear Programs", Operations Research, Vol. 34, pp 250-256, 1986.
461. Teterev AG, "On a Generalisation of Linear and Piecewise-Linear Programming," Matekon, Vol. 6, pp 246-259, 1970.
462. Theil H, "A Note on Certainty Equivalence in Dynamic Planning," Econometrica, Vol. 25, pp 346-349, 1957.
463. Theil H, "Some Reflections on Static Programming Under Uncertainty," Weltwirtschaftliches, Vol. 87, pp 124-138, 1961.
464. Theil H, and C van de Panne, "Quadratic Programming as an Extension of Conventional Quadratic Maximization," Management Science, Vol. 7, pp 1-20, 1961.
465. Tintner G, "Stochastic Linear Programming with Applications to Agricultural Economics," Proceed Second Symposium in Linear Programming, HA Antosiewicz (ed), Washington, 1955.
466. Tintner G, "A Note on Stochastic Linear Programming," Econometrica, Vol. 28, pp 490-495, 1960.
467. Tintner G, C Millham, and JK Sengupta, "A Weak Duality Theorem for Stochastic Linear Programming," Unternechmensforschung, Vol. 7, pp 1-8, 1963.
468. Todd MJ, "Exploiting Special Structure in Karmarkar's Linear Programming Algorithm," Mathematical Programming, Vol. 41, pp 97-114, 1988.
469. Todd MJ and BP Burrell, "An Extension of Karmarkar's Algorithm for Linear Programming Using Dual Variables," Algorithmica, Vol. 1, pp 409-424, 1986.
470. Tomlin JA, "An Experimental Approach to Karmarkar's Projective Method for Linear Programming," Ketron Inc Mountain View, CA 94040, 1985.
471. Topkis DM and AF Veinott, "On the Convergence of Some Feasible Direction Algorithms for Nonlinear Programming," SIAM J Control Vol. 5, pp 268-279, 1967.
472. Tucker AW, "Linear and Nonlinear Programming" Operations Research, Vol. 5, pp 244-257, 1957.
473. Tucker AW, "Dual Systems of Homogenous Linear Relations," in (292), 1956.
474. Tuteja GC, "A Note on Programming with a Nonlinear Fractional Program," J Decision and Mathematika Sciences, Vol. 4, pp 77-82, 1999.

475. Tuteja GC, "Programming With the Sum of a Linear and Quotient Objective Function," Opsearch, Vol. 37, pp 177-180, 2000.

476. Tuteja GC, on Mathematical Programming, PhD Thesis, University of Delhi, Delhi, February, 2003.

477. Uzawa H, "The Kuhn–Tucker Theorem in Concave Programming," in [16], 1958.

478. Uzawa H, "The Gradient Method for Concave Programming, II, Global results," in [16] 1958.

479. Uzawa H, "Iterative Methods for Concave Programming," in [16], 1958.

480. Vajda S, Mathematical Programming, Addison-Wesley, Reading, Mass, 1961.

481. Vajda S, "Nonlinear Programming and Duality," in [1], 1967.

482. Vajda S, "Stochastic Programming" in [3], 1970.

483. Vajda S, Probablistic Programming, Academic Press, New York, 1972.

484. Vajda S, Theory of Linear and Nonlinear Programming, Longman, London, 1974.

485. Vajda S, "Tests of Optimality in Constrained Optimization," J Institute of Mathematics and its Applications, Vol. 13, pp 187-200, 1974.

486. Veinott AF, "The Supporting Hyperplane Method for Unimodal Programming," Operations Research, Vol. 15, pp 147-152, 1967.

487. vön Neumann J, "Zur Theoric der Gesellshaftspiele," Mathematische Annalen, Vol. 100, pp 295-320, 1928.

488. vön Neumann J, "Uber ein okonomisches Gleichungssystem und ein Verallgemeinerung des Brouwerschen Fixpunktsatzes," Ergebnisse eines Mathematischen Kolloquiums, No 8, 1937. English translation, "A Model of General Economic Equilibrium," Rev Econ Studies, Vol. 13, pp 1-9, 1945-1946.

489. vön Neumann J, and O Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, Princeton, NJ, 3rd ed, 1953.

490. Wadhwa V, "Programming with Separable Fractional Functionals", J Mathematical Sciences, Vol. 4, pp 51-60, 1969.

491. Wadhwa V, "On Some Specialized Problems in Mathematical Programming, Ph D Thesis, Unversity of Delhi, Delhi, January, 1972.

492. Wadhwa V, "Linear Fractional Programs with Variable Coefficients," Cahiers du centre d' Etudes de Recherche Operationnelle, Vol. 14, pp 223-232, 1972.

493. Wadhwa V, "Parametric Linear Fractional Programming," SCIMA; Journal of Management Sciences and Applied Cybernetics, Vol. 3, pp 21-29, 1974.

494. Wagner HM, "The Dual Simplex Algorithms for Bounded Variables," Naval Research Logistics Quarterly, Vol. 5, pp 257-261, 1958.

495. Wagner HM, "On the Distribution of Solutions in Linear Programming Problems," J American Statistical Association, Vol. 53, pp 161-163, 1958.

496. Wagner HM, "On a Class of Capaciated Transportation Problems," Management Science, Vol. 5, pp 304-318, 1959.
497. Wagner HM, Principles of Operations Research with Applications to Managerial Decisions, Prentice Hall, Englewood Cliffs, NJ, 1969.
498. Wagner HM and JSC Yuan, "Algorithmic Equivalence in Linear Fractional Programming," Management Science, Vol. 14, pp 301-306, 1968.
499. Walras L, Elements of Pure Economics or the Theory of Social Wealth, Translated by W. Jaffe, London; Allen and Unwin, 1954.
500. Wegner P, "A Nonlinear Extension of the Simplex Method," Management Science, Vol. 7, pp 43-55, 1960.
501. Wets RJB, "Programming Under Uncertainty: The Equivalent Convex Program," SIAM J Applied Mathematics, Vol. 14, pp 89-105, 1966.
502. Wets RJB, "Programming Under Uncertainty: The Complete Problem," Z Wahrswcheinlichkeits Theorie und Verwandte, Gebiete, 4, pp 316-339, 1966.
503. Wets RJB, "Characterisation Theorems for Stochastic Programs," Mathematical Programming, Vol. 2, pp 166-175, 1972.
504. Wets RJB, "Stochastic Programs with Recourse: A Basic Theorem for Multistage Problems", Z Wahrswcheinlichkeits Theorie und Verwandte Gebiete, 21, pp 201-206, 1972.
505. Wets RJB, "Stochastic Programs with Fixed Recourse: The Equivalent Deterministic Problem," SIAM Review, Vol. 16, pp 309-339, 1974.
506. Wets RJB, "Solving Stochastic Programs with Simple Recourse II," in Proceed John Hopkins Symposium on System and Information Science, John Hopkins University, Baltimore, 1975.
507. Wets RJB, "Duality Relations in Stochastic Programming," in Symposia Mathe. Atica 19, Academic Press, New York, 1976.
508. Wets RJB, "Stochastic Programs with Chance Constraints: Generalized Convexity and Approximation Issues," in Generalized Convexity, Generalized Monotonicity: Recent Results, JP Crouzeix, JE Martinez-Legaz & M. Volle (eds), Kluwer, Dordrecht, 1998.
509. Whinston A, "A Dual Decomposition Algorithm for Quadratic Programming," Cahiers du Centre d'Etudes de Recherche Operationnelle, Vol. 6, pp 188-201, 1964.
510. Whinston A, "Conjugate Functions and Dual Programs," Naval Research Logistics Quarterly, Vol. 12, pp 315-322, 1965.
511. Whinston, A, "Some Applications of the Conjugate Function Theory to Duality," in [1], 1967.
512. White DJ, Dynamic Programming, Oliver & Boyd, London, 1969.
513. White DJ, "A Linear Programming Analogue, a Duality Theorem and a Dynamic Algorithm," Management Science, Vol. 21, pp 47-59, 1974.

514. White WB, SM Johnson, and GB Dantzig, "Chemical Equilibrium in Complex Mixtures," J. Chem Physics, Vol. 28, pp 751-755, 1958.

515. Wilde DJ, Optimum Seeking Methods, Prentice-Hall, Englewood Chiffs, New Jersey, 1964.

516. Williams AC, "A Treatment of Transportation Problems by Decomposition," SIAM J, Vol. 10, pp 35-48, 1962.

517. Williams AC, "A Stochastic Transportation Problems," Operations Research, Vol. II, pp 759-770, 1963.

518. Williams AC, "On Stochastic Linear Programming," SIAM J Applied Mathematics, Vol. 13, pp 927-940, 1965.

519. Williams AC, Approximation Formulas for Stochastic Linear Programming," SIAM J Applied Mathematics, Vol. 14, pp 668-677, 1966.

520. Williams AC, "Complementary Theorems for Linear Programming," SIAM Review, Vol. 12, pp 135-137, 1970.

521. Williams AC, "Nonlinear Activity Analysis," Management Science, Vol. 17, pp 127-139, 1970.

522. Wilson R, "On Programming Under Uncertainty," Operations Research, Vol 14, pp 652-657, 1966.

523. Wismer DA (ed), Optimization Methods for Large Scale Systems, McGraw-Hill Book Co, New York, 1971.

524. Wolfe P, "The Simplex Method for Quadratic Programming," Econometrica, Vol. 27, pp 382-398, 1959.

525. Wolfe P, "A Duality Theorem for Nonlinear Programming," Quarterly of Applied Mathematics, Vol. 19, pp 239-244, 1961.

526. Wolfe P, "Some Simplex-like Nonlinear Programming Procedures," Operations Research, Vol 10, pp 438-447, 1962.

527. Wolfe P, "Methods of Nonlinear Programming" in [211], 1963.

528. Wolfe P, "A Technique for Resolving Degeneracy in Linear Programming," SIAM J, Vol. 11, pp 205-211, 1963.

529. Wolfe P, "The Composite Simplex Algorithm," SIAM Review, Vol. 7, 1965.

530. Wolfe P, "Methods of Nonlinear Programming," in [1], 1967.

531. Wolfe P, "Convergence, Theory in Nonlinear Programming," in [3], 1970.

532. Wolfe P, "On the Convergence of Gradient Methods Under Constraint," IBMJ Research and Development, Vol. 16, pp 407-411, 1972.

533. Wolfe P, "A Method of Conjugate Subgradients for Minimizing Nondifferentiable functions," in [25] 1975.

534. Wright SJ, "An Infeasible-Interior-Point Algorithm for Linear Complementarity Problems," Mathematical Programming, Vol. 67, pp 29-51, 1994.

535. Wright SJ, Primal-Dual-Interior-Point Methods, SIAM, Philadelphia, 1997.
536. Ye Y, "Karmarkar's Algorithm and the Ellipsoid Method," Operations Research Letters, 6, pp 177-182, 1987.
537. Ye Y, "A Class of Projective Transformations for Linear Programming," SIAM J Computing, Vol. 19, pp 457-466, 1990.
538. Ye Y, "On an Affine Scaling Algorithm for Nonconvex Quadratic Programming," Mathematical Programming, Vol. 52, pp 285-300, 1992.
539. Ye Y, "A Potential Reduction Algorithm Allowing Column Generation," SIAM J Optimization, Vol. 2, pp 7-20, 1992.
540. Ye Y, "Toward Probabilistic Analysis of Interior-Point Algorithms for Linear Programming," Mathematics of Operations Research, Vol. 19, pp 38-52, 1994.
541. Ye Y, Interior Point Algorithms: Theory and Analysis, John Wiley & Sons, New York, 1997.
542. Ye Y, and M Kojima, "Recovering Optimal Dual Solutions in Karmarkar's Polynomial Algorithm for Linear Programming," Mathematical Programming, Vol. 39, pp 305-317, 1987.
543. Ye Y, and E Tse, "An Extension of Karmarkar's Projective Algorithm for Convex Quadratic Programming", Mathematical Programming, Vol. 44, pp 157-179, 1989.
544. Yu PL and M Zeleny, "The Set of all Nondominated Solutions in Linear Cases and a Multicriteria simplex method," J Math Analysis and Applications, Vol 49, pp 430-468, 1975.
545. Zangwill WI, "The Convex Simplex Method," Management Science, Vol. 14, pp 221-283, 1967.
546. Zangwill WI, "Minimising a Function without Calculating Derivatives," Computer Journal, Vol. 10, pp 293-296, 1967.
547. Zangwill WI, "Nonlinear Programming via Penalty Functions," Management Science, Vol. 13, pp 344-358, 1967.
548. Zangwill WI, "The Piecewise Concave Function," Management Science, Vol. 13, pp 900-912, 1967.
549. Zangwill WI, "A Decomposable Nonlinear Programming Approach," Operations Research, Vol. 15, pp 1068-1087, 1967.
550. Zangwill WI, Nonlinear Programming: A Unified Approach, Prentice-Hall, Englewood Cliffs, NJ, 1969.
551. Zeleny M, Linear Multi-Objective Programming, Lecture Notes in Economics and Mathematical Systems No 95, Springer-Verlag, Berlin, 1974.
552. Zelleny M(ed), Multiple Criteria Decision Making: Kyoto 1975; Lecture Notes in Economics and Mathematical systems No 123, Springer-Verlag, Berlin, 1976.

553. Zeleny M, Multiple Criteria Decision Making, McGraw-Hill Book Co, New York, 1982.

554. Zellner A, "Linear Regression with Inequality Constraints on the Coefficients: An Application of Quadratic Programming and Linear Decision Rules," Report No 6109, Econ Inst Netherlands School of Economics, Rotterdam, 1961.

555. Ziemba WT, "Computational Algorithms for Convex Stochastic Programs with Simple Recourse," Operations Research, Vol. 18, pp 414-431, 1970.

556. Ziemba WT, "Transforming Stochastic Dynamic Programming Problems into Nonlinear Programs," Management Science, Vol. 17, pp 450-462, 1971.

557. Ziemba WT, "Duality Relations, Certainty Equivalents and Bounds for Convex Stochastic Programs with Simple Recourse," Cahiers du Centre d' Etudes de Recherche Operationnelle, Vol. 13, pp 85-97, 1971.

558. Ziemba WT, "Stochastic Programs with Simple Recourse," in Mathematical Programming in Theory and Practice, PL Hammer and G Zoutendijk (eds), North-Holland, Amsterdam, 1974.

559. Ziemba WT, C Parkan, and R Brooks-Hill, "Calculation of Investment Portfolios with Risk Free Borrowing and Lending," Management Science, Vol. 21, pp 209-222, 1974.

560. Zimmerman HJ, "Fuzzy Programming and Linear Programming with Several Objective Functions," Fuzzy Sets and Systems, Vol. 1, pp 45-56, 1978.

561. Zionts S, "Programming with Linear Fractional Functions," Naval Research Logistics Quarterly, Vol. 15, pp 449-452, 1968.

562. Zionts S, "A Survey of Multiple Criteria Integer Programming Methods," Annals of Discrete Mathematics, Vol. 5, pp 389-398, 1979.

563. Zoutendijk G, Methods of Feasible Directions, Elsevier, Amsterdam, 1960.

564. Zoutendijk, G, "Nonlinear Programming: A Numerical Survey," SIAM J Control, Vol. 4, pp 194-210, 1966.

565. Zoutendijk G, Mathematical Programming Methods, North Holland, Amsterdam, 1976.

566. Zwart PB, "Nonlinear Programming: Global Use of the Lagrangian," J Optimization Theory and Applications, Vol. 6, pp 150-160, 1970.

567. Zwart PB, "Nonlinear Programming – The Choice of Direction by Gradient Projection," Naval Research Logistics Quarterly, Vol. 17, pp 431-438, 1970.

568. Zwart PB, "Global Maximization of a Convex Function with Linear Inequality Constraints," Operations Research, Vol. 22, pp 602-609, 1974.

This Page Intentionally Left Blank

Index

S

T

U

V