Springer Optimization and Its Applications 153


Asoke Kumar Bhunia
Laxminarayan Sahoo
Ali Akbar Shaikh


# Advanced Optimization and Operations Research


Springer

# Springer Optimization and Its Applications

Volume 153

*Series Editor*
Panos M. Pardalos, University of Florida, Gainesville, FL, USA

*Honorary Editor*
Ding-Zhu Du, University of Texas at Dallas, Richardson, TX, USA

*Advisory Editors*
J. Birge, University of Chicago, Chicago, IL, USA
S. Butenko, Texas A&M University, College Station, TX, USA
F. Giannessi, University of Pisa, Pisa, Italy
S. Rebennack, Karlsruhe Institute of Technology, Karlsruhe, Baden-Württemberg, Germany
T. Terlaky, Lehigh University, Bethlehem, PA, USA
Y. Ye, Stanford University, Stanford, CA, USA

**Aims and Scope**
Optimization has continued to expand in all directions at an astonishing rate. New algorithmic and theoretical techniques are continually developing and the diffusion into other disciplines is proceeding at a rapid pace, with a spot light on machine learning, artificial intelligence, and quantum computing. Our knowledge of all aspects of the field has grown even more profound. At the same time, one of the most striking trends in optimization is the constantly increasing emphasis on the interdisciplinary nature of the field. Optimization has been a basic tool in areas not limited to applied mathematics, engineering, medicine, economics, computer science, operations research, and other sciences.

The series **Springer Optimization and Its Applications (SOIA)** aims to publish state-of-the-art expository works (monographs, contributed volumes, textbooks, handbooks) that focus on theory, methods, and applications of optimization. Topics covered include, but are not limited to, non-linear optimization, combinatorial optimization, continuous optimization, stochastic optimization, Bayesian optimization, optimal control, discrete optimization, multi-objective optimization, and more. New to the series portfolio include Works at the intersection of optimization and machine learning, artificial intelligence, and quantum computing.

*Volumes from this series are indexed by Web of Science, zbMATH, Mathematical Reviews, and SCOPUS.*

More information about this series at http://www.springer.com/series/7393

Asoke Kumar Bhunia · Laxminarayan Sahoo ·
Ali Akbar Shaikh

# Advanced Optimization and Operations Research

Springer

Asoke Kumar Bhunia
Department of Mathematics
The University of Burdwan
Burdwan, West Bengal, India

Laxminarayan Sahoo
Department of Mathematics
Raniganj Girls' College
Raniganj, West Bengal, India

Ali Akbar Shaikh
Department of Mathematics
The University of Burdwan
Burdwan, West Bengal, India

ISSN 1931-6828 ISSN 1931-6836 (electronic)
Springer Optimization and Its Applications
ISBN 978-981-32-9966-5 ISBN 978-981-32-9967-2 (eBook)
https://doi.org/10.1007/978-981-32-9967-2

Mathematics Subject Classification (2010): 90Cxx, 90Bxx, 91Axx, 90B05

© Springer Nature Singapore Pte Ltd. 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

*Dedicated to*
*our respected teachers*

*Prof. M. Maiti and Prof. T. K. Pal*
*Vidyasagar University*
*West Bengal, India*

*and*

*Prof. M. Basu*
*University of Kalyani*
*West Bengal, India*

# Preface

Due to the gradual complexity of day-to-day decision-making problems, engineers, system analysts and managers face challenges in offering the best solutions for these problems. Most of these real-life decision-making problems can be modelled as non-linear constrained or unconstrained optimization problems. Optimization is the art of finding the best alternative(s) among a given set of options. The process of finding the maximum or minimum possible value which a given function can attain in its domain of definition is known as optimization. Operations research, on the other hand, a relatively new discipline, is an interdisciplinary subject. Mainly concerned with different techniques, it provides new insights and capabilities for determining better solutions of decision-making problems with higher accuracy, competence and confidence.

This book, *Advanced Optimization and Operations Research*, is intended to introduce the fundamental and advanced concepts of optimization and operations research. It has been written in simple and lucid language, according to the syllabi of the all major universities which include optimization and operations research in their graduate and post-graduate courses in areas of science, arts, commerce, engineering and management. Maximum care has been taken in preparing the chapters to include as many different topics as possible.

There are several textbooks available on the market on optimization and operations research as separate topics. Our primary motivation to write this book was the absence of one book covering these two areas together. Existing books have been written in a conventional way, whereas in this book, all 18 chapters are written in details with equal emphasis. Each chapter states its objective in the beginning and includes a sufficient number of solved examples and exercises which will be of much help for aspirants preparing for competitive examinations like NET, SET, GATE, CAT and MAT.

Chapter 1 highlights the historical perspective, various definitions, characteristic scope and purpose of operations research. Chapter 2 discusses convex and concave functions and their properties. Chapter 3 introduces the simplex method. In Chaps. 4–6, advanced linear programming techniques, like revised simplex method, dual simplex method and bounded variable technique, are discussed. Chapter 7 introduces the post-optimality analysis of linear programming problems (LPPs), and Chap. 8 deals with techniques for solving LPPs with integer values. Chapters 9 through 11 deal with the basics of unconstrained and constrained non-linear optimization problems with equality and inequality constraints. In Chap. 12, quadratic programming problems and their solution by Wolfe's modified simplex method and Beale's method are explained in details. Chapter 13 talks about game theory, and Chap. 14 discusses project management, which includes the topics of critical path analysis, PERT analysis and time-cost trade-off. Chapter 15 presents the concepts of queueing theory, including Poisson and non-Poisson queues. Chapter 16 presents ideas on flow and potential in networks, including the maximum flow problem and its solution procedure. Chapter 17 introduces deterministic, probabilistic and fuzzy inventory control models. Chapter 18 discusses the preliminaries of several topics, like matrices, determinant, vectors, probability and Markov process, which will be essential for studying the different chapters of this book.

Salient features of this book are as follows:

- It is written in a self-instructional mode so that the readers can easily understand the subject matter without any help from other sources.
- Each topic has been discussed in detail with numerous worked-out examples to help students properly understand the underlying theories.
- Advanced-level (research-level) topics have also been included.
- A set of exercises has been included at the end of each chapter.
- It is suitable to help students to prepare for different competitive examinations.

We would like to express our sincere gratitude to all the faculty members of the Department of Mathematics at the University of Burdwan, who have contributed greatly to the success of this project. We also express our heartfelt thanks to our research scholars—Avijit Duary, Subash Chandra Das, Tanmoy Banerjee, Rajan Mondal, Md. Akhtar, Md. Sadikur Rahaman and Nirmal Kumar—for their constant help in preparing the manuscript. Finally, we wish to thank Shamim Ahmad, Tooba Shafique and the other publishing teams of Springer Nature for producing the book in its present form.

West Bengal, India

Asoke Kumar Bhunia
Laxminarayan Sahoo
Ali Akbar Shaikh

# Contents

# About the Authors

**Asoke Kumar Bhunia** is Professor at the Department of Mathematics, The University of Burdwan, West Bengal, India. He obtained his Ph.D. degree in Mathematics and M.Sc. in Applied Mathematics from Vidyasagar University, India. His research interests include computational optimization, soft computing, interval mathematics and interval ranking. He has published more than 125 research papers in various national and international journals of repute. He is a reviewer of several Science Citation Index (SCI) journals. He has guided 13 Ph.D. and two M.Phil. students. An author of four research monographs and six book chapters, he is an Indian National Science Academy (INSA) Visiting Fellow and former Associate Editor of the Springer journal *OPSEARCH*.

**Laxminarayan Sahoo** is Assistant Professor at the Department of Mathematics, Raniganj Girls' College, West Bengal, India. He obtained his Ph.D. in Mathematics from the University of Burdwan and his M.Sc. in Applied Mathematics from Vidyasagar University, India. His research interests include reliability optimization, computational optimization, artificial intelligence, soft computing, interval mathematics, interval ranking and fuzzy mathematics. He has published several research papers in various national and international journals of repute. He has received a Ministry of Human Resource Development (MHRD) fellowship from the Government of India and the Professor M.N. Gopalan Award for the best Ph.D. thesis in operations research from the Operational Research Society of India.

**Ali Akbar Shaikh** is Assistant Professor at the Department of Mathematics in The University of Burdwan, West Bengal, India. Earlier, he was a postdoctoral fellow at the School of Engineering and Sciences, Tecnológico de Monterrey, Mexico. He received the award SNI of level 1 (out of 0–3) presented by the National System of Researchers of Mexico from the Government of Mexico in the year 2017. He earned his Ph.D. and M.Phil. in Mathematics from the University of Burdwan, and his

M.Sc. in Applied Mathematics from the University of Kalyani, India. Dr. Shaikh has published more than 37 research papers in different international journals of repute. His research interests include inventory control, interval optimization and particle swarm optimization.

# Chapter 1
# Introduction to Operations Research

## 1.1 Objectives

The objectives of this chapter are to discuss:

- The needs of using Operations Research (OR)
- The historical perspective of the approach of OR
- Several definitions of OR in different contexts and various phases of scientific study
- The classification and usage of various models for solving a problem under consideration
- Applications of OR in different areas.

## 1.2 Introduction

The term 'operations research (OR)' was first introduced by McClosky and Trefthen in the year 1940. This subject came into existence in a military context. The main origin of the development of OR was the Second World War. During that war, the military services of England gathered some scientists/researchers from various disciplines and formed a team to assist them in solving strategic and tactical problems relating to air and land defence of the country. As they had very limited military resources, it was necessary to decide upon their most effective utilization, e.g. efficient ocean transport, effective bombing, etc. The mission of the team was to formulate specific proposals and plans to help the military services make decisions on optimal utilization of military resources and efforts and also to implement the decisions effectively. The members of this team were not actually engaged in the military operations. However, they worked as advisors in winning the war to the extent that the scientific and systematic approaches involved in OR provided good

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_1

intellectual support to the strategic initiatives of the military services. In this context, Arthur C. Clarke gave an appropriate definition of OR, saying '*OR is an art of winning the war without actually fighting it'*.

After seeing the encouraging results by the British OR team, the US military department was quickly motivated to start similar activities in their country. The major developments of the US teams were the invention of new fighting patterns, planning, sea mining and effective utilization of electronic equipment and their successful application in war. In USA, the work of the OR team was given different names, like Operational Analysis, Operational Evaluations, Operations Research, System Analysis, System Evaluation, etc. Among these, the name Operations Research became widely accepted.

After the Second World War, industrial managers/executives who were seeking better solutions to their complex problems were attracted by the success of the military teams. The most common problem arising in industry was which method(s) to adopt so that the total cost/profit is optimum. The first mathematical technique in this field, called the simplex method of linear programming problems (LPPs), was developed in 1947 by George B. Dantzig of the US Air Force. Since then, new techniques and applications have been developed through the efforts and cooperation of interested researchers in both academic institutions and industry.

The impact of OR can now be felt in different areas of government as well as in the private sector. A number of consulting firms are currently engaged in OR activities. Besides military and business applications, OR activities include solving the problems of transportation systems, libraries, hospitals, city planning, financial decision-making, etc. Most of the Indian industries, like Indian Railways, Indian airlines, defence organizations, Hindustan Unilever Ltd., TISCO, FCI, Telecom, Union Carbide and Hindustan Steel are using OR activities to obtain their best alternatives.

## 1.3 Definitions of Operations Research

OR has been defined in various ways. It is not possible to provide uniformly acceptable definitions of OR due to its wide scope of applications. A few opinions about the definition of OR are given as follows:

1. 'OR is the art of winning the war without actually fighting it'. [Arthur C. Clarke]
2. 'OR is a scientific method of providing executive departments with a quantitative basis for decisions regarding the operations under their control'. [P. M. Morse and G. E. Kimbol (1946)]
3. 'OR is the scientific method of providing the executive with an analytical and objective basis for decisions'. [P. M. S. Blackett (1948)]
4. 'OR is the art of giving bad answers to problems to which otherwise, worse answers are given'. [T. L. Saaty]

5. 'OR is a scientific approach to problem solving for executive management'. [H. M. Wagner]
6. 'OR is an experimental and applied science devoted to observing, understanding and predicting the behaviour of purposeful man-machine systems and OR workers are actively engaged in applying this knowledge to practical problems in business, government and society'. [OR Society of America]
7. 'OR is the application of scientific methods, techniques and tools to problems involving the operations of a system so as to provide those in control of the system with optimum solutions to the problem'. [Churchman, Ackoff and Arnoff (1957)]
8. 'OR is applied decision theory. It uses many scientific, mathematical or logical means to attempt to cope with the problems that confront the executive when he tries to achieve a thorough going rationality in dealing with decision problems'. [D. W. Miller and M. K. Starr]
9. 'OR has been characterized by the use of scientific knowledge through interdisciplinary team effort for the purpose of determining the best utilization of limited resources'. [H. A. Taha]
10. 'The term "OR" has hitherto-fore been used to connate various attempts to study operations of war by scientific methods. From a more general point of view, OR can be considered to be an attempt to study those operations of modern society which involved organizations of men or of men and machines'. [P. M. Morse (1948)]
11. 'OR is a management activity pursued in two complementary ways—one half by the free and bold exercise of commonsense untrammeled by any routine, and other half by the application of a repertoire of well established precreated methods and techniques'. [Jagjit Singh (1968)]
12. 'OR is the application of the methods of science to complex problems in the direction and management of large systems of men, machines, materials and money in industry, business and defence. The distinctive approach is to develop a scientific model of the system, incorporating measurements of factors such as chance and risk with which to predict and compare the outcomes of alternative decisions, strategies or controls. The purpose is to help management to determine its policy and actions scientifically'. [Operational Research Society, UK (1971)]

## 1.4 Applications of OR

OR is mainly concerned with different techniques. It provides new insights and capabilities to determine better solutions to decision-making problems with higher speed, competence and confidence. During the last few decades, OR has been applied successfully in various areas such as defence, government sectors, service organizations and industry. Some areas of management decision-making where the tools and techniques of OR are applied are outlined as follows:

1. *Budgeting and investments*
   - (i) Cash flow analysis, long-range capital requirements, dividend policies, investment portfolios
   - (ii) Credit policies, credit risk, etc.
2. *Purchasing*
   - (i) Rules for buying, supplies with stable or varying prices
   - (ii) Determination of quantities and timing of purchases
   - (iii) Replacement policies, etc.
3. *Production management*
   - (i) Physical distribution
     - (a) Location and size of warehouses, distribution centres and retail outlets
     - (b) Distribution policy
   - (ii) Facilities planning
     - (a) Numbers and location of factories, warehouses, hospitals, etc.
     - (b) Loading and unloading facilities for railways and trucks determining the transport schedule
   - (iii) Manufacturing
     - (a) Production scheduling and sequencing
     - (b) Stabilization of production and employment training layoffs and optimum product mix
   - (iv) Maintenance and project scheduling
     - (a) Maintenance policies and preventive maintenance
     - (b) Project scheduling and allocation of resources
4. *Marketing*
   - (i) Product selection, timing, competitive actions
   - (ii) Number of sales representatives
   - (iii) Advertising media with respect to cost and time
5. *Personnel management*
   - (i) Selection of suitable personnel on minimum salary
   - (ii) Mixes of age and skills
   - (iii) Recruitment policies and assignment of jobs
6. *Research and development*
   - (i) Determination of the areas of concentration of research and development

(ii) Project selection
(iii) Reliability and alternative design
(iv) Determination of time-cost trade-off and control of development projects.

## 1.5 Main Phases of OR

In the first half of the twentieth century, it was difficult to obtain an OR method to find a procedure for conducting an OR project. The procedure for an OR study generally involves the following major phases:

**Phase I: Formulation of the Problem**
To find the solution of a problem, first it has to be formulated. Formulation of a problem for research consists of identifying, defining and specifying the measure of the computation of a decision-making problem. To formulate a problem, the following information is required:

(i) What are the objectives?
(ii) Controlled variables and their ranges
(iii) Uncontrolled variables that may affect the outcomes of the available choices.

Since it is difficult to extract a right answer from the wrong problem, this phase should be executed with considerable care.

**Phase II: Construction of a Mathematical Model**
After formulation of the problem, the next phase is to reformulate this problem into a specific form which is most useful for the analysis. The most convenient approach is to construct a mathematical model which represents the system. Three types of models are commonly used in OR and other sciences. These are (i) iconic, (ii) analogue and (iii) symbolic/mathematical.

**Phase III: Deriving a Solution from the Model**
After formulating the mathematical model for the problem under consideration, the next phase is to find a solution from this model. In OR, our objective is to search for an optimal solution which optimizes (maximizes or minimizes) the objective function of the model.

In addition to solving the model, sometimes it is also essential to perform a sensitivity analysis for studying the effect of changes of different system parameters.

**Phase IV: Testing the Model and its Solution (Updating the Model)**
After deriving a solution from the model of the problem, the entire solution should be tested again to determine whether or not there is any error. This may be done by reexamining the formulation of the problem and comparing it with that model, which may help to reveal any mistakes.

The solution of an OR problem should yield a better performance than a solution obtained by some alternative procedure. A solution derived from the model of the problem should be tested further for its optimality.

**Phase V: Implementing the Solution**
The final phase of an OR study is to implement the optimal solution derived by the OR team. Due to the constantly changing scenarios in the world, the model and its solution may not remain valid for a long period of time. With changes of scenario, the model, its solution and the resulting course of action can be modified accordingly.

## 1.6 Scope of OR

OR has considerable scope in the different sectors of our society. In general, we can say that wherever there is a problem, there is an OR technique for solving that problem. In addition to military applications, OR is widely used in many business and industry organizations among others. Here, we shall discuss the scope of OR in some particular areas of the government sector.

(i) **Defence:** During World War II, the OR teams of Britain and America developed different techniques and strategies which helped them to win different battles. In modern warfare, military operations are carried out by the air force, army and navy. Therefore, it is necessary to formulate optimum strategies which will give maximum benefits. Because OR helps military personnel to select the best possible strategy to win battles, OR has extensive applications in defence.

(ii) **Industry:** After observing successful applications of OR in the military, industry personnel became interested in this field. When a small organization expands, and in larger organizations in general, it is not possible to manage everything with a single department. As a result, the administration of an organization can establish the following departments with different objectives:

  (a) **Production department**
  The objective of this department is to produce an item with better quality by minimizing the manufacturing cost/time of the produced item.

  (b) **Marketing department**
  The objectives of this department are to:

  - Promote the item/product to the customers by attractive advertisements with the help of electronic and print media and also through sales representatives
  - Maximize the amount of profit
  - Minimize the cost of sales.

(c) **Finance department**
The objective of this department is to minimize the capital required to maintain any level of business.
These departments can come into conflict with each other as their policies may conflict. This difficulty can be overcome by the application of OR techniques.

(iii) **Life Insurance Company:** OR techniques can be applied to determine the premium rates of different policies for the better interest of the organization.
(iv) **Agriculture:** With an increasing population and consequent shortages of food, there is a need to increase the agricultural output of a country. However, many problems arise, e.g. climatic conditions and the problem of optimal distribution of water from resources, etc., for the agricultural department of a country. Thus, there is a need for best policies under the given restrictions. OR techniques may be beneficial in determining the best policies.
(v) **Planning:** Careful planning plays a crucial role in the economic development of a country, and OR techniques may be used in such planning. Thus, OR has extensive applications in the planning of different sectors of a country.

## 1.7 Modelling in OR

A model in OR may be defined as an idealized representation of a real-life system in which only the basic aspects or the most important features of a typical problem under investigation are considered.

The objective of a model is to analyse the behaviour of the system for the purpose of improving its performance.

**Advantages of a Model**
A model has many advantages over a verbal description of a problem. Some of them are as follows:

(i) It describes a problem much more concisely.
(ii) It provides a logical and systematic approach to the problem.
(iii) It enables the use of high-powered mathematical techniques to analyse the problem.
(iv) It indicates the limitations and scope of the problem.
(v) It tends to make the overall structure of the problem more comprehensible.

**Disadvantages of a Model**
Models also have a few disadvantages; these are as follows:

1. Models are only an attempt to understand an operation and should never be considered absolute in any sense.
2. The validity of any model with regard to the corresponding operation can only be verified by experimentation and the use of relevant data characteristics.

### Types of Models

Three types of models are commonly used in OR: (i) iconic models, (ii) analogue models, (iii) mathematical models.

### Iconic Models

These models represent the system as it is, but in a different size. Thus, iconic models are obtained by enlarging or reducing the size of the system. An iconic model is a pictorial representation of the system. Some common examples are photographs, drawings, maps and globes. Iconic models of the Sun and its planets are scaled down, while the model of an atom is scaled up so as to make it visible. These models have some advantages as well as disadvantages. The advantages are that they are specific, concrete and easy to construct. They can be studied more easily than the system itself.

The disadvantages are that these models are difficult to manipulate for experimental purposes. To make any modification or improvement in these models is not easy. Adjustments with changing situations cannot be performed with these models.

### Analogue Models

These models are more abstract than iconic models; thus, there is no 'look-alike' correspondence between these models and real-life items. In analogue models, one set of properties is used to represent another set of properties. For example, graphs are analogues, as distance is used to represent a wide variety of variables such as time, percentage, age, weight, etc. These models are easier to manipulate than iconic models. However, they are less specific and less concrete.

### Mathematical (Symbolic) Models

Mathematical models are more abstract and more general in nature. They employ a set of mathematical symbols to represent the components of a real system. Thus, these models are mathematical equations or inequalities reflecting the structure of the system they represent. Inventory models and queueing models are some examples of symbolic models.

These models can be classified into two categories: (i) deterministic (crisp) models and (ii) non-deterministic models. In deterministic models, all the parameters and functional relationships are assumed to be known and precise. Linear programming and classical inventory models are examples of deterministic models. On the other hand, models in which at least one parameter or decision variable is a random variable or a fuzzy number or an interval-valued number are called non-deterministic models. These models reflect the complexity of the real world and the uncertainty surrounding it.

In the context of generality, models can also be categorized as (i) specific models and (ii) general models. When a model presents a system at some specific time, it is known as a specific model. In these models, if a time factor is considered, they are termed dynamic models; otherwise, they are called static models. An inventory problem of determining the economic order quantity for the next period with a uniform demand rate is an example of a static model, whereas the same problem with a time-dependent demand is an example of a dynamic model. On the other hand, simulation and heuristic models are called general models. These models do not yield an optimum solution to the problem, but they give a solution near to the optimal one.

**General Solution Methods for OR Models**

Generally, OR models are solved by the following three methods.

(i) Analytical methods
(ii) Numerical methods
(iii) Monte Carlo technique.

**Analytical Methods**

In these methods, all the tools of basic mathematics (such as differential calculus, integral calculus, differential equations, difference equations, etc.) are used for solving an OR model.

**Numerical Methods**

These methods are concerned with iterative or trial and error methods. These methods are primarily used when analytical methods fail to derive a solution. The analytical methods may fail because of the complexity of the constraints and/or number of variables.

In this procedure, the algorithm is started with an initial approximate (trial) solution and continued with a set of given rules for improving it towards optimality. Then the trial solution is replaced by the improved one, and the process is repeated for a fixed number of times or until no further improvement is possible.

**Monte Carlo Technique**

This technique is based on the random sampling of a variable's possible values. For this technique, some random numbers are required which may be converted into random variates whose behaviour is known from past experience. Darker and Kac developed this technique, combining probability and a sampling technique to provide the solution of partial or integral differential equations. Hence, this technique is concerned with experiments on random numbers, and it provides solutions to complicated OR problems. The Monte Carlo technique is useful in the following situations:

(i) When the mathematical and statistical problems are too complicated and alternative methods are needed
(ii) When it is not possible to gain any information from past experience for a particular problem
(iii) To estimate the parameters of a model.

**Algorithm for Monte Carlo Technique**

The following steps are associated with the Monte Carlo technique:

Step 1. To illustrate the general idea of the system, draw a flow diagram.
Step 2. Make sample observations to select a suitable model for the system and determine the probability distribution for the variables of interest.
Step 3. Convert the probability distribution to a cumulative distribution.

Step 4. Select a sequence of random numbers with the help of random number tables.
Step 5. Determine the sequence of values of the variables of interest with the sequence of random numbers obtained in Step 4.
Step 6. Fit an appropriate standard mathematical function to the sequence of values obtained in Step 5.

**Advantages and Disadvantages of Monte Carlo Techniques**
The advantages of Monte Carlo techniques are as follows:

(i) They are very helpful in finding solutions of complicated mathematical problems which cannot be solved otherwise.
(ii) The difficulties of trial and error experimentation are avoided.

Monte Carlo techniques have the following disadvantages:

(i) These methods are time-consuming and costly ways of getting a solution of any problem.
(ii) These techniques do not provide an optimal solution to a problem. The solution is nearly optimal only when the size of the sample is sufficiently large.

## 1.8 Development of OR in India

In India, OR started to appear when an OR unit was established at the Regional Research Laboratory in Hyderabad in the year 1949. At the same time, Professor R. S. Verma of Delhi University set up an OR team in the Defence Science Laboratory to solve different problems involving storing, purchasing and planning. In 1953, Professor P. C. Mahalanobis established an OR team at the Indian Statistical Institute in Kolkata for solving the problem of national planning and survey. In 1957, the OR Society of India (ORSI) was formed. India now publishes a number of research journals, namely, *OPSEARCH* (published by ORSI), *Industrial Engineering and Management*, *Journal of Engineering Production*, *Defence Science Journal*, *SCIMA*, *IJOMAS*, *IAPQR*, etc.

Regarding OR education in India, the University of Delhi first introduced a complete M.Sc. course in 1963. Simultaneously, the Institute of Management at Kolkata and Ahmedabad started teaching OR in their MBA courses. Nowadays, OR has become such a popular subject that it has been introduced in almost all institutes and universities in various disciplines, like Mathematics, Statistics, Commerce, Economics, Computer Science, Engineering, Business Administration, etc. Also, because of the importance of this subject, it has been introduced in different competitive examinations, like IAS, CA, ICWA and NET.

Profesor P. C. Mahalanobis first applied OR techniques in the formulation of a second five-year plan. The Planning Commission made use of OR techniques for planning the optimum size of the Caravelle fleet of Indian airlines. Some industries, such as HLL, Union Carbide, TELCO, Hindustan Steel, TISCO, FCI, etc., have used these techniques, as well as other Indian organizations, like CSIR, TIFR, the Indian Institute of Science, the State Trading Corporation, and Indian Railways.

## 1.9 Role of Computers in OR

In the development of OR, computers play a significant role. Without computers it would not be possible to achieve OR's current position. In most OR techniques, the computations are very complicated and require a large number of iterations. Without a computer, these techniques have no practical use. Many large-scale applications of OR techniques, which require only a few seconds/minutes on the computer, would take a very long time (maybe weeks/months/years) to obtain the same results manually. In this context, the computer has become an essential and integral part of OR.

# Chapter 2
# Convex and Concave Functions

## 2.1 Objective

The objective of this chapter is to study convex & concave functions and their properties.

## 2.2 Introduction

Convex and concave functions are the key concepts of mathematical analysis and have interesting consequences in the areas of optimization theory, statistical estimation, inequalities and applied probability. These functions have some properties which can be used in developing suitable optimality conditions and computational schemes for optimization problems.

## 2.3 Convex and Concave Functions

A function $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ (where $S$ is a non-empty convex set) is said to be convex at $\bar{x} \in S$, if

$$f(\lambda\bar{x} + (1-\lambda)x) \leq \lambda f(\bar{x}) + (1-\lambda)f(x) \text{ for all } x \in S,\ 0 \leq \lambda \leq 1 \text{ and } \lambda\bar{x} + (1-\lambda)x \in S.$$

The function $f(x)$ is said to be convex on $S$ if it is convex at each $x \in S$.

The general definition of a convex function is as follows:

A function $f : S \to \mathbb{R}$ is said to be convex on $S$, if $x_1, x_2 \in S$ and $0 \leq \lambda \leq 1$ implies

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_2

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2).$$

By induction, this inequality can be extended as follows:

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i) where \sum_{i=1}^{n} \lambda_i = 1$$

for any convex combination of points from $S$.

If there is a constant $c > 0$ such that for any $x_1, x_2 \in S$,

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) - \frac{1}{2}c\lambda(1-\lambda)\|x_1 - x_2\|^2,$$

then $f$ is called a uniformly convex or strongly convex function on $S$.

A function $f$ defined on a convex set $S \subset \mathbb{R}^n$ is said to be strictly convex at $\bar{x} \in S$ if

$$f((1-\lambda)\bar{x} + \lambda x) < (1-\lambda)f(\bar{x}) + \lambda f(x)$$

for $x \in S, 0 < \lambda < 1$ and $(1-\lambda)\bar{x} + \lambda x \in S$.

$f(x)$ is said to be strictly convex on $S$ if it is strictly convex at each $x \in S$.

A function $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ (where $S$ is a non-empty convex set) is said to be a concave function at $\bar{x} \in S$ if

$$f(\lambda\bar{x} + (1-\lambda)x) \geq \lambda f(\bar{x}) + (1-\lambda)f(x) \text{ for all } x \in S,\ 0 \leq \lambda \leq 1 \text{ and } \lambda\bar{x} + (1-\lambda)x \in S.$$

The function $f(x)$ is said to be concave on $S$ if it is concave at each $x \in S$.

Then the general definition is as follows:

A function $f : S \to \mathbb{R}$ is said to be concave on $S$ if $x_1, x_2 \in S$ and $0 \leq \lambda \leq 1$ implies

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda)f(x_2).$$

By induction, the preceding inequality can be extended as follows:

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \geq \sum_{i=1}^{n} \lambda_i f(x_i) where \sum_{i=1}^{n} \lambda_i = 1$$

for any convex combination of points from $S$.

If there is a constant $c > 0$ such that for any $x_1, x_2 \in S$,

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda)f(x_2) - \frac{1}{2}c\lambda(1-\lambda)\|x_1 - x_2\|^2,$$

then $f$ is called a uniformly convex or strongly convex function on $S$.

A function $f$ defined on a convex set $S \subset \mathbb{R}^n$ is said to be strictly concave at $\bar{x} \in S$ if

$$f((1-\lambda)\bar{x} + \lambda x) > (1-\lambda)f(\bar{x}) + \lambda f(x)$$

for $x \in S, 0 < \lambda < 1$ and $(1-\lambda)\bar{x} + \lambda x \in S$.

$f(x)$ is said to be strictly concave on $S$ if it is strictly concave at each $x \in S$.

Note that a strictly convex (strictly concave) function on a convex set $S \subseteq \mathbb{R}^n$ is convex (concave) on $S$, but the converse is not always true.

For example, a constant function on $\mathbb{R}^n$ is both convex and concave on $\mathbb{R}^n$, but neither strictly convex nor strictly concave on $\mathbb{R}^n$.

## 2.4 Geometrical Interpretation

The geometrical interpretation of the convexity of a single variable function is as follows. For a convex function, the function values are below the corresponding chord (see Fig. 2.1); i.e. the values of a convex function at any point in the interval $x_1 \leq x \leq x_2$ are less than or equal to the height of the chord joining the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$. This equality holds only for linear functions (See Fig. 2.1 and Fig. 2.3a, c).

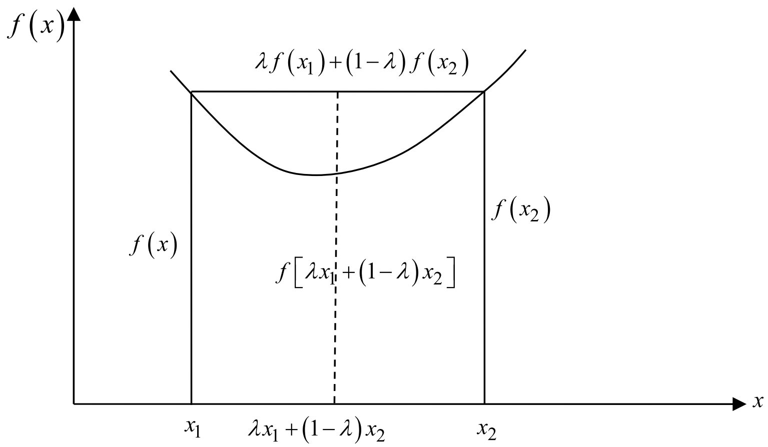



**Fig. 2.1** Convexity of single variable function $f(x)$

For a concave function the function values are above the corresponding chord; i.e. the values of a concave function at any point in the interval $x_1 \leq x \leq x_2$ are greater than or equal to the height of the chord joining the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$. (See Figs. 2.2 and 2.3b, c).

**Examples of Convex Functions**

The following are some examples of convex functions:

(i) $f(x) = e^{ax}$ is convex on $\mathbb{R}$, for any $a \in \mathbb{R}$.

(ii) $f(x) = x^a$ is convex on $\mathbb{R}_+$, when $a \geq 1$ or $a \leq 0$.

(iii) $f(x) = |x|^p$ for $p \geq 1$ is convex on $\mathbb{R}$.

(iv) The function $f$ with domain [0, 1] defined by
$f(0) = f(1) = 1$
$f(x) = 0$ for $0 < x < 1$
is convex. It is continuous on the open interval (0, 1) but not continuous at $x = 0$ and $x = 1$.

(v) Every affine function of the form $f(x) = \langle a, x \rangle + b$ in $\mathbb{R}$ is convex. Here instead of inequality equality holds, i.e.

$$f(\alpha x + (1 - \alpha)y) = \alpha f(x) + (1 - \alpha) f(y), \ \alpha \in (0, 1), \ x, \ y \in S.$$

(vi) Every Euclidean norm in $\mathbb{R}$, i.e. $f(x) = \|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$, is convex.

(vii) The distance to a convex set, i.e. the function $d(x, S)$ from points $x \in \mathbb{R}^n$ to a convex set $S$, is convex. Since for any convex combination $\alpha x + (1 - \alpha)y$, if we take sequences $u_{(k)}$ and $v_{(k)}$ from $S$ such that

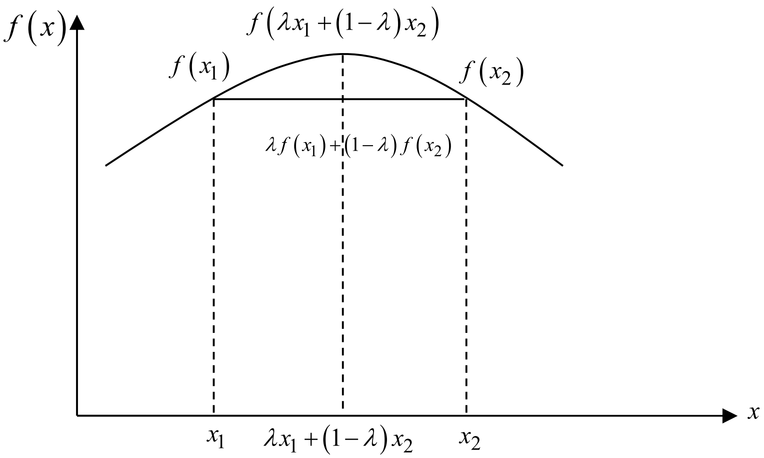



**Fig. 2.2** Concavity of single variable function $f(x)$

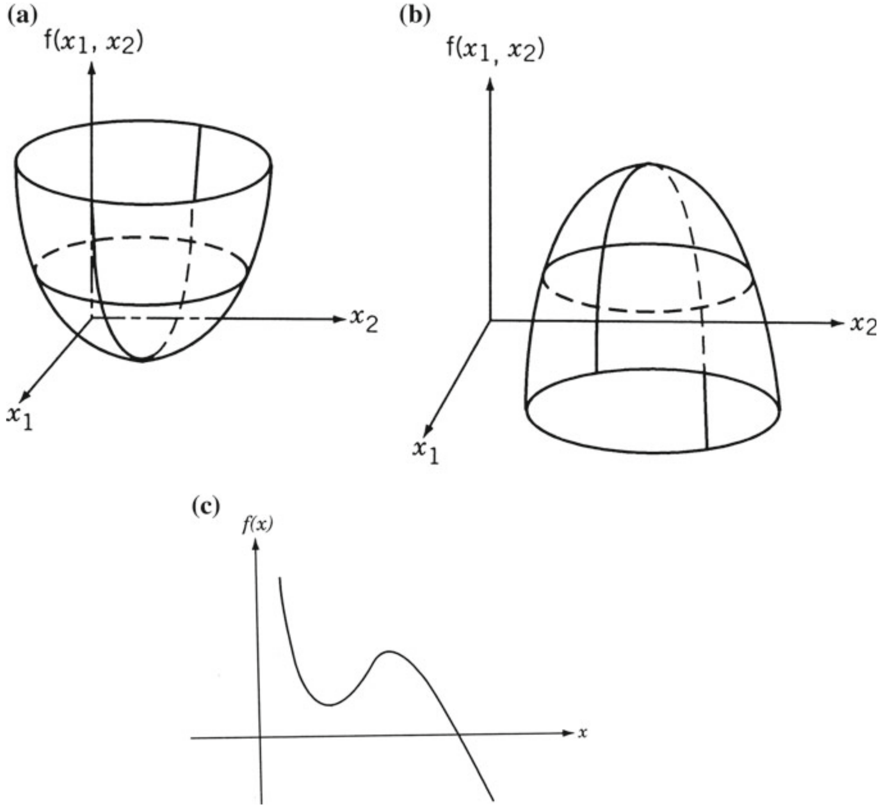



**Fig. 2.3** **a** Convex function of two variables. **b** Concave function of two variables. **c** Function (of single variable) that is convex over a certain interval and concave over another certain interval

$$d(x, S) = \lim_{k\to\infty} \left\|x - u_{(k)}\right\|$$
$$d(y, S) = \lim_{k\to\infty} \left\|y - v_{(k)}\right\|,$$

the points $\alpha u_{(k)} + (1-\alpha)v_{(k)}$ lie in $S$, and taking limits in the inequality,

$$\begin{aligned} d[\alpha x + (1-\alpha)y, S] &\leq \left\|\alpha x + (1-\alpha)y - \alpha u_{(k)} - (1-\alpha)v_{(k)}\right\| \\ &\leq \alpha\left\|x - u_{(k)}\right\| + (1-\alpha)\left\|y - v_{(k)}\right\|, \text{i.e.} \\ d[\alpha x + (1-\alpha)y, S] &\leq \alpha\, d(x, S) + (1-\alpha)d(y, S). \end{aligned}$$

(viii) The function $f(x) = \text{Max}\{x_1, x_2, \ldots, x_n\}$ is convex on $\mathbb{R}^n$.
The function $f(x) = \text{Max}_i\, x_i$ satisfies

$$\begin{aligned} f(\theta x + (1-\theta)y) &= \mathrm{Max}_i(\theta x_i + (1-\theta)y_i) \quad [0 \le \theta \le 1] \\ &\le \theta\, \mathrm{Max}_i\, x_i + (1-\theta)\, \mathrm{Max}_i\, y_i \\ &= \theta f(x) + (1-\theta) f(y) \end{aligned}$$

(ix) The function $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$ is convex on $\mathbb{R}^n$.

**Examples of Concave Functions**
Some examples of concave functions are as follows:

(i) $f(x) = \log x$ on $\mathbb{R}_+$ is a concave function.
(ii) $f(x) = x^a$ for $0 \le a \le 1$ is concave.
(iii) The geometric mean $f(x) = \left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}}$ on $\mathbb{R}^n$ is a concave function.
(iv) $f(x) = -8x^2$ is a strictly concave function.
(v) The function $f(x) = \sin x$ is concave on the interval $[0, \pi]$.

## 2.5 Properties of Convex and Concave Functions

The properties of convex and concave functions are described as follows:

(i) A function $f$ is said to be concave if $-f$ is convex.
(ii) Every linear function is both concave and convex on $\mathbb{R}^n$.

*Proof* Let $f(x) = \langle a, x\rangle + b(a \in \mathbb{R}^n)$ be a linear function on $\mathbb{R}^n$.
Let $x_1 \in \mathbb{R}^n, x_2 \in \mathbb{R}^n$ and $0 \le \lambda \le 1$.
So, $\lambda x_1 + (1-\lambda)x_2 \in \mathbb{R}^n$.

$$\begin{aligned} \text{Then} f(\lambda x_1 + (1-\lambda)x_2) &= \langle a, \lambda x_1 + (1-\lambda)x_2\rangle + b, \forall \lambda \\ &= \lambda\langle a, x_1\rangle + (1-\lambda)\langle a, x_2\rangle + b \\ &= \lambda\{\langle a, x_1\rangle + b\} + (1-\lambda)\{\langle a, x_2\rangle + b\} \\ &= \lambda f(x_1) + (1-\lambda) f(x_2). \end{aligned}$$

Hence, $f(\lambda x_1 + (1-\lambda)x_2) = \lambda f(x_1) + (1-\lambda) f(x_2), 0 \le \lambda \le 1$,
i.e. $f(x)$ is both convex and concave on $\mathbb{R}^n$.

(iii) If $f : S \subset \mathbb{R}^n \to \mathbb{R}$ is concave, then $S$ is a convex set on $\mathbb{R}^n$.
(iv) If $f$ is linear and $g$ is convex, then the composite function $g \circ f$ is convex.
(v) If $f$ is convex and $g$ is convex and non-decreasing, then $g \circ f$ is convex.

*Proof* Let $S(\subset \mathbb{R}^n)$ be a convex set and $0 \le \lambda \le 1$. Since $f(x)$ is a convex function on a convex set $S$, then

$$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda)f(x_2).$$

Since $g$ is non-decreasing, we have

$$g(f(\lambda x_1 + (1-\lambda)x_2)) \le g(\lambda f(x_1) + (1-\lambda)f(x_2)). \tag{2.1}$$

Again as $g$ is convex, then

$$g(\lambda f(x_1) + (1-\lambda)f(x_2)) \le \lambda g(f(x_1)) + (1-\lambda)g(f(x_2)). \tag{2.2}$$

From (2.1) and (2.2) we can write

$$g(f(\lambda x_1 + (1-\lambda)x_2)) \le \lambda g(f(x_1)) + (1-\lambda)g(f(x_2)),$$

which shows that $g \circ f$ is convex on $S$.

(vi) The sum of convex (concave) functions is convex (concave).
Let $f_1, f_2, \ldots, f_k : \mathbb{R}^n \to \mathbb{R}$ be convex functions. Then we have the following:

(a) $f(x) = \sum_{j=1}^{k} \alpha_j f_j(x)$, where $\alpha_j > 0$ for $j = 1, 2, \ldots, k$, is convex,
(b) $f(x) = \text{Max}\{f_1(x), f_2(x), \ldots, f_k(x)\}$ is convex.

(vii) Let us suppose that $g : \mathbb{R}^n \to \mathbb{R}$ is a concave function. Let $S = \{x : g(x) > 0\}$ and define $f : S \to \mathbb{R}$ as $f(x) = \frac{1}{g(x)}$. Then $f$ is convex over $S$.
(viii) An $n$-dimensional vector function $f$ defined on a set $S$ in $\mathbb{R}^n$ is convex (concave) at $\bar{x} \in S$, convex (concave) on $S$, if each of its components $f_i (i = 1, 2, \ldots, n)$ is convex (concave) at $\bar{x} \in S$, convex (concave) on $S$.

**Theorem** *Let $f(x)$ be a function defined on a convex set $S \subset \mathbb{R}^n$. If $f(x)$ is convex on $S$, then the set $A_\alpha = \{x : x \in S, f(x) \le \alpha\} \subset S$ is convex for each real number $\alpha$.*

*Proof*

$$\text{Let}\, x_1, x_2 \in A_\alpha \;\therefore f(x_1) \le \alpha \;\text{ and }\; f(x_2) \le \alpha. \tag{2.3}$$

Since $f(x)$ is convex on $S$,

$$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda)f(x_2) \le \lambda\alpha + (1-\lambda)\alpha = \alpha,$$

i.e. $f(\lambda x_1 + (1-\lambda)x_2) \le \alpha$.
Hence, $\lambda x_1 + (1-\lambda)x_2 \in A_\alpha$ and $A_\alpha$ is a convex set.

Note that the condition of the preceding theorem is not sufficient; i.e. if $A_\alpha$ is convex for each real number $\alpha$, it does not follow that $f(x)$ is a convex function on $S$. Let us consider the function $f(x) = x^3$ which is not convex on $\mathbb{R}$. However, the set $A_\alpha = \left\{x : x \in \mathbb{R}, x^3 \le \alpha\right\} = \left\{x : x \in \mathbb{R}, x \le \alpha^{1/3}\right\}$ is obviously convex for any $\alpha$.

**Corollary** *Let $f(x)$ be a function defined on a convex set $S \subset \mathbb{R}^n$. If $f(x)$ is concave on S, then the set $B_\alpha = \{x : x \in S, f(x) \geq \alpha\} \subset S$ is a convex set for each real number $\alpha$.*

**Theorem** *Let $f(x) = (f_1, f_2, \ldots, f_m)$ be an m-dimensional vector function defined on $S \subset \mathbb{R}^n$. If $f$ is convex (concave) at $\bar{x} \in S$, then each non-negative linear combination of its components $f_i$, $\phi(x) = \langle c, f(x)\rangle, c = (c_1, c_2, \ldots, c_m) \geq 0$ is convex (concave) at $\bar{x}$.*

*Proof* Here we shall prove the theorem for a convex function.

Let $x \in S, 0 \leq \lambda \leq 1$ and let $\lambda\bar{x} + (1-\lambda)x \in S$.

$$\begin{aligned}\text{Then } \phi[\lambda\bar{x} + (1-\lambda)x] &= <c, f[\lambda\bar{x} + (1-\lambda)x]> \;\leq\; <c, [\lambda f(\bar{x}) + (1-\lambda)f(x)]> \\ &= \lambda <c, f(\bar{x})> + (1-\lambda) <c, f(x)> \text{ [By the convexity of } f(x) \text{ at } \bar{x} \text{ and } c \geq 0] \\ &= \lambda\phi(\bar{x}) + (1-\lambda)\phi(x).\end{aligned}$$

This proves that $\phi(x)$ is convex at $\bar{x}$.

*Example* Let the function $f(x)$ be convex on a convex set $S \subset \mathbb{R}^n$. Show that for every $x_1 \in S, x_2 \in S$, the one-dimensional function $g(y) = f(yx_1 + (1-y)x_2)$ is convex for all $0 \leq y \leq 1$.

**Solution** Let $y_1, y_2 \in \mathbb{R}$. Hence, $0 \leq y_1, y_2 \leq 1$.

So, $\lambda y_1 + (1-\lambda)y_2 \in \mathbb{R}$ for $0 \leq \lambda \leq 1$.

Now, $g(\lambda y_1 + (1-\lambda)y_2) = f[\{\lambda y_1 + (1-\lambda)y_2\}x_1 + \{1 - \lambda y_1 - (1-\lambda)y_2\}x_2]$

$$\left[\begin{array}{l}\text{As } \{\lambda y_1 + (1-\lambda)y_2\}x_1 + \{1 - \lambda y_1 - (1-\lambda)y_2\}x_2 \\ = \lambda\{y_1x_1 + (1-y_1)x_2\} + (1-\lambda)\{y_2x_1 + (1-y_2)x_2\}\end{array}\right]$$

$$\begin{aligned}&= f[\lambda\{y_1x_1 + (1-y_1)x_2\} + (1-\lambda)\{y_2x_1 + (1-y_2)x_2\}] \\ &\leq \lambda f[y_1x_1 + (1-y_1)x_2] + (1-\lambda)f[y_2x_1 + (1-y_2)x_2] \\ &= \lambda g(y_1) + (1-\lambda)g(y_2),\end{aligned}$$

which implies that $g(y)$ is convex for all $0 \leq y \leq 1$.

## 2.6 Differentiable Convex and Concave Functions

Here we shall discuss some of the properties of differentiable convex and concave functions.

Let $f$ be a function defined on an open set $S$ in $R^n$. If $f$ is differentiable at $\bar{x} \in S$, then for $x \in R^n$ and $\bar{x} + x \in S$, $f(\bar{x} + x) = f(\bar{x}) + \langle \nabla f(\bar{x}), x\rangle + \alpha(\bar{x}, x)\|x\|$ and $\lim_{x \to 0} \alpha(\bar{x}, x) = 0$, where $\nabla f(\bar{x})$ is the $n$-dimensional gradient vector of $f$ at $\bar{x}$ whose $n$ components are the partial derivatives of $f$ with respect to $x_1, x_2, \ldots, x_n$ evaluated at $\bar{x}$, and $\alpha$ is a function of $x$.

**Theorem** *Let f be a function defined on an open set $S \subset R^n$ and let f be differentiable at $\bar{x} \in S$. If f is convex at $\bar{x} \in S$, then*

$$f(x) - f(\bar{x}) \geq \langle \nabla f(\bar{x}), (x - \bar{x}) \rangle \text{ for each } x \in S.$$

*If f is concave at $\bar{x} \in S$, then $f(x) - f(\bar{x}) \leq \langle \nabla f(\bar{x}), (x - \bar{x}) \rangle$ for each $x \in S$.*

*Proof* Let $f$ be convex at $\bar{x} \in S$.

Since $S$ is open, then there exists an open ball $B_\delta(\bar{x})$ around $\bar{x}$, which is contained in $S$.
Let $\bar{x} \in S$ and $x \neq \bar{x}$.
Then for some $\mu$ such that $0 < \mu < 1$ and $\mu < \delta / \|x - \bar{x}\|$,
we have

$$\begin{aligned} \hat{x} &= \bar{x} + \mu(x - \bar{x}) \\ &= (1 - \mu)\bar{x} + \mu x \in B_\delta(\bar{x}) \subset S. \end{aligned}$$

Since $f$ is convex at $\bar{x}$, from the convexity of $B_\delta(\bar{x})$ and the fact that $\hat{x} \in B_\delta(\bar{x})$ and $0 < \lambda < 1$, it follows that:

$$\begin{aligned} &(1 - \lambda) f(\bar{x}) + \lambda f(\hat{x}) \geq f[(1 - \lambda)\bar{x} + \lambda \hat{x}] \\ &\text{or,} \quad f(\bar{x}) - \lambda f(\bar{x}) + \lambda f(\hat{x}) \geq f[(1 - \lambda)\bar{x} + \lambda \hat{x}] \\ &\text{or,} \quad \lambda[f(\hat{x}) - f(\bar{x})] \geq f[(1 - \lambda)\bar{x} + \lambda \hat{x}] - f(\bar{x}) \\ &\text{or,} \quad f(\hat{x}) - f(\bar{x}) \geq \tfrac{f[\bar{x} + \lambda(\hat{x} - \bar{x})] - f(\bar{x})}{\lambda} \\ &\qquad = \tfrac{\lambda \langle \nabla f(\bar{x}), (\hat{x} - \bar{x}) \rangle + \alpha[\bar{x}, \lambda(\hat{x} - \bar{x})] \lambda \|\hat{x} - \bar{x}\|}{\lambda} \\ &\qquad\qquad [\text{As } f \text{ is differentiable at } \bar{x} \in S] \\ &\qquad = \langle \nabla f(\bar{x}), (\hat{x} - \bar{x}) \rangle + \alpha[\bar{x}, \lambda(\hat{x} - \bar{x})] \|\hat{x} - \bar{x}\| \\ &\therefore f(\hat{x}) - f(\bar{x}) \geq \langle \nabla f(\bar{x}), (\hat{x} - \bar{x}) \rangle + \alpha[\bar{x}, \lambda(\hat{x} - \bar{x})] \|\hat{x} - \bar{x}\|. \end{aligned}$$

Now taking the limit of both sides as $\lambda \to 0$, we have

$$f(\hat{x}) - f(\bar{x}) \geq \langle \nabla f(\bar{x}), (\hat{x} - \bar{x}) \rangle, \text{since} \lim_{\lambda \to 0} \alpha[\bar{x}, \lambda(\hat{x} - \bar{x})] = 0. \tag{2.4}$$

$$\text{Since } f \text{ is convex at } \bar{x}, \hat{x} \in S \text{ and } \hat{x} = (1 - \mu)\bar{x} + \mu x, \tag{2.5}$$

we have $(1 - \mu) f(\bar{x}) + \mu f(x) \geq f[(1 - \mu)\bar{x} + \mu x]$

$$\text{or } \mu[f(x) - f(\bar{x})] \geq f(\hat{x}) - f(\bar{x}) \tag{2.6}$$

$$\text{or } \mu[f(x) - f(\bar{x})] \geq \langle \nabla f(\bar{x}), (\hat{x} - \bar{x}) \rangle \quad \begin{bmatrix} \text{Since from (2.4)} \\ f(\hat{x}) - f(\bar{x}) \geq \langle \nabla f(\bar{x}), (\hat{x} - \bar{x}) \rangle \end{bmatrix}$$

$$\text{or} \quad \mu[f(x) - f(\bar{x})] \geq \langle \nabla f(\bar{x}), \mu(x - \bar{x}) \rangle \quad \begin{bmatrix} \text{From (2.5), since } \hat{x} = (1 - \mu)\bar{x} + \mu x \\ \therefore \hat{x} = \bar{x} - \mu \bar{x} + \mu x \text{ or,} \quad \hat{x} - \bar{x} = \mu(x - \bar{x}) \end{bmatrix}$$

$$\text{or } f(x) - f(\bar{x}) \geq \langle \nabla f(\bar{x}), (x - \bar{x}) \rangle \quad [\because \mu > 0].$$

The second part of the theorem can be proved similarly.

**Theorem** *Let f be a differential function on an open convex set $S \subset R^n$. f is convex on S if and only if*

$$f(x_2) - f(x_1) \geq \langle \nabla f(x_1), (x_2 - x_1) \rangle \text{ for each } x_1, x_2 \in S.$$

*f is concave on S if and only if*

$$f(x_2) - f(x_1) \leq \langle \nabla f(x_1), (x_2 - x_1) \rangle \text{ for each } x_1, x_2 \in S.$$

*Proof* Necessary part:

This part of the proof follows from the previous theorem.

Sufficient part:

$$\text{Let} f(x_2) - f(x_1) \geq \langle \nabla f(x_1), (x_2 - x_1) \rangle \text{ for each } x_1, x_2 \in S. \tag{2.7}$$

Since $S$ is convex and $x_1, x_2 \in S$, $\therefore (1-\lambda)x_1 + \lambda x_2 \in S$ for $0 \leq \lambda \leq 1$.

Now from (2.4), we have

$$\begin{aligned} & f(x_1) - f[(1-\lambda)x_1 + \lambda x_2] \geq \langle \nabla f[(1-\lambda)x_1 + \lambda x_2], [x_1 - \{(1-\lambda)x_1 + \lambda x_2\}] \rangle \\ & \qquad \text{for } x_1 \in S \text{ and } (1-\lambda)x_1 + \lambda x_2 \in S \\ & \text{or } f(x_1) - f[(1-\lambda)x_1 + \lambda x_2] \geq \langle \nabla f[(1-\lambda)x_1 + \lambda x_2], \{x_1 - x_1 + \lambda x_1 - \lambda x_2\} \rangle \\ & \text{or } f(x_1) - f[(1-\lambda)x_1 + \lambda x_2] \geq \lambda \langle \nabla f[(1-\lambda)x_1 + \lambda x_2], (x_1 - x_2) \rangle. \end{aligned} \tag{2.8}$$

Again for $x_2 \in S, (1-\lambda)x_1 + \lambda x_2 \in S$, we have from (2.7)

$$\begin{aligned} & f(x_2) - f[(1-\lambda)x_1 + \lambda x_2] \geq \langle \nabla f[(1-\lambda)x_1 + \lambda x_2], [x_2 - \{(1-\lambda)x_1 + \lambda x_2\}] \rangle \\ & \text{or } f(x_2) - f[(1-\lambda)x_1 + \lambda x_2] \geq \langle \nabla f[(1-\lambda)x_1 + \lambda x_2], [-(1-\lambda)x_1 + (1-\lambda)x_2] \rangle \\ & \text{or } f(x_2) - f[(1-\lambda)x_1 + \lambda x_2] \geq -(1-\lambda) \langle \nabla f[(1-\lambda)x_1 + \lambda x_2], (x_1 - x_2) \rangle. \end{aligned} \tag{2.9}$$

Now multiply (2.8) by $(1-\lambda)$ and (2.9) by $\lambda$; then adding, we have

$$\begin{aligned} & (1-\lambda)f(x_1) - (1-\lambda)f[(1-\lambda)x_1 + \lambda x_2] + \lambda f(x_2) - \lambda f[(1-\lambda)x_1 + \lambda x_2] \\ & \quad \geq \lambda(1-\lambda)\langle \nabla f[(1-\lambda)x_1 + \lambda x_2], (x_1 - x_2) \rangle - \lambda(1-\lambda)\langle \nabla f[(1-\lambda)x_1 + \lambda x_2], (x_1 - x_2) \rangle \\ & \text{or } (1-\lambda)f(x_1) + \lambda f(x_2) - (1-\lambda+\lambda)f[(1-\lambda)x_1 + \lambda x_2] \geq 0 \\ & \text{or } (1-\lambda)f(x_1) + \lambda f(x_2) \geq f[(1-\lambda)x_1 + \lambda x_2]. \end{aligned}$$

Hence, $f$ is convex on $S$.

Similarly, we can prove that $f$ is concave on $S$.

**Theorem** *Let f be a differentiable function on an open convex set $S \subset R^n$. A necessary and sufficient condition that f be convex (concave) on S is that for each $x_1, x_2 \in S$*

$$\langle [\nabla f(x_2) - \nabla f(x_1)], (x_2 - x_1)\rangle \geq 0 \quad (\leq 0).$$

*Proof* We shall establish the theorem for the convex case only. The concave case is similar.

The condition is necessary.

Let $f$ be convex on $S$ and let $x_1, x_2 \in S$.

Then we have $f(x_2) - f(x_1) \geq \langle \nabla f(x_1), (x_2 - x_1)\rangle$,

$$\text{i.e. } f(x_2) - f(x_1) - \langle \nabla f(x_1), (x_2 - x_1)\rangle \geq 0 \tag{2.10}$$

$$\text{and } f(x_1) - f(x_2) \geq \langle \nabla f(x_2), (x_1 - x_2)\rangle,$$

$$\text{i.e. } f(x_1) - f(x_2) \geq -\langle \nabla f(x_2), (x_2 - x_1)\rangle$$

$$\text{or } f(x_1) - f(x_2) + \langle \nabla f(x_2), (x_2 - x_1)\rangle \geq 0. \tag{2.11}$$

Now adding these two inequalities (2.10) and (2.11), we have

$$f(x_2) - f(x_1) - \langle \nabla f(x_1), (x_2 - x_1)\rangle + f(x_1) - f(x_2) + \langle \nabla f(x_2), (x_2 - x_1)\rangle \geq 0$$
$$\text{or } \langle \nabla f(x_2), (x_2 - x_1)\rangle - \langle \nabla f(x_1), (x_2 - x_1)\rangle \geq 0$$
$$\text{or } \langle [\nabla f(x_2) - \nabla f(x_1)], (x_2 - x_1)\rangle \geq 0.$$

The condition is sufficient.

$$\text{Let } \langle [\nabla f(x_2) - \nabla f(x_1)], (x_2 - x_1)\rangle \geq 0 \text{ for each } x_1, x_2 \in S. \tag{2.12}$$

We shall prove that $f$ is convex on $S$.

Since $S$ is a convex set and $x_1, x_2 \in S$, then for $0 \leq \lambda \leq 1$, $(1-\lambda)x_1 + \lambda x_2 \in S$.

Now by the mean value theorem, we have for some $\bar{\lambda}, 0 < \bar{\lambda} < 1$,

$$f(x_2) - f(x_1) = \left\langle \nabla f\left[x_1 + \bar{\lambda}(x_2 - x_1)\right], (x_2 - x_1)\right\rangle. \tag{2.13}$$

But by inequality (2.12),

$$\left\langle \left\{\nabla f\left[x_1 + \bar{\lambda}(x_2 - x_1)\right] - \nabla f(x_1)\right\}, \bar{\lambda}(x_2 - x_1)\right\rangle \geq 0$$
$$\text{or } \left\langle \left\{\nabla f\left[x_1 + \bar{\lambda}(x_2 - x_1)\right] - \nabla f(x_1)\right\}, (x_2 - x_1)\right\rangle \geq 0$$
$$\text{or } \left\langle \nabla f\left[x_1 + \bar{\lambda}(x_2 - x_1)\right], (x_2 - x_1)\right\rangle \geq \langle \nabla f(x_1), (x_2 - x_1)\rangle$$
$$\text{or } f(x_2) - f(x_1) \geq \langle \nabla f(x_1), (x_2 - x_1)\rangle \qquad [\text{using } (2.4)].$$

Hence, $f$ is convex on $S$ [By the theorem 'if $f$ is a differentiable function on an open convex set $S \subset R^n$, $f$ is convex on $S$ and if $f(x_2) - f(x_1) \geq \langle \nabla f(x_1), (x_2 - x_1) \rangle$ then $f$ is convex on $S$.']

**Note** If $f$ is an $n$-dimensional function on $S \subset R^n$ and $\langle [f(x_2) - f(x_1)], (x_2 - x_1) \rangle \geq 0$ for all $x_1, x_2 \in S$, then $f$ is said to be monotonic on $S$.

So from the previous theorem, a differentiable function on the open convex set $S \subset R^n$ is convex if and only if $\nabla f$ is monotonic on $S$.

## 2.7 Differentiable Strictly Convex and Concave Functions

**Theorem** *Let $f$ be a function defined on an open set $S \subset R^n$ and let $f$ be differentiable at $\bar{x} \in S$. If $f$ is strictly convex at $\bar{x} \in S$, then*

$$f(x) - f(\bar{x}) > \langle \nabla f(\bar{x}), (x - \bar{x}) \rangle \, \textit{for each} \, x \in S, x \neq \bar{x}.$$

*If $f$ is strictly concave at $\bar{x} \in S$, then*

$$f(x) - f(\bar{x}) < \langle \nabla f(\bar{x}), (x - \bar{x}) \rangle \, \textit{for each} \, x \in S, \; x \neq \bar{x}.$$

*Proof* Let $f$ be strictly convex at $\bar{x}$.

$$\text{Then} \, f[(1-\lambda)\bar{x} + \lambda x] < (1-\lambda) f(\bar{x}) + \lambda f(x) \text{ for all} \, x \in S, x \neq \bar{x} \tag{2.14}$$

$$0 < \lambda < 1, \text{ and } (1-\lambda)\bar{x} + \lambda x \in S.$$

Since $f$ is convex and differentiable at $\bar{x} \in S$,

$$\text{then} \, f(x) - f(\bar{x}) \geq \langle \nabla f(\bar{x}), (x - \bar{x}) \rangle \text{ for all} \, \bar{x} \in S. \tag{2.15}$$

We shall now show that if equality holds in (2.15) for some $x$ in $S$ which is distinct from $\bar{x}$, then a contradiction arises. Let the equality hold in (2.15) for $x = \tilde{x}, \tilde{x} \in S$ and $\tilde{x} \neq \bar{x}$. Then

$$\begin{aligned} f(\tilde{x}) - f(\bar{x}) &= \langle \nabla f(\bar{x}), (\tilde{x} - \bar{x}) \rangle \\ \text{or } f(\tilde{x}) &= f(\bar{x}) + \langle \nabla f(\bar{x}), (\tilde{x} - \bar{x}) \rangle. \end{aligned} \tag{2.16}$$

Since $\tilde{x} \in S$, from (2.14) we have

$$\begin{aligned} & f[(1-\lambda)\bar{x} + \lambda\tilde{x}] < (1-\lambda) f(\bar{x}) + \lambda f(\tilde{x}) \quad [\text{putting } x = \tilde{x} \text{ in } (2.14)] \\ & \text{or } f[(1-\lambda)\bar{x} + \lambda\tilde{x}] < (1-\lambda) f(\bar{x}) + \lambda[f(\bar{x}) + \langle \nabla f(\bar{x}), (\tilde{x} - \bar{x}) \rangle] \; [\text{using } (2.16)] \\ & \qquad \text{for } 0 < \lambda < 1 \text{ and } (1-\lambda)\bar{x} + \lambda\tilde{x} \in S \\ & \text{or } f[(1-\lambda)\bar{x} + \lambda\tilde{x}] < f(\bar{x}) + \lambda \langle \nabla f(\bar{x}), (\tilde{x} - \bar{x}) \rangle. \end{aligned} \tag{2.17}$$

Again for the points $x = (1 - \lambda)\bar{x} + \lambda\tilde{x} \in S$ and $\bar{x}$, from (2.15) we have

$$\begin{aligned} & f[(1 - \lambda)\bar{x} + \lambda\tilde{x}] - f(\bar{x}) \geq \langle \nabla f(\bar{x}), \lambda(\tilde{x} - \bar{x}) \rangle \\ & \text{or}\, f[(1 - \lambda)\bar{x} + \lambda\tilde{x}] \geq \lambda \nabla f(\bar{x})(\tilde{x} - \bar{x}) + f(\bar{x}). \end{aligned} \tag{2.18}$$

Relation (2.18) contradicts relation (2.13).

For small $\lambda, (1 - \lambda)\bar{x} + \lambda\tilde{x} \in S$ is open.

Hence, equality cannot hold in (2.15) for any $x \in S$ distinct.

**Theorem** *Let $f(x)$ be a twice differentiable function on an open convex set $S \subset \mathbb{R}^n$ and $H(x)$ be the Hessian matrix of $f(x)$. Then $f(x)$ is convex on $S$ if and only if $H(x)$ is positive semi-definite for each $x \in S$.*

*Proof* Let $x \in S$ and $y \in \mathbb{R}^n$. Since $S$ is convex and open, there exists a $\overline{\lambda} > 0$ such that $x + \lambda y \in S$ for $0 < \lambda < \overline{\lambda}$.

Since $f(x)$ is a twice differentiable function on $S$, then

$$f(x + \lambda y) = f(x) + \lambda \langle \nabla f(x), y \rangle + \frac{\lambda^2}{2} \langle H(x)y, y \rangle + \lambda^2 \beta(x, \lambda y) \|y\|^2 \forall \lambda \, (0 < \lambda < \overline{\lambda}).$$

When $\beta \to 0$ as $\lambda \to 0$,

$$f(x + \lambda y) - f(x) - \lambda \langle \nabla f(x), y \rangle = \frac{\lambda^2}{2} \left[ \langle H(x)y, y \rangle + 2\beta \|y\|^2 \right].$$

For a differentiable convex function, we know that

$$\begin{aligned} & f(x + \lambda y) - f(x) - \lambda \langle \nabla f(x), y \rangle \geq 0 \\ \text{or} \quad & \tfrac{\lambda^2}{2} \left[ \langle H(x)y, y \rangle + 2\beta \|y\|^2 \right] \geq 0 \\ \text{or} \quad & \langle H(x)y, y \rangle + 2\beta \|y\|^2 \geq 0 \text{ for } 0 < \lambda < \overline{\lambda}. \end{aligned}$$

Letting $\lambda \to 0$, we have $\langle H(x)y, y \rangle \geq 0$ for all $y \in \mathbb{R}^n$, which means that $H(x)$ is positive semi-definite for all $x \in S$. Let the Hessian matrix $H(x)$ be positive semi-definite for each $x \in S$.

Hence, $\langle H(x)y, y \rangle \geq 0$ for all $y \in \mathbb{R}^n$.

Now by Taylor's theorem, we have for $x_1, x_2 \in S$

$$f(x_2) - f(x_1) - \langle \nabla f(x_1), (x_2 - x_1) \rangle = \frac{1}{2} \langle H(x_1 + \theta(x_2 - x_1))(x_2 - x_1), (x_2 - x_1) \rangle$$

for some $0 < \theta < 1$.

Since $x_1 + \theta(x_2 - x_1) \in S$, the right-hand side of the preceding equation is non-negative as $\langle H(x)y, y \rangle \geq 0$ for all $y \in \mathbb{R}^n$.

Hence, $f(x_2) - f(x_1) - \langle \nabla f(x_1), (x_2 - x_1) \rangle \geq 0$ for $x_1, x_2 \in S$,

i.e. $f(x)$ is convex on $S$.

**Theorem** *Let $f(x)$ be a twice differentiable function on an open convex set $S(\subset \mathbb{R}^n)$ and $H(x)$ be the Hessian matrix of $f(x)$. Then $f(x)$ is strictly convex on $S$ if $H(x)$ is positive definite for each $x \in S$.*

*Remark*

(i) The converse of the previous theorem is not true. To prove this, let us consider the function $y = x^4, x \in \mathbb{R}$. Hence, $H(x) = 12x^2$, which is not positive definite since $H(0) = 0$, but $f(x)$ is strictly convex on $\mathbb{R}$ because $f'(x) = 4x^3$ is strictly increasing on $\mathbb{R}$.

(ii) Let $f(x)$ be a twice differentiable function of a single variable $x$ on an open set $S(\subset \mathbb{R}^n)$. Then $f(x)$ is strictly convex if $f''(x)$ is positive for each $x \in S$.

*Example* Determine whether the following functions are convex or concave:

(a) $f(x) = e^x$
(b) $f(x) = -8x^2$
(c) $f(x_1, x_2) = 2x_1^3 - 6x_2^2$
(d) $f(x_1, x_2, x_3) = 4x_1^2 + 3x_2^2 + 5x_3^2 + 6x_1x_2 + x_1x_3 - 3x_1 - 2x_2 + 15$

**Solution**

(a) $f(x) = e^x$
$\therefore f''(x) = e^x > 0$ for all real values of $x$.
Hence, $f(x)$ is strictly convex.

(b) $f(x) = -8x^2$, $f''(x) = -16 < 0$
for all real values of $x$. Hence, $f(x)$ is strictly concave.

(c) $f(x_1, x_2) = 2x_1^3 - 6x_2^2$

$$H(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 12x_1 & 0 \\ 0 & -12 \end{bmatrix}.$$

$$\text{Since } \frac{\partial^2 f}{\partial x_1^2} = 12x_1 \le 0 \text{ for } x_1 \le 0 \text{ and}$$
$$\ge 0 \text{ for } x_1 \ge 0$$

$$\text{and } |H(x)| = -144x_1 \ge 0 \text{ for } x_1 \le 0$$
$$\le 0 \text{ for } x_1 \ge 0,$$

then $f(x)$ will be negative semi-definite for $x_1 \le 0$, i.e. $f(x)$ is convex for $x_1 \le 0$.

(d) $f(x_1, x_2, x_3) = 4x_1^2 + 3x_2^2 + 5x_3^2 + 6x_1x_2 + x_1x_3 - 3x_1 - 2x_2 + 15$

$$H(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_1 \partial x_3} & \frac{\partial^2 f}{\partial x_2 \partial x_3} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix} = \begin{bmatrix} 8 & 6 & 1 \\ 6 & 6 & 0 \\ 1 & 0 & 10 \end{bmatrix}.$$

Here the leading principal minors are given by $|8| = 8 > 0$:

$$\begin{vmatrix} 8 & 6 \\ 6 & 6 \end{vmatrix} = 12 > 0 \text{ and} \begin{vmatrix} 8 & 6 & 1 \\ 6 & 6 & 0 \\ 1 & 0 & 10 \end{vmatrix} = 114 > 0.$$

Hence, the matrix $H(x)$ is positive definite for all real values of $x_1, x_2$ and $x_3$. Therefore, $f(x)$ is a strictly convex function.

# Chapter 3
# Simplex Method

## 3.1 Objective

The objective of this chapter is to discuss:

- The concept of a linear programming problem (LPP)
- The concepts of slack/surplus variables and canonical and standard forms of LPPs
- The theories and algorithm of the simplex method
- The Big M and two-phase methods
- Finding the solution of linear simultaneous equations and inverse of a non-singular matrix by the simplex method.

## 3.2 General Linear Programming Problem

The general linear programming problem (LPP) is a problem of determining the value of $n$ decision variables $x_1, x_2, \ldots, x_n$ which optimizes (maximizes or minimizes) the linear function

$$z = c_1x_1 + c_2x_2 + \cdots + c_nx_n \tag{3.1}$$

subject to the inequalities

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq, =, \geq b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \leq, =, \geq b_2 \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leq, =, \geq b_m \end{array}\right\} \tag{3.2}$$

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_3

and

$$x_j \geq 0, \quad j = 1, 2, \ldots, n. \tag{3.3}$$

The linear function $z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$ which is to be maximized (or minimized) is called the objective function of the general LPP.

The inequalities (3.2) are called the constraints of the general LPP, whereas the set of inequalities (3.3) is usually known as the non-negativity conditions of the general LPP.

The right-hand quantities $b_i(i = 1,\ 2,\ \ldots,\ m)$ of the constraints (3.2) are known as the requirement parameters of the problem.

In a compact form, the general LPP can be written as follows:

$$\begin{aligned}
&\text{Optimize } z = \sum_{j=1}^{n} c_j x_j \\
&\text{subject to the constraints} \\
&\qquad \sum_{j=1}^{n} a_{ij}x_j \leq,\ =,\ \geq b_i,\ \ i = 1,\ 2,\ \ldots,\ m \\
&\quad \text{and } x_j \geq 0,\ \ j = 1, 2, \ldots, n.
\end{aligned}$$

In matrix notation, the general LPP can be written as follows:

$$\begin{aligned}
&\text{Optimize } z = \langle c^T,\ x\rangle \\
&\text{subject to the constraints} \\
&\qquad Ax \leq,\ =,\ \geq b \\
&\quad \text{and } x \geq 0,
\end{aligned}$$

where $c, x^T \in R^n, b^T \in R^m$, and $A$ is an $m \times n$ real matrix.

**Note** The notation $\langle c^T, x\rangle$ represents the inner product of two vectors. In some books, it is denoted by either $c^Tx$ or $cx$. In a mathematical sense, $\langle c^T, x\rangle$ is the appropriate representation as $c^Tx$ or $cx$ is a matrix.

## 3.3 Slack and Surplus Variables

In the formulation of a general LPP, if the constraint is in the form of a $\leq$ type of inequality, such as

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \leq b_i,$$

then to convert this inequality into an equation, a non-negative quantity $x_{n+i}$ is added on the left-hand side of the constraint. The converted equation is as follows:

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + x_{n+i} = b_i.$$

Hence, the non-negative quantity $x_{n+i}$ is known as a slack variable.

Similarly, if a constraint is in the form of a $\geq$ type of inequality, such as

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \geq b_i,$$

then to transform this inequality into an equation, a non-negative quantity $x_{n+i}$ is subtracted from the left-hand side of the constraint. The converted equation is as follows:

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n - x_{n+i} = b_i.$$

Hence, the non-negative quantity $x_{n+i}$ is known as a surplus variable. Note that the coefficients of slack and/or surplus variables in the objective function are always assumed to be zero. This means that the conversion of the constraints to a system of simultaneous linear equations does not change the objective function.

## 3.4 Canonical and Standard Forms of LPP

After the formulation of the LPP, the next step is to determine its solution. But, to solve the LPP, the problem must be expressed in the following two forms:

(i) Canonical form
(ii) Standard form.

**Canonical form**

The formulation of a general LPP can always be expressed in the following form:

$$\begin{aligned} &\text{Maximize } z = \sum_{j=1}^{n} c_j x_j \\ &\text{subject to the constraints} \\ &\qquad \sum_{j=1}^{n} a_{ij}x_j \leq b_i, \quad i = 1,\, 2,\, \ldots,\, m \\ &\quad \text{and } x_j \geq 0, \quad j = 1, 2, \ldots, n \end{aligned}$$

This form of LPP is called the canonical form of the LPP. The characteristics of this form are as follows:

1. The objective function is of the maximization type. If the objective function is of the minimization type, it can be expressed as a maximization type by using the following mini-max property:

$$\text{Minimize } f(x) = -\text{Maximize } \{-f(x)\}$$

   For example, the objective function

$$\text{Minimize } z = 3x_1 + x_2 - 2x_3$$

   is equivalent to

$$\text{Maximize } z' = -z = -3x_1 - x_2 + 2x_3.$$

2. All the constraints are of $\leq$ type, except the non-negativity restrictions. An inequality of $\geq$ type can changed to an inequality of $\leq$ type by multiplying both sides of the inequality by $(-1)$.

   For example, the constraint

$$3x_1 + 2x_2 - x_3 \geq 6$$

   is equivalent to

$$-3x_1 - 2x_2 + x_3 \leq -6.$$

   On the other hand, an equation can be replaced by two weak inequalities.

   For example, the constraint

$$5x_1 - x_2 + 7x_3 = 15$$

   is equivalent to

$$5x_1 - x_2 + 7x_3 \leq 15$$

   and $5x_1 - x_2 + 7x_3 \geq 15$,

   i.e. $5x_1 - x_2 + 7x_3 \leq 15$

$$-5x_1 + x_2 - 7x_3 \leq -15.$$

3. All the variables are non-negative. If a variable is unrestricted in sign (i.e. the value of that variable may be positive, negative or zero), it can be replaced by the difference between two non-negative variables. For example, if a variable $x_1$ is unrestricted in sign, we can replace this variable by $x_1' - x_1''$ where both $x_1'$ and $x_1''$ are non-negative, i.e. $x_1 = x_1' - x_1''$ where $x_1' \geq 0$ and $x_1'' \geq 0$.

**Standard form**

A general LPP in the form

$$\begin{aligned}
&\text{Maximize or Minimize } z = \sum_{j=1}^{n} c_j x_j \\
&\text{subject to the constraints} \\
&\qquad \sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = 1, 2, \ldots, m \\
&\quad \text{and } x_j \geq 0, \ \ j = 1, 2, \ldots, n
\end{aligned}$$

is known as the standard form of the LPP.

The characteristics of this form are as follows:

1. All the constraints are expressed in the form of equations, except the non-negativity restrictions.
2. The right-hand side of each constraint equation is non-negative.

The inequality constraint can be converted to an equation by introducing slack/surplus variables on the left-hand side of that constraint.

## 3.5 Fundamental Theorem of LPP

**Theorem** *If the LPP admits an optimal solution, then the optimal solution will coincide with at least one basic feasible solution of the problem.*

*Proof* Here we shall prove the theorem for the maximization problem.

Let us assume that $x^*$ is an optimal solution of the following LPP:

$$\begin{aligned}
&\text{Maximize } z = \langle c^T, x \rangle \\
&\text{subject to } Ax = b \\
&\text{and } x \geq 0.
\end{aligned} \tag{3.4}$$

Here $A$ is an $m \times n$ $(m < n)$ matrix and is given by $A = [a_1, a_2, \ldots, a_n]$,

where $a_j$ is an $m$-component column vector given by $a_j = [a_{1j}, a_{2j}, \ldots, a_{mj}]^T$, $j = 1, 2, \ldots, n$.

Without any loss of generality, let us assume that the first $k$ components of $x^*$ are non-zero and the remaining $(n - k)$ components of $x^*$ are zero.

Thus, $x^* = [x_1, x_2, \ldots, x_k, 0, 0, \ldots, 0]$.

Then, from (3.4), $Ax^* = b$ gives

$$\sum_{j=1}^{k} a_{ij}x_j = b_i,\ i = 1, 2, \ldots, m$$

in which $A = [a_1, a_2, \ldots, a_k, a_{k+1}, \ldots, a_n]$

so that

$$a_1x_1 + a_2x_2 + \cdots, a_kx_k = b \tag{3.5}$$

Also $z^* = z_{\text{Max}} = \sum_{j=1}^{k} c_jx_j$.

Now, if the vectors $a_1, a_2, \ldots, a_k$ corresponding to the non-zero components of $x^*$ are linearly independent, then by definition $x^*$ is a basic feasible solution; hence, the theorem holds in this case. If $k = m$, then the basic feasible solution will be non-degenerate. Again, if $k < m$, then it will form a degenerate basic feasible solution with $m - k$ of the basic variables equal to zero. But if the vectors $a_1, a_2, \ldots, a_k$ are not linearly independent, then they must be linearly dependent and there exist scalars $\lambda_j, j = 1, 2, \ldots, k$, with at least one non-zero $\lambda_j$ such that

$$\lambda_1a_1 + \lambda_2a_2 + \cdots, \lambda_ka_k = 0 \tag{3.6}$$

Let us suppose at least one $\lambda_j > 0$; if the non-zero $\lambda_j$ is not positive, then by multiplying (3.6) by $(-1)$, one can get a positive $\lambda_j$.

$$\text{Let}\ \mu = \underset{1 \le j \le k}{\text{Max}} \left\{ \frac{\lambda_j}{x_j} \right\}. \tag{3.7}$$

This $\mu$ is positive as $x_j > 0$ for all $j = 1, 2, \ldots, k$ and at least one $\lambda_j$ is positive. Dividing (3.6) by $\mu$ and subtracting it from (3.5), we have

$$\left(x_1 - \frac{\lambda_1}{\mu}\right)a_1 + \left(x_2 - \frac{\lambda_2}{\mu}\right)a_2 + \cdots + \left(x_k - \frac{\lambda_k}{\mu}\right)a_k = b$$

and hence

$$X_1 = \left[\left(x_1 - \frac{\lambda_1}{\mu}\right), \ldots, \left(x_k - \frac{\lambda_k}{\mu}\right), 0, 0, \ldots, 0\right] \tag{3.8}$$

is a solution of the system of equations $Ax = b$.

Again, from (3.7), we have

$\mu \geq \frac{\lambda_j}{x_j}$, for $\quad j = 1, 2, \ldots, k$

or $x_j \geq \frac{\lambda_j}{\mu}$, for $\quad j = 1, 2, \ldots, k$

or $x_j - \frac{\lambda_j}{\mu} \geq 0$, for $j = 1, 2, \ldots, k$.

This means that all the components of $X_1$ are non-negative and hence $X_1$ is a feasible solution of $Ax = b$, $x \geq 0$.

Again, for at least one value of $j$, we have, from (3.7), $\mu = \frac{\lambda_j}{x_j}$

or, in other words, $x_j - \frac{\lambda_j}{\mu} = 0$, for at least one value of $j$.

Hence, it is seen that the feasible solution $X_1$ will contain one more zero; i.e. the feasible solution $X_1$ cannot contain more than $(k-1)$ non-zero variables.

Therefore, the number of positive variables in an optimal solution can be reduced.

Now we have to show that $X_1$ is optimal. Let $z'$ be the value of the objective function for $x = X_1$.

Then

$$\begin{aligned} z' &= \sum_{j=1}^{k} c_j \left( x_j - \frac{\lambda_j}{\mu} \right) = \sum_{j=1}^{k} c_j x_j - \sum_{j=1}^{k} c_j \frac{\lambda_j}{\mu} \\ &= z^* - \frac{1}{\mu} \sum_{j=1}^{k} c_j \lambda_j \qquad [\because z^* = \sum_{j=1}^{k} c_j x_j]. \end{aligned} \tag{3.9}$$

If possible, let us suppose that

$$\sum_{j=1}^{k} c_j \lambda_j \neq 0. \tag{3.10}$$

Then for a suitable real number $\beta$, we can write

$$\beta(c_1\lambda_1 + c_2\lambda_2 + \cdots, + c_k\lambda_k) > 0,$$

That is, $c_1(\beta\lambda_1) + c_2(\beta\lambda_2) + \cdots, + c_k(\beta\lambda_k) > 0$

Adding $(c_1x_1 + c_2x_2 + \cdots, + c_kx_k)$ to both sides, we get

$$\begin{aligned} & c_1(x_1 + \beta\lambda_1) + c_2(x_2 + \beta\lambda_2) + \cdots, + c_k(x_k + \beta\lambda_k) \\ & > c_1x_1 + c_2x_2 + \cdots, + c_kx_k = z^*. \end{aligned} \tag{3.11}$$

Again, multiplying (3.6) by $\beta$ and adding it with (3.5), we get

$$(x_1 + \beta\lambda_1)a_1 + (x_2 + \beta\lambda_2)a_2 + \cdots, + (x_k + \beta\lambda_k)a_k = b$$

This relation implies that

$$[(x_1 + \beta\lambda_1), \ldots, (x_k + \beta\lambda_k), 0, 0, \ldots, 0]^T \tag{3.12}$$

is also the solution of the system $Ax = b$.

Now we choose $\beta$ in such a way that

$$\begin{aligned} & x_j + \beta\lambda_j \geq 0 \text{ for all } j = 1, 2, \ldots, k \\ & \text{or } \ \beta\lambda_j \geq -x_j \\ & \text{or } \ \beta \geq -\tfrac{x_j}{\lambda_j}, \text{ if } \lambda_j > 0, \\ & \beta \leq -\tfrac{x_j}{\lambda_j}, \text{ if } \lambda_j < 0 \end{aligned}$$

and $\beta$ is unrestricted, if $\lambda_j = 0$.

Then

$$\operatorname*{Max}_j\left\{-\frac{x_j}{\lambda_j}, \lambda_j > 0\right\} \leq \beta \leq \operatorname*{Min}_j\left\{-\frac{x_j}{\lambda_j}, \lambda_j < 0\right\}.$$

For this $\beta$, (3.12) becomes a feasible solution of $Ax = b, x \geq 0$.

Hence, from (3.11) it is clear that the feasible solution (3.12) gives a greater value of the objective function than $z^*$ for $x^*$. But this contradicts our assumption that $z^*$ is the optimal value, and thus we must have

$$\sum_{j=1}^{k} c_j\lambda_j = 0.$$

Hence, $X_1$ is also an optimal solution.

Thus, it is seen that from the given optimal solution, we can construct a new optimal solution, whose number of non-zero variables is less than that of the given solution. If the vectors associated with the new non-zero variables are linearly independent, then the new solution will be a basic feasible solution, and hence it proves the theorem.

If the new solution is not a basic feasible solution, then we can further diminish the number of non-zero variables as above to get a new set of optimal solutions. We may continue the process until the optimal solution obtained is a basic feasible solution. This proves the theorem.

For minimization problems, the theorem can be proved in a similar way.

## 3.6 Replacement of a Basis Vector

**Theorem** *Let an LPP have a basic feasible solution. If one of the basis vectors is removed from the basis and a non-basic vector is introduced in a basis, then the new solution obtained is also a basic feasible solution.*

*Proof* Let us consider the LPP

$$\begin{aligned}&\text{Maximize } z = \langle c^T, x\rangle\\&\text{subject to } Ax = b \text{ and } x \geq 0\\&\text{where } c, x^T \in R^n, b^T \in R^m \text{ and } A \text{ is an } m \times n \text{ real matrix.}\end{aligned}$$

Also, let the rank of $A$ be $m$.

Let $x_B$ be a basic feasible solution to the LPP.

$$\therefore Bx_B = b \tag{3.13}$$

where $B$ is the basis matrix.

Let $a_j$ be any column vector of $A$. Then

$$a_j = y_{1j}B_1 + y_{2j}B_2 + \cdots + y_{rj}B_r + \cdots + y_{mj}B_m \tag{3.14}$$

where $B_i \in B$ and $y_{ij}$ are suitable scalars.

Now, we know that if a basis vector $B_r$ with non-zero coefficient is replaced by $a_j \in A$, then the new set of vectors also forms a basis.

Now, for $y_{rj} \neq 0$, from (3.14) we can write

$$B_r = \frac{a_j}{y_{rj}} - \sum_{\substack{i=1\\ i\neq r}}^{m} \frac{y_{ij}}{y_{rj}} B_i. \tag{3.15}$$

From (3.13), $Bx_B = b$

which implies

$$\sum_{\substack{i=1\\ i\neq r}}^{m} x_{B_i}B_i + x_{B_r}B_r = b$$

$$\text{or} \quad \sum_{\substack{i=1\\ i\neq r}}^{m} x_{B_i}B_i + x_{B_r}\left[\frac{a_j}{y_{rj}} - \sum_{\substack{i=1\\ i\neq r}}^{m} \frac{y_{ij}}{y_{rj}}B_i\right] = b$$

$$\text{or} \quad \sum_{\substack{i=1\\ i\neq r}}^{m} \left[x_{B_i} - x_{B_r}\frac{y_{ij}}{y_{rj}}\right]B_i + \frac{x_{B_r}}{y_{rj}}a_j = b$$

From the preceding relation, it is clear that the components of the new basic solution $\hat{x}_B$ are given by

$$\hat{x}_{B_i} = x_{B_i} - x_{B_r}\frac{y_{ij}}{y_{rj}}, \quad i = 1, 2, \ldots, m, i \neq r$$

and $\hat{x}_{B_r} = \frac{x_{B_r}}{y_{rj}}$.

Now we shall show that $\hat{x}_B$ is also feasible; i.e. all the components of $\hat{x}_B$ are non-negative.

According to the value of $x_{B_r}$, two cases may arise.

**Case 1** $x_{B_r} = 0$.

In this case, $\hat{x}_{B_i} = x_{B_i} \geq 0$, $i = 1, 2, \ldots, m$, $i \neq r$ and $x_{B_r} = 0$
i.e. $\hat{x}_B \geq 0$.

**Case 2** $x_{B_r} \neq 0$.

In this case, we must have $y_{rj} > 0$.

For the remaining $y_{ij} (i \neq r)$, either $y_{ij} = 0$ or $y_{ij} \neq 0$ for $i \neq r$.

Now for $y_{ij} \neq 0$ for $i \neq r$, we have

$$\frac{x_{B_i}}{y_{ij}} \geq \frac{x_{B_r}}{y_{rj}} \text{ for } y_{ij} > 0 \text{ and } i \neq r$$
$$\text{and } \frac{x_{B_i}}{y_{ij}} \leq \frac{x_{B_r}}{y_{rj}} \text{ for } y_{ij} < 0 \text{ and } i \neq r.$$

Therefore, if we select $r(y_{rj} \neq 0)$ in such a way that $\frac{x_{B_r}}{y_{rj}} = \underset{i}{\text{Min}}\left\{\frac{x_{B_i}}{y_{ij}}; y_{ij} > 0,\ i \neq r\right\}$,

then the new set of basic variables is non-negative; hence, the basic solution $\hat{x}_B$ is feasible.

## 3.7 Improved Basic Feasible Solution

**Theorem** *Let $x_B$ be a basic feasible solution to the LPP:*

$$\begin{aligned}&\text{Maximize } z = \langle c^T, x\rangle\\&\text{subject to } Ax = b \text{ and } x \geq 0,\end{aligned}$$

*where $c, x^T \in R^n$, $b^T \in R^m$ and $A$ is an $m \times n$ real matrix.*

If $\hat{x}_B$ is another basic feasible solution obtained by introducing a non-basic column vector $a_j$ in the basis, for which $z_j - c_j$ is negative, then $\hat{x}_B$ is an improved basic feasible solution to the problem, i.e.

$$\left\langle \hat{c}_B^T, \hat{x}_B \right\rangle > \left\langle c_B^T, x_B \right\rangle.$$

*Proof* It is given that $x_B$ is a basic feasible solution.

Let $z_0 = \left\langle c_B^T, x_B \right\rangle$.

Let $\hat{a}_j$ be the column vector introduced in the basis such that $z_j - c_j < 0$.

Let $B_r$ be the vector removed from the basis and let $\hat{x}_B$ be the new basic feasible solution. Then

$$\hat{x}_{B_i} = x_{B_i} - x_{B_r}\frac{y_{ij}}{y_{rj}}\,(i \neq r) \text{ and } \hat{x}_{B_r} = \frac{x_{B_r}}{y_{rj}}.$$

Hence, the new value of the objective function is given by

$$\begin{aligned}\hat{z} = \sum_{i=1}^{m} \hat{c}_{B_i}\hat{x}_{B_i} &= \sum_{\substack{i=1\\ i\neq r}}^{m} c_{B_i}\hat{x}_{B_i} + \hat{c}_{B_r}\hat{x}_{B_r}\\ &= \sum_{\substack{i=1\\ i\neq r}}^{m} c_{B_i}\left(x_{B_i} - x_{B_r}\frac{y_{ij}}{y_{rj}}\right) + \hat{c}_{B_r}\frac{x_{B_r}}{y_{rj}}\\ &= \sum_{\substack{i=1\\ i\neq r}}^{m} c_{B_i}\left(x_{B_i} - x_{B_r}\frac{y_{ij}}{y_{rj}}\right) + c_j\frac{x_{B_r}}{y_{rj}} \qquad [\because \hat{c}_{B_r} = c_j]\\ &= \sum_{i=1}^{m} c_{B_i}x_{B_i} - \sum_{\substack{i=1\\ i\neq r}}^{m} c_{B_i}x_{B_r}\frac{y_{ij}}{y_{rj}} - c_{B_r}x_{B_r} + c_j\frac{x_{B_r}}{y_{rj}}\end{aligned}$$

$$
\begin{aligned}
&= z_0 - \sum_{i=1}^{m} c_{B_i} y_{ij} \frac{x_{B_r}}{y_{rj}} + c_j \frac{x_{B_r}}{y_{rj}} \\
&= z_0 - \left( \sum_{i=1}^{m} c_{B_i} y_{ij} - c_j \right) \frac{x_{B_r}}{y_{rj}} \\
&= z_0 - \left( z_j - c_j \right) \frac{x_{B_r}}{y_{rj}} \\
&> z_0.
\end{aligned}
$$

Hence, the new basic feasible solution $\hat{x}_B$ gives an improved value of the objective function.

**Corollary** *If $z_j - c_j = 0$ for at least one $j$ corresponding to the non-basis vector for which $y_{ij} > 0$, $i = 1, 2, \ldots, m$, then a new basic feasible solution is obtained which gives an unchanged value of the objective function.*

*Proof* From the preceding theorem,

$$
\begin{aligned}
\hat{z} &= z_0 - \left( z_j - c_j \right) \frac{x_{B_r}}{y_{rj}} \quad , y_{rj} > 0 \\
&= z_0 \quad [\because z_j - c_j = 0].
\end{aligned}
$$

**Note** According to the preceding theorem, for the maximum improvement, determine the entering vector corresponding to the net evolution as follows:

$$
z_k - c_k = \operatorname*{Min}_{j} \left\{ z_j - c_j : z_j - c_j < 0 \right\}.
$$

## 3.8 Optimality Condition

**Theorem** *For a basic feasible solution $x_B$ of an LPP*

$$
\begin{aligned}
&\text{Maximize } z = \left\langle c^T, x \right\rangle \\
&\text{subject to } Ax = b \text{ and } x \geq 0 \\
&\text{where } c, x^T \in R^n \text{ and } b^T \in R^m,
\end{aligned}
$$

if $z_j - c_j \geq 0$ for every column vector $a_j$ of $A$, then $x_B$ is an optimal solution.

*Proof* Let $B = (B_1, B_2, \ldots, B_m)$ be the basis matrix corresponding to the basic feasible solution $x_B$ so that $Bx_B = b$.

Also, let $z_B$ be the value of the objective function corresponding to $x_B$.

Hence, $z_B = \left\langle c_B^T, x_B \right\rangle = \sum_{i=1}^{m} c_{B_i} x_{B_i}$.

Let $X_1 = \left[x_1', x_2', \ldots, x_n'\right]^T$ be any other feasible solution of the LPP and let $z'$ be the value of the objective function corresponding to the solution. Then we have $Bx_B = b = AX_1$

so that

$$x_B = B^{-1}(AX_1) = (B^{-1}A)X_1 = yX_1, \tag{3.16}$$

where $B^{-1}A = y = [y_1, y_2, \ldots, y_n]$.

Now equating the $i$th component of (3.16) from both sides, we have

$$x_{B_i} = \sum_{j=1}^{n} y_{ij} x_j'. \tag{3.17}$$

It is given that for every column vector $a_j$ of $A$,

$$z_j - c_j \geq 0, \text{ i.e. } z_j \geq c_j, \; j = 1, 2, \ldots, n. \tag{3.18}$$

Hence, for all $j = 1, 2, \ldots, n$,

$$\begin{aligned}
& z_j x_j' \geq c_j x_j', \quad \text{as } x_j' \geq 0 \\
\text{or} \quad & \sum_{j=1}^{n} z_j x_j' \geq \sum_{j=1}^{n} c_j x_j' \\
\text{or} \quad & \sum_{j=1}^{n} x_j'(c_B^T y_j) \geq z' \\
\text{or} \quad & \sum_{j=1}^{n} x_j' \left( \sum_{i=1}^{m} c_{B_i} y_{ij} \right) \geq z' \\
\text{or} \quad & \sum_{i=1}^{m} c_{B_i} \left( \sum_{j=1}^{n} x_j' y_{ij} \right) \geq z' \\
\text{or} \quad & \sum_{i=1}^{m} c_{B_i} x_{B_i} \geq z', \quad \left[ \text{From (3.17) } x_{B_i} = \sum_{j=1}^{n} y_{ij} x_j' \right] \\
\text{or} \quad & z_B \geq z'.
\end{aligned}$$

This shows that the value of $z_B$ is greater than the objective function for any other feasible solution, and hence $z_B$ is the maximum value of the objective function.

For a minimization problem, the optimality condition is as follows.

**Theorem** *For a basic feasible solution $x_B$ of an LPP*

$$\begin{aligned}&\text{Minimize}\, z = \langle c^T, x\rangle\\&\text{subject to}\, Ax = b\,\text{and}\, x \geq 0,\\&\text{where}\, c, x^T \in R^n, b^T \in R^m\,\text{and}\, A\,\text{is an}\, m \times n\,\text{matrix},\end{aligned}$$

*if $z_j - c_j \leq 0$ for every column vector $a_j$ of $A$, then $x_B$ is an optimal solution.*

## 3.9 Unbounded Solution

If the objective function has no finite optimum value and it can be increased or decreased arbitrarily, then we say that the problem has an unbounded solution.

**Theorem** *If at any iteration of the simplex algorithm, $z_j - c_j < 0$ for at least one $j$, and for this $j$, $y_{ij} < 0$ for all $i = 1, 2, \ldots, m$, then the LPP admits an unbounded solution in a maximization problem.*

*Proof* Let $x_B$ be the basic feasible solution of an LPP with basis matrix $B$ at any iteration in which $z_j - c_j < 0$ and $y_{ij} < 0$, for all $i = 1, 2, \ldots, m$. If $z_B$ is the value of the objective function for this basic feasible solution, then

$$z_B = \langle c_B^T, x_B\rangle \,\text{and}\, Bx_B = b. \tag{3.19}$$

Again, $Bx_B = b$ implies that

$$\sum_{i=1}^{m} x_{B_i} B_i = b, \tag{3.20}$$

where $B = (B_1, B_2, \ldots, B_m)$, i.e. the $B_i$'s are column vectors of $B$.

In (3.20), adding and subtracting $\gamma a_j$, where $\gamma$ is a positive scalar, we get

$$\sum_{i=1}^{m} x_{B_i} B_i + \gamma a_j - \gamma a_j = b. \tag{3.21}$$

Again, we know that $a_j = \sum_{i=1}^{m} y_{ij} B_j$,

i.e. $-\gamma a_j = -\gamma \sum_{i=1}^{m} y_{ij} B_i$.

Using this result in (3.21), we have

$$\sum_{i=1}^{m} x_{B_i}B_i + \gamma a_j - \gamma \sum_{i=1}^{m} y_{ij}B_i = b$$
$$\text{or} \quad \sum_{i=1}^{m} (x_{B_i} - \gamma y_{ij})B_i + \gamma a_j = b. \tag{3.22}$$

If now $\gamma$ is assumed to be positive and since $y_{ij} < 0$, for $i = 1, 2, \ldots, m$, then we have

$$x_{B_i} - \gamma y_{ij} \geq 0. \tag{3.23}$$

Now, from (3.22) and (3.23), it is seen that we have obtained a new set of feasible solutions in which the $(m+1)$ variables $x_{B_i} - \gamma y_{ij}$, together with $\gamma$, can be different from zero and the remaining components are zero. This solution in general is feasible but not a basic solution. Then

$$\begin{aligned} z' &= \sum_{i=1}^{m} c_{B_i}(x_{B_i} - \gamma y_{ij}) + c_j\gamma \\ &= \sum_{i=1}^{m} c_{B_i}x_{B_i} - \gamma \sum_{i=1}^{m} c_{B_i}y_{ij} + c_j\gamma \\ &= z_B - \gamma(z_j - c_j). \end{aligned} \tag{3.24}$$

From (3.24), it is observed that $z'$ will always be greater than $z_B$ as $\gamma > 0$ and $(z_j - c_j) < 0$.

As $\gamma$ increases, $z'$ also goes on increasing, and by making $\gamma$ sufficiently large, we can make $z'$ as large as desired. Hence, we conclude that under such circumstances there is no upper bound of the objective function, although the solution is feasible.

Hence, the problem admits an unbounded solution.

For a minimization problem, the corresponding property is as follows:

If at any iteration of the simplex algorithm we get $(z_j - c_j) > 0$ for at least one $j$, and for this $j, y_{ij} < 0$, for all $i = 1, 2, \ldots, m$, then the LPP admits an unbounded solution in a minimization problem.

**Theorem** *Any convex combination of k different optimal solutions to an LPP is again an optimum solution to the problem.*

*Proof* Let $X_1, X_2, \ldots, X_k$ be $k$ different optimum solutions of the following LPP:

Maximize $z = \langle c^T, x \rangle$
subject to $Ax = b$ and $x \geq 0$,
where $c, x^T \in R^n, b^T \in R^m$ and $A$ is an $m \times n$ real matrix.

Let $z_0$ be the optimum value of $z$.

Hence,

$$z_0 = \langle c, X_1 \rangle = \langle c, X_2 \rangle = \cdots = \langle c, X_k \rangle \tag{3.25}$$

and

$$AX_1 = AX_2 = \cdots = AX_k = b, \tag{3.26}$$

where

$$X_1 \geq 0, X_2 \geq 0, \ldots, X_k \geq 0.$$

Now, any combination of $X_1, X_2, \ldots, X_k$ can be written as

$$x = \sum_{i=1}^{k} \alpha_i X_i, \text{ where the } \alpha_i\text{'s are positive scalars such that } \sum_{i=1}^{k} \alpha_i = 1.$$

Now

$$\begin{aligned} Ax = A\left(\sum_{i=1}^{k} \alpha_i X_i\right) &= \sum_{i=1}^{k} \alpha_i (AX_i) \\ &= \sum_{i=1}^{k} \alpha_i b \quad [\text{using (3.26)}] \\ &= b \sum_{i=1}^{k} \alpha_i = b, \quad \text{since } \sum_{i=1}^{k} \alpha_i = 1. \end{aligned}$$

Again, since $X_i \geq 0$ and $\alpha_i \geq 0$ for all $i = 1, 2, \ldots, k$, then every component of $x$ will be non-negative.

Thus, $x$ is a feasible solution to the LPP.

Again,

$$\begin{aligned} \langle c^T,\ x \rangle = \left\langle c^T,\ \sum_{i=1}^{k} \alpha_i X_i \right\rangle &= \sum_{i=1}^{k} \langle c^T,\ \alpha_i X_i \rangle \\ = \sum_{i=1}^{k} \alpha_i \langle c^T,\ X_i \rangle &= \sum_{i=1}^{k} \alpha_i z_0 = z_0 \quad \left[\because \sum_{i=1}^{k} \alpha_i = 1\right]. \end{aligned}$$

Hence, a convex combination of $k$ different optimum solutions to an LPP is also an optimal solution to the LPP.

## 3.10 Simplex Algorithm

The iterative procedure for finding an optimal basic feasible solution (BFS) to an LPP is as follows:

Step 1: Check whether the given LPP is a maximization or minimization problem. If it is a minimization problem, then convert it into a problem of maximization, using

$$\text{Minimize}\, z = -\text{Maximize}\,(-z),$$

where $z$ is the objective function of the given LPP.

Step 2: Check whether all $b_i\,(i = 1, 2, \ldots, m)$ are non-negative or not. If any one of $b_i$ is negative, then multiply the corresponding inequality/equation of the constraint by $(-1)$ so that all $b_i\,(i = 1, 2, \ldots, m)$ will be positive.

Step 3: Convert all the inequality constraints into equations by introducing slack and/or surplus variables in the constraints. Put the coefficients of these variables in the objective function as zero.

Step 4: Obtain an initial BFS $x_B$ using $x_B = B^{-1}b$, where $B$ is the initial unit basis matrix. Alternatively, by taking non-basic variables as zero, obtain the initial BFS.

Step 5: Construct the simplex table.

Step 6: Compute the net evaluations $z_j - c_j\,(j = 1, 2, \ldots, n)$ by using the formula $z_j - c_j = \langle c_B, y_j \rangle - c_j$ and examine their sign. Three cases may arise:

- (i) If all $z_j - c_j \geq 0$ and either no artificial variable is in the basis or artificial variables appear in the basis at zero level, then the current BFS is optimal. Stop the process.
- (ii) If all $z_j - c_j \geq 0$ but artificial variable(s) are in the basis at a positive level, then the given LPP has no feasible solution.
- (iii) If at least one $z_j - c_j < 0$, go to the next step.

Step 7: If at least one $z_j - c_j < 0$, choose the most negative of $z_j - c_j$. Let it be $z_r - c_r$ for some $j = r$.
Two cases may arise:

- (i) If all $y_{ir} \leq 0\,(i = 1, 2, \ldots, m)$, then there is an unbounded solution to the given LPP. Stop the process.
- (ii) If at least one $y_{ir} > 0$, then the corresponding vector $y_r$ enters the basis. This vector is known as an incoming or entering vector.

Step 8: To select the departing or outgoing vector, compute $\underset{i}{\text{Min}}\left\{\frac{x_{B_i}}{y_{ir}}, y_{ir} > 0\right\}$. Let the minimum value be $\frac{x_{B_k}}{y_{kr}}$. Then the vector $y_k$ will leave the basis. This vector is known as a departing or outgoing vector. The common element $y_{kr}$ which is in the $k$th row and the $r$th column is known as a key or leading element.

Step 9: Convert the key element to unity by dividing all the elements of its row by the key element and all other elements in its column to zeros using the following operations:

$$\hat{y}_{ij} = y_{ij} - \frac{y_{kj}}{y_{kr}} y_{ir},\ i = 1, 2, \ldots, m,\ i \neq k,\ j = 1, 2, \ldots, n$$

and $\hat{y}_{kj} = \frac{y_{kj}}{y_{kr}}$, $j = 1, 2, \ldots, n$.

Then modify the components of $x_B$ by using the relations

$$\hat{x}_{B_i} = x_{B_i} - \frac{y_{kj}}{y_{kr}} x_{B_k},\ i = 1, 2, \ldots, m\ (i \neq k) \text{ and } \hat{x}_{B_k} = \frac{x_{B_k}}{y_{kr}}.$$

Using these values, construct the next simplex table and go to Step 6.

## 3.11 Simplex Table

Let us consider the following standard LPP:

Maximize $z = \langle c^T, x \rangle$
subject to $Ax = b$ and $x \geq 0$
where $c = (c_1, c_2, \ldots, c_n)$ = the profit vector
$x = (x_1, x_2, \ldots, x_n)^T$ = the decision vector
$A = (a_1, a_2, \ldots, a_n)$ = the coefficient matrix
$b = (b_1, b_2, \ldots, b_m)^T$ = the requirement vector
$c_B = (c_{B_1}, c_{B_2}, \ldots, c_{B_m})^T$ = the m-component column vector with elements corresponding to the basic variables
$B = (B_1, B_2, \ldots, B_m)$ = the basis matrix
$x_B = B^{-1}b$ = the basic feasible solution
$y_j = B^{-1}a_j, j = 1, 2, \ldots, n.$

| $c_j$ | | | $c_1$ | $c_2$ | ... | $c_j$ | ... | $c_n$ |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | ... | $y_j$ | ... | $y_n$ |
| $c_{B_1}$ | $y_{B_1}$ | $x_{B_1}$ | $y_{11}$ | $y_{12}$ | ... | $y_{1j}$ | ... | $y_{1n}$ |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| $c_{B_m}$ | $y_{B_m}$ | $x_{B_m}$ | $y_{m1}$ | $y_{m2}$ | ... | $y_{mj}$ | ... | $y_{mn}$ |
| $z_B = \langle c_B^T, x_B \rangle$ | | | $z_1 - c_1$ | $z_2 - c_2$ | ... | $z_j - c_j$ | ... | $z_n - c_n$ |

Initially, $B = I_m$ = a unit matrix of order $m$

For the initial simplex table, $x_B = B^{-1}b = I_m b = b$ $[\because B = I_m \therefore B^{-1} = I_m]$

$$y_j = B^{-1}a_j = a_j.$$

*Example 1* Solve the following LPP:

$$\begin{aligned}&\text{Maximize}\, z = 5x_1 + 4x_2\\&\text{subject to}\, 3x_1 + 4x_2 \le 24\\&\qquad 3x_1 + 2x_2 \le 18\\&\qquad x_2 \le 5\,\text{and}\, x_1,\ x_2 \ge 0.\end{aligned}$$

**Solution:** Here the given problem is:

$$\begin{aligned}&\text{Maximize}\, z = 5x_1 + 4x_2\\&\text{subject to}\, 3x_1 + 4x_2 \le 24\\&\qquad 3x_1 + 2x_2 \le 18\\&\qquad x_2 \le 5\,\text{and}\, x_1,\ x_2 \ge 0.\end{aligned}$$

This is a maximization problem. All the $b_i$'s are positive. Now introducing the slack variables $x_3$, $x_4$, $x_5$, one to each constraint, we get the system of equations as follows:

$$\begin{aligned}3x_1 + 4x_2 + x_3 &= 24\\3x_1 + 2x_2 + x_4 &= 18\\x_2 + x_5 &= 5.\end{aligned}$$

Then the new objective function is

$$z = 5x_1 + 4x_2 + 0\cdot x_3 + 0\cdot x_4 + 0\cdot x_5.$$

The set of above equations (1) can be written as

$$\begin{pmatrix}3 & 4 & 1 & 0 & 0\\3 & 2 & 0 & 1 & 0\\0 & 1 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix} = \begin{pmatrix}24\\18\\5\end{pmatrix}$$

$$\text{or}\ A\ x = b,$$

where $A = \begin{pmatrix}3 & 4 & 1 & 0 & 0\\3 & 2 & 0 & 1 & 0\\0 & 1 & 0 & 0 & 1\end{pmatrix} \sim \begin{pmatrix}0 & 0 & 1 & 0 & 0\\3 & 2 & 0 & 1 & 0\\0 & 1 & 0 & 0 & 1\end{pmatrix} \sim \begin{pmatrix}0 & 0 & 1 & 0 & 0\\0 & 0 & 0 & 1 & 0\\0 & 0 & 0 & 0 & 1\end{pmatrix}.$

Hence, the rank of $A$ is 3. Then $\begin{pmatrix}1\\0\\0\end{pmatrix}$, $\begin{pmatrix}0\\1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\0\\1\end{pmatrix}$ are three linearly independent column vectors of $A$. Therefore, the submatrix $B = \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$ is a non-singular basis submatrix of $A$. The basic variables are therefore $x_3, x_4, x_5$ and an obvious initial basic feasible solution is

$$\begin{aligned}x_B &= B^{-1}b\\ &= \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}24\\18\\5\end{pmatrix} = \begin{pmatrix}24\\18\\5\end{pmatrix}\end{aligned}$$

$$\text{i.e.}\ \begin{pmatrix}x_3\\x_4\\x_5\end{pmatrix} = \begin{pmatrix}24\\18\\5\end{pmatrix} \quad \therefore x_3 = 24,\ x_4 = 18,\ x_5 = 5$$

and $c_B = (c_3, \quad c_4, \quad c_5) = (0 \quad 0 \quad 0)$.

Now we construct the starting simplex table.

| $c_j$ | | | 5 | 4 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_3$ | $x_3 = 24$ | 3 | 4 | 1 | 0 | 0 | |
| 0 | $y_4$ | $x_4 = 18$ | 3 | 2 | 0 | 1 | 0 | |
| 0 | $y_5$ | $x_5 = 5$ | 0 | 1 | 0 | 0 | 1 | |
| $z_B = 0$ | | | −5 | −4 | 0 | 0 | 0 | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j$ are not greater than or equal to zero. Hence, the current solution is not optimal. Now, $\text{Minimum}\{z_j - c_j,\ z_j - c_j < 0\} = \text{Minimum}\{-5, -4\} = -5$ which is $z_1 - c_1$.

Hence, $y_1$ is the incoming vector.

Again, $\text{Minimum}\left\{\frac{x_{B_i}}{y_{i1}},\ y_{i1} > 0\right\} = \text{Minimum}\left(\frac{24}{3}, \frac{18}{3}\right) = 6$,

which occurs for the element $y_{21} (= 3)$. Hence, $y_{21} = 3$ is the key element and $y_4$ is the outgoing vector. In the key row, we then replace $0,\ y_4,\ x_4$ under $c_B, y_B, x_B$ by $5,\ y_1,\ x_1$ respectively, the corresponding elements of the key column.

Now converting the key element to unity and all other elements of this column to zero by suitable operations, we get the new simplex table as follows:

| $c_j$ | | | 5 | 4 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_3$ | $x_3 = 6$ | 0 | 2 | 1 | $-1$ | 0 | |
| 5 | $y_1$ | $x_1 = 6$ | 1 | 2/3 | 0 | 1/3 | 0 | |
| 0 | $y_5$ | $x_5 = 5$ | 0 | 1 | 0 | 0 | 1 | |
| $z_B = 30$ | | | 0 | $-2/3$ | 0 | 5/3 | 0 | $\leftarrow z_j - c_j$ |

Here $z_2 - c_2 < 0$. Hence, the current solution is not optimal.

Now Minimum $\{z_j - c_j, z_j - c_j < 0\}$ = Minimum $\{-2/3\}$ which is $z_2 - c_2$. Hence, $y_2$ is the incoming vector.

Again, Minimum $\left\{\frac{x_{B_i}}{y_{i2}},\ y_{i2} > 0\right\}$ = Minimum $\left(\frac{6}{2}, \frac{6}{2/3}, \frac{5}{1}\right) = 3$, which occurs for the element $y_{12}(=2)$. Hence, $y_{12} = 2$ is the key element and $y_3$ is the outgoing vector.

In the key row, we then replace 0, $y_3$, $x_3$ under $c_B$, $y_B$, $x_B$ by 4, $y_2$, $x_2$ respectively, the corresponding elements of the key column.

Now converting the key element to unity and all other elements of this column to zero by suitable operations, we construct the new simplex table as follows:

| $c_j$ | | | 5 | 4 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 4 | $y_2$ | $x_2 = 3$ | 0 | 1 | 1/2 | $-1/2$ | 0 | |
| 5 | $y_1$ | $x_1 = 4$ | 1 | 0 | $-1/3$ | 2/3 | 0 | |
| 0 | $y_5$ | $x_5 = 2$ | 0 | 0 | $-1/2$ | 1/2 | 1 | |
| $z_B = 32$ | | | 0 | 0 | 1/3 | 4/3 | 0 | $\leftarrow z_j - c_j$ |

Here all the $z_j - c_j \geq 0$. Hence, the current BFS is optimal, and the optimal solution is

$$x_1 = 4,\ \ x_2 = 3 \text{ and Max } z = 32.$$

*Example 2* Solve the following LPP:

$$\begin{aligned} \text{Maximize } z &= 10x_1 + x_2 + 2x_3 \\ \text{subject to } x_1 + x_2 - 2x_3 &\leq 10 \\ 4x_1 + x_2 + x_3 &\leq 20 \\ \text{and } x_1, x_2, x_3 &\geq 0. \end{aligned}$$

**Solution:**

This is a maximization problem. All the $b_i$'s are positive. Now introducing the slack variables $x_4$, $x_5$, one to each constraint, we get the problem in the standard form as follows:

$$\begin{aligned}\text{Maximize}\, z &= 10x_1 + x_2 + 2x_3 + 0x_4 + 0x_5\\ \text{subject to}\, x_1 + x_2 - 2x_3 + x_4 &= 10\\ 4x_1 + x_2 + x_3 + x_5 &= 20\\ x_1, x_2, x_3, x_4, x_5 &\geq 0.\end{aligned}$$

Clearly, we may consider the variables $x_4$ and $x_5$ as basic variables. Considering the other variables as non-basic, we get $x_4 = 10$ and $x_5 = 20$, which is the initial basic feasible solution.

Now we construct the initial simplex table.

| $c_j$ | | | 10 | 1 | 2 | 0 | 0 | Mini ratio |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_4$ | $x_4$ = 10 | 1 | 1 | −2 | 1 | 0 | 10/1 = 10 |
| 0 | $y_5$ | $x_5$ = 20 | 4 | 1 | 1 | 0 | 1 | 20/4 = 5 |
| $z_B = 0$ | | | −10 | −1 | −2 | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_1$ is the incoming vector, $y_5$ is the outgoing vector and $y_{21} = 4$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | 10 | 1 | 2 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_4$ | $x_4$ = 5 | 0 | 3/4 | −9/4 | 1 | −1/4 | |
| 10 | $y_1$ | $x_1$ = 5 | 1 | 1/4 | 1/4 | 0 | 1/4 | |
| $z_B = 50$ | | | 0 | 6/4 | 2/4 | 0 | 10/4 | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j \geq 0$. Hence, an optimum solution has been attained, and this solution is given by $x_1 = 5,\ x_2 = 0,\ x_3 = 0$ and Max $Z = 50$.

*Example 3*

$$\begin{aligned}\text{Maximize}\, z &= 3x_1 + x_2 + 3x_3\\ \text{subject to}\, 2x_1 + x_2 + x_3 &\leq 2\\ x_1 + 2x_2 + 3x_3 &\leq 5\\ 2x_1 + 2x_2 + x_3 &\leq 6 \text{ and } x_1,\, x_2, x_3 \geq 0.\end{aligned}$$

**Solution:**

This is a maximization problem. All the $b_i$'s are positive. Now introducing the slack variables $x_4$, $x_5, x_6$, one to each constraint, we get the problem in the standard form as follows:

$$\begin{aligned} \text{Maximize } z &= 3x_1 + x_2 + 3x_3 + 0x_4 + 0x_5 + 0x_6 \\ \text{subject to } & 2x_1 + x_2 + x_3 + x_4 = 2 \\ & x_1 + 2x_2 + 3x_3 + x_5 = 5 \\ & 2x_1 + 2x_2 + x_3 + x_6 = 6 \text{ and } x_i \geq 0 i = 1, 2, \ldots, 6. \end{aligned}$$

Let us take the variables $x_4, x_5$ and $x_6$ as basic variables. Considering the other variables as non-basic, we get $x_4 = 2, x_5 = 5$ and $x_6 = 6$, which is the initial basic feasible solution.

Now we construct the initial simplex table.

| $c_j$ | | | 3 | 1 | 3 | 0 | 0 | 0 | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 0 | $y_4$ | $x_4 = 2$ | 2 | 1 | 1 | 1 | 0 | 0 | 2/2 = 1 |
| 0 | $y_5$ | $x_5 = 5$ | 1 | 2 | 3 | 0 | 1 | 0 | 5/1 = 5 |
| 0 | $y_6$ | $x_6 = 6$ | 2 | 2 | 1 | 0 | 0 | 1 | 6/2 = 3 |
| $z_B = 0$ | | | −3 | −1 | −3 | 0 | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_1$ is the incoming vector, $y_4$ is the outgoing vector and $y_{11} = 2$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | 3 | 1 | 3 | 0 | 0 | 0 | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 3 | $y_1$ | $x_1 = 1$ | 1 | 1/2 | 1/2 | 1/2 | 0 | 0 | 1 |
| 0 | $y_5$ | $x_5 = 4$ | 0 | 3/2 | 5/2 | −1/2 | 1 | 0 | 8/5 |
| 0 | $y_6$ | $x_6 = 4$ | 0 | 1 | 0 | −1 | 0 | 1 | – |
| $z_B = 3$ | | | 0 | 1/2 | −3/2 | 3/2 | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_3$ is the incoming vector, $y_5$ is the outgoing vector and $y_{23} = 5/2$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | 3 | 1 | 3 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 3 | $y_1$ | $x_1 = 1/5$ | 1 | 1/5 | 0 | 3/5 | 1/5 | 0 | |
| 3 | $y_3$ | $x_3 = 8/5$ | 0 | 3/5 | 1 | −1/5 | 2/5 | 0 | |
| 0 | $y_6$ | $x_6 = 4$ | 0 | 1 | 0 | −1 | 0 | 1 | |
| $z_B = 27/5$ | | | 0 | 7/5 | 0 | 6/5 | 9/5 | 0 | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j \geq 0$. Hence, the optimality condition has been satisfied, and the optimal solution is

$$x_1 = 1/5,\ x_2 = 0,\ x_3 = 8/5 \text{ and Max } z = \frac{27}{5}.$$

*Example 4* Use simplex method to solve the following LPP:

$$\begin{aligned} \text{Minimize} \quad & z = x_1 - 3x_2 + 2x_3 \\ \text{subject to} \quad & 3x_1 + x_2 + 2x_3 \leq 7 \\ & -2x_1 + 4x_2 \leq 12 \\ & -4x_1 + 3x_2 + 8x_3 \leq 10 \text{ and } x_1,\ x_2,\ x_3 \geq 0. \end{aligned}$$

**Solution:** This is a minimization problem. All the $b_i$'s are positive.

Let $z' = -z = -x_1 + 3x_2 - 2x_3$.

Hence,

$$\text{Minimize } z = -\text{Maximize } z'.$$

Now introducing the slack variables $x_4$, $x_5, x_6$, one to each constraint, we get the reduced problem in the standard form as follows:

$$\begin{aligned} \text{Maximize } z' = -x_1 + 3x_2 - 2x_3 + 0x_4 + 0x_5 + 0x_6 \\ \text{subject to } 3x_1 + x_2 + 2x_3 + x_4 &= 7 \\ -2x_1 + 4x_2 + x_5 &= 12 \\ -4x_1 + 3x_2 + 8x_3 + x_6 &= 10 \\ \text{and } x_1,\ x_2,\ x_3,\ \ldots,\ x_6 \geq 0. \end{aligned}$$

Clearly, we can select the variables $x_4, x_5$ and $x_6$ as the basic variables. Considering the other variables as non-basic, we get $x_4 = 7, x_5 = 12$ and $x_6 = 10$, which is the initial basic feasible solution.

Now we construct the initial simplex table.

| $c_j$ | | | −1 | 3 | −2 | 0 | 0 | 0 | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 0 | $y_4$ | $x_4 = 7$ | 3 | −1 | 2 | 1 | 0 | 0 | – |
| 0 | $y_5$ | $x_5 = 12$ | −2 | 4 | 0 | 0 | 1 | 0 | 12/4 = 3 |
| 0 | $y_6$ | $x_6 = 10$ | −4 | 3 | 8 | 0 | 0 | 1 | 10/3 = 10/3 |
| $z'_B = 0$ | | | 1 | −3 | 2 | 0 | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_2$ is the incoming vector, $y_5$ is the outgoing vector and $y_{22} = 4$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | −1 | 3 | −2 | 0 | 0 | 0 | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 0 | $y_4$ | $x_4 = 10$ | 5/2 | 0 | 2 | 1 | 1/4 | 0 | 10/5/2 = 4 |
| 3 | $y_2$ | $x_2 = 3$ | −1/2 | 1 | 0 | 0 | 1/4 | 0 | – |
| 0 | $y_6$ | $x_6 = 1$ | −5/2 | 0 | 8 | 0 | −1/4 | 1 | – |
| $z'_B = 9$ | | | −1/2 | 0 | 0 | 0 | 3/4 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_1$ is the incoming vector, $y_4$ is the outgoing vector and $y_{11} = 5/2$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | −1 | 3 | −2 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| −1 | $y_1$ | $x_1 = 4$ | 1 | 0 | 4/5 | 2/5 | 1/10 | 0 | |
| 3 | $y_2$ | $x_2 = 5$ | 0 | 1 | 2/5 | 1/5 | 3/10 | 0 | |
| 0 | $y_6$ | $x_6 = 11$ | 0 | 0 | 10 | 1 | −1/2 | 1 | |
| $z'_B = 11$ | | | 0 | 0 | 11/5 | 1/5 | 4/5 | 0 | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j \geq 0$. Hence, optimality has been reached, and the optimal solution is

$$x_1 = 4,\ x_2 = 5,\ x_3 = 0 \,\text{and}\, \text{Min}\, z = -\text{Max}\, z' = -11.$$

## 3.12 Use of Artificial Variables

In the standard form of an LPP, if the coefficient matrix does not contain a unit basis matrix, then a non-negative variable is added to the left side of each equation which lacks the starting basic variables. This variable is known as an artificial variable. This variable plays the same role as a slack variable in providing the initial basic feasible solution. However, since such artificial variables have no physical meaning from the standpoint of the original problem, the method will be valid only if these variables are driven out or at zero level when the optimum solution is attained. In other words, to return to the original problem, artificial variables must be driven out or at zero level in the final solution. Otherwise, the resulting solution may be infeasible.

For solving an LPP involving artificial variables, the following two methods are generally employed:

(i) Big M method (also known as Charne's Big M method or the method of penalties)
(ii) Two-phase method.

### *3.12.1 Big M Method*

The Big M method is an alternative method for solving an LPP having artificial variables. This method is applied only when the coefficient matrix of standard form of an LPP does not give the unit basis matrix. In this method, a minimum number of artificial variables are introduced, and a very large negative price, say $-M$ ($M$ is positive and very large), is assigned in the objective function corresponding to the artificial variable. Due to the very large negative price, the objective function cannot be improved if one or more artificial variables appear in the basis. The corresponding problem can be solved in the usual simplex algorithm, and finally the following cases may arise:

(i) At any iteration, if the optimality conditions are not satisfied, we have to proceed further to get an optimal solution by omitting the vectors corresponding to the artificial variables. Here the artificial variables are used as agents to get a unit basis.
(ii) At any iteration, if the optimality conditions are satisfied and all the vectors corresponding to artificial variables are driven out from the basis, the given LPP has no feasible solution.
(iii) At any iteration, if the optimality conditions are satisfied and all the artificial variables are at zero level but at least one artificial vector is present in the basis, the solution obtained is an optimal solution of the given LPP.
(iv) At any iteration, if the optimality conditions are satisfied and some of the artificial vectors are present in the basis, i.e. some of the vectors corresponding to artificial variables are at a positive level, then the given LPP has no feasible solution.

### 3.12.2 Solution of Simultaneous Linear Equations by Simplex Method

Let us consider the system of $m$ simultaneous linear equations in $n$ unknowns in the form

$$Ax = b, \tag{3.27}$$

where $A$ is an $m \times n$ real matrix and $b$ is an $m \times 1$ real matrix. The variables to be determined are given by the vector $x$. Now we shall solve this by the simplex method. For this purpose, we have to form an LPP with a dummy objective function which is to be optimized.

To solve the given system of equations, we introduce artificial variables in the given system of equations, and the corresponding price in the dummy objective function $z$ will be $-M$ for each artificial variable. Thus, the given problem of solving the system of linear equations can be reduced to an LPP in which the objective function will be

$$z = -M \sum_{i=1}^{m} x_{a_i},$$

where $x_{a_i}$ $(i = 1, 2, \ldots, m)$ is the artificial variable.

Since there is no non-negativity restriction for the required variables $x$, these variables are considered to be unrestricted in sign. Now to introduce the non-negativity restrictions, let us assume that

$$x = x' - x'' \text{ where } x',\ x'' \geq 0.$$

Therefore, the transformed LPP becomes

$$\begin{aligned} &\text{Maximize}\, z = -M \sum_{i=1}^{m} x_{a_i} \\ &\text{subject to}\, A(x' - x'') + x_a = b \\ &\text{and}\, x',\ x'',\ x_a \geq 0, \end{aligned}$$

where $x_a$ is the vector of artificial variables.

Solving this LPP by the Big M method:

(i) If we get Max $z = 0$ for which $x_{a_i}(i = 1,\ 2,\ \ldots,\ m)$ becomes zero, then a basic feasible solution of the system of equations will be obtained.
(ii) At any iteration, if the optimality conditions are satisfied and some of the artificial vectors are present in the basis, then the given system of equations has no solution.

### 3.12.3 Inversion of a Matrix by Simplex Method

To find the inverse of a given non-singular square matrix, an identity matrix of the same order is considered. Now, performing the same sequence of elementary row transformations on the given matrix and the identity matrix to reduce the given matrix to an identity matrix, we get the inverse matrix.

Let us consider a real non-singular square matrix of order $n$. Let $x$ be an $n$-component column vector and $b$ be an $n$-component dummy column vector. Then let us consider the system of linear equations as follows:

$$Ax = b, \quad x \geq 0. \tag{3.28}$$

Now we introduce the artificial vectors $x_a(\geq 0)$ and construct the dummy objective function as

$$z = -M \sum_{i=1}^{m} x_{a_i},$$

where $x_{a_i}(\geq 0)$ is the $i$th component of the artificial vector $x_a$.

Hence, our problem is

$$\begin{aligned}&\text{Maximize } z = -M \sum_{i=1}^{m} x_{a_i}\\&\text{subject to the constraints}\\&\qquad Ax + Ix_a = b, \quad x,\, x_a \geq 0.\end{aligned}$$

Solving this LPP, we get the optimal solution. If the basis of the optimum simplex table contains all the variables of vector $x$, then the inverse of $A$ is directly obtained from the optimum simplex table. It consists of those column vectors in the last iteration of the simplex method which are present in the initial basis.

If the optimum (final) simplex table does not contain all the variables of vector $x$ in the basis, the simplex method is continued until all the variables of the vector are in the basis. In this situation, the solution may be either feasible or infeasible but optimum.

*Example 5* Solve the following LPP by the Big M method:

$$\begin{aligned}\text{Maximize} \quad & z = 2x_1 - 3x_2 - 2x_3\\ \text{subject to} \quad & 2x_1 + x_2 - 2x_3 \leq 3\\ & x_1 + x_2 - x_3 \geq 1 \quad x_1,\ x_2,\ x_3 \geq 0.\end{aligned}$$

**Solution:** The given problem is a maximization problem. Here all $b_i$'s are positive. Now, introducing a slack variable $x_4$ and a surplus variable $x_5$ to the first and second constraints respectively, we get the following converted equations:

$$\begin{aligned} 2x_1 + x_2 - 2x_3 + x_4 &= 3 \\ x_1 + x_2 - x_3 - x_5 &= 1, \end{aligned}$$

where $x_i \geq 0, \quad i = 1, 2, \ldots, 5.$

The coefficient matrix of this system of equations does not contain a unit matrix. To get a unit matrix, an artificial variable $x_6$ is added to the second equation; then the transformed equations are as follows:

$$\begin{aligned} 2x_1 + x_2 - 2x_3 + x_4 &= 3 \\ x_1 + x_2 - x_3 - x_5 + x_6 &= 1 \end{aligned}$$

where $x_i \geq 0, \quad i = 1, 2, \ldots, 6.$

Then the new objective function is given by

$z = 2x_1 - 3x_2 - 2x_3 + 0 \cdot x_4 + 0 \cdot x_5 - Mx_6$ [assign a very large negative price to the artificial variable].

| $c_j$ | | | 2 | −3 | −2 | 0 | 0 | $-M$ | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 0 | $y_4$ | $x_4 = 3$ | 2 | 1 | −2 | 1 | 0 | 0 | 3/2 = 3/2 |
| $-M$ | $y_6$ | $x_6 = 1$ | 1 | 1 | −1 | 0 | −1 | 1 | 1/1 = 1 |
| $z_B = -M$ | | | $-M - 2$ | $-M + 3$ | $M + 2$ | 0 | $M$ | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_1$ is the incoming vector, $y_6$ is the outgoing vector and $y_{21} = 1$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | 2 | −3 | −2 | 0 | 0 | $-M$ | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 0 | $y_4$ | $x_4 = 1$ | 0 | −1 | 0 | 1 | 2 | −2 | 1/2 |
| 2 | $y_1$ | $x_1 = 1$ | 1 | 1 | −1 | 0 | −1 | 1 | – |
| $z_B = 2$ | | | 0 | 5 | 0 | 0 | −2 | $2 + M$ | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_5$ is the incoming vector, $y_4$ is the outgoing vector and $y_{15} = 1$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | 2 | −3 | −2 | 0 | 0 | $-M$ | |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 0 | $y_5$ | $x_5$ = 1/2 | 0 | −1/2 | 0 | 1/2 | 1 | −1 | |
| 2 | $y_1$ | $x_1$ = 3/2 | 1 | 1/2 | −1 | 1/2 | 0 | 0 | |
| $z_B = 3$ | | | 0 | 4 | 0 | 1 | 0 | $M$ | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j \geq 0$. Hence, optimality has been reached, and the optimal solution is

$$x_1 = 3/2, \quad x_2 = 0, \quad x_3 = 0 \text{ with Max } z = 3.$$

*Example 6* Use Charne's Big M method to solve the following LPP:

$$\begin{array}{ll} \text{Maximize} & z = x_1 + 2x_2 \\ \text{subject to} & x_1 - 5x_2 \leq 10 \\ & 2x_1 - x_2 \geq 2 \\ & x_1 + x_2 = 0 \text{ and } x_1,\ x_2 \geq 0. \end{array}$$

**Solution:** The given problem is a maximization problem. Here all $b_i$'s are positive. Now, introducing a slack variable $x_3$ and a surplus variable $x_4$ to the first and second constraints respectively, we get the following converted equations:

$$\begin{aligned} x_1 - 5x_2 + x_3 &= 10 \\ 2x_1 - x_2 - x_4 &= 2 \\ x_1 + x_2 &= 0 \end{aligned}$$

where $x_i \geq 0, \quad i = 1,\ 2,\ \ldots,\ 4.$

The coefficient matrix of this system of equations does not contain a unit matrix. To get a unit matrix, two artificial variables $x_5$ and $x_6$ are added to second and third equations respectively, and the transformed equations are as follows:

$$\begin{aligned} x_1 - 5x_2 + x_3 &= 10 \\ 2x_1 - x_2 - x_4 + x_5 &= 2 \\ x_1 + x_2 + x_6 &= 0 \end{aligned}$$

where $x_i \geq 0, \quad i = 1,\ 2,\ \ldots,\ 6.$

Then the new objective function is given by

$z = x_1 + 2x_2 + 0x_3 + 0x_4 - Mx_5 - Mx_6$ [assign a very large negative price to the artificial variable].

Clearly, the variables $x_3$, $x_5$ and $x_6$ are basic. Considering the other variables as non-basic, we get $x_3 = 10$, $x_5 = 2$ and $x_6 = 0$ which is the initial basic feasible solution.

Now we construct the initial simplex table.

| $c_j$ | | | 1 | 2 | 0 | 0 | $-M$ | $-M$ | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 0 | $y_3$ | $x_3 = 10$ | 1 | $-5$ | 1 | 0 | 0 | 0 | $10/1 = 10$ |
| $-M$ | $y_5$ | $x_5 = 2$ | 2 | $-1$ | 0 | $-1$ | 1 | 0 | $2/2 = 1$ |
| $-M$ | $y_6$ | $x_6 = 0$ | 1 | 1 | 0 | 0 | 0 | 1 | $0/1 = 0$ |
| $z_B = -2M$ | | | $-3M - 2$ | $-2$ | 0 | $M$ | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_1$ is the incoming vector, $y_6$ is the outgoing vector and $y_{31} = 1$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | 1 | 2 | 0 | 0 | $-M$ | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_3$ | $x_3 = 10$ | 0 | $-6$ | 1 | 0 | 0 | |
| $-M$ | $y_5$ | $x_5 = 2$ | 0 | $-3$ | 0 | $-1$ | 1 | |
| 1 | $y_1$ | $x_1 = 0$ | 1 | 1 | 0 | 0 | 0 | |
| $z_B = -2M$ | | | 0 | $3M - 1$ | 0 | $M$ | 0 | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j \geq 0$ for all $j$; i.e. the optimality condition has been satisfied. But the artificial vector $y_5$ appears in the basis at a positive level, and the value of the corresponding artificial variable $x_5$ is 2. Hence, there is no feasible solution to the given LPP.

*Example 7* Solve the LPP

$$\begin{aligned} \text{Maximize} \quad & z = 3x_1 - x_2 \\ \text{subject to} \quad & -x_1 + x_2 \geq 2 \\ & 5x_1 - 2x_2 \geq 2 \text{ and } x_1,\ x_2 \geq 0. \end{aligned}$$

**Solution:**

The given problem is a maximization problem. Here all $b_i$'s are positive. Now, introducing surplus and artificial variables, the given problem can be written in the standard form as follows:

$$\begin{array}{ll} \text{Maximize} & z = 3x_1 - x_2 + 0 \cdot x_3 + 0 \cdot x_4 - Mx_5 - Mx_6 \\ \text{subject to} & -x_1 + x_2 - x_3 + \quad x_5 \quad = 2 \\ & 5x_1 - 2x_2 \quad - x_4 \quad + x_6 = 2 \\ & \text{and } x_i \geq 0, \quad i = 1, 2, \ldots, 6. \end{array}$$

Clearly, the variables $x_5$ and $x_6$ are basic. Considering the other variables as non-basic, we get $x_5 = 2$ and $x_6 = 2$, which is the initial basic feasible solution.

Now we construct the initial simplex table.

| $c_j$ | | | 3 | −1 | 0 | 0 | $-M$ | $-M$ | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| $-M$ | $y_5$ | $x_5 = 2$ | −1 | 1 | −1 | 0 | 1 | 0 | – |
| $-M$ | $y_6$ | $x_6 = 2$ | 5 | −2 | 0 | −1 | 0 | 1 | 2/5 = 2/5 |
| $z_B = -4M$ | | | $-4M - 3$ | $M + 1$ | $M$ | $M$ | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_1$ is the incoming vector, $y_6$ is the outgoing vector and $y_{21} = 5$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | | 3 | −1 | 0 | 0 | $-M$ | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $B$ | $x_B$ | $b$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| $-M$ | $y_5$ | $x_5$ | 12/ 5 | 0 | 3/5 | −1 | −1/5 | 1 | 12/5/3/ 5 = 4 |
| 3 | $y_1$ | $x_1$ | 2/5 | 1 | −2/5 | 0 | −1/5 | 0 | – |
| $z_B = -12M/5 + 6/5$ | | | | 0 | $-3/5M - 1/5$ | $M$ | $M/5 - 3/5$ | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_2$ is the incoming vector, $y_5$ is the outgoing vector and $y_{12} = 3/5$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | | 3 | −1 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $B$ | $x_B$ | $b$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| −1 | $y_2$ | $x_2$ | 4 | 0 | 1 | −5/3 | −1/3 | |
| 3 | $y_1$ | $x_1$ | 2 | 1 | 0 | −2/3 | −1/3 | |
| $z_B = 2$ | | | | 0 | 0 | −1/3 | −2/3 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. The value of $z_4 - c_4$ is mostly negative, with $y_{i4} < 0$, $i = 1, 2$. Hence, the outgoing vector cannot be selected because there is an unbounded solution to the given problem.

*Example 8* Solve the following LPP using the Big M method:

$$\begin{aligned}&\text{Maximize}\, z = 3x_1 - x_2\\&\text{subject to the constraints}\\&\quad 2x_1 + x_2 \geq 2, x_1 + 3x_2 \leq 3, x_2 \leq 4\, \text{and}\, x_1, x_2 \geq 0.\end{aligned}$$

**Solution:** The given problem is a maximization problem. Here all $b_i$'s are positive. Now introducing surplus variable $x_3$ and slack variables $x_4$ and $x_5$ in the given constraints, we get the following converted equations:

$$\begin{aligned}2x_1 + x_2 - x_3 &= 2\\x_1 + 3x_2 + x_4 &= 3\\x_2 + x_5 \quad &= 4\end{aligned}$$

where $x_i \geq 0, \quad i = 1,\ 2,\ \ldots,\ 5.$

The coefficient matrix of the preceding system of equations does not contain a unit matrix. To get a unit matrix, an artificial variable $x_6$ is added to the first equation, and the transformed equations are as follows:

$$\begin{aligned}&2x_1 + x_2 - x_3 + x_6 = 2\\&x_1 + 3x_2 + x_4 \qquad\;\; = 3\\&x_2 + x_5 \qquad\qquad\quad\; = 4\end{aligned}$$

where $x_i \geq 0, \quad i = 1,\ 2,\ \ldots,\ 6$ .

Then the new objective function is given by

$z = 3x_1 - x_2 + 0 \cdot x_3 + 0 \cdot x_4 + 0 \cdot x_5 - Mx_6$ [assign a very large negative price to the artificial variable].

Clearly, the variables $x_4,\ x_5$ and $x_6$ are basic. Considering the other variables as non-basic, we get $x_6 = 2,\ x_4 = 3$ and $x_5 = 4$, which is the initial basic feasible solution. Now we construct the initial simplex table.

| $c_j$ | | | 3 | $-1$ | 0 | 0 | 0 | $-M$ | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| $-M$ | $y_6$ | $x_6 = 2$ | 2 | 1 | $-1$ | 0 | 0 | 1 | 2/2 = 1 |
| 0 | $y_4$ | $x_4 = 3$ | 1 | 3 | 0 | 1 | 0 | 0 | 3/1 = 3 |
| 0 | $y_5$ | $x_5 = 4$ | 0 | 1 | 0 | 0 | 1 | 0 | – |
| $z_B = -2M$ | | | $-2M - 3$ | $-M + 1$ | $M$ | 0 | 0 | 0 | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j \ngeq 0$. Hence, the current BFS is not optimal. Here $y_1$ is the incoming vector, $y_{11} = 4$ is the key element and $y_6$ is the outgoing vector. Now we construct the next simplex table.

| $c_j$ | | | 3 | −1 | 0 | 0 | 0 | Mini ratio |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 3 | $y_1$ | $x_1 = 1$ | 1 | 1/2 | −1/2 | 0 | 0 | – |
| 0 | $y_4$ | $x_4 = 2$ | 0 | 5/2 | $1/2$ | 1 | 0 | 2/½ = 4 |
| 0 | $y_5$ | $x_5 = 4$ | 0 | 1 | 0 | 0 | 1 | – |
| $z_B = 3$ | | | 0 | 5/2 | −3/2 | 0 | 0 | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j \ngeq 0$. Hence, the current BFS is not optimal. Here $y_3$ is the incoming vector, $y_{23} = 1/2$ is the key element and $y_4$ is the outgoing vector. Now we construct the next simplex table.

| $c_j$ | | | 3 | −1 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 3 | $y_1$ | $x_1 = 3$ | 1 | 3 | 0 | 1 | 0 | |
| 0 | $y_3$ | $x_3 = 4$ | 0 | 5 | 1 | 2 | 0 | |
| 0 | $y_5$ | $x_5 = 4$ | 0 | 1 | 0 | 0 | 1 | |
| $z_B = 9$ | | | 0 | 10 | 0 | 3 | 0 | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j \geq 0$ for all $j$. Hence, the current BFS is optimal, and the optimal solution is given by

$$x_1 = 3, x_2 = 0 \,\text{with}\, \text{Max}\, z = 9.$$

*Example 9* Using the simplex method, solve the following system of linear equations:

$$x_1 + x_2 = 1, \;\; 2x_1 + x_2 = 3.$$

**Solution:** Since the non-negativity restrictions of $x_1$ and $x_2$ are not given, these variables are unrestricted in sign. Now to introduce the non-negativity restrictions, let us assume that

$$x_1 = x_1' - x_1'' \,\text{and}\, x_2 = x_2' - x_2'',$$

where $x_1', x_1'', x_2', x_2'' \geq 0$.

To solve the given system of equations, we introduce two artificial variables $x_3 \geq 0$ and $x_4 \geq 0$ in the first and second equations respectively and a dummy objective function

$z = -Mx_3 - Mx_4$ (assigning a large negative price $-M$ corresponding to each artificial variable)

which is to be maximized subject to the constraints as follows:

$$x_1' - x_1'' + x_2' - x_2'' + x_3 \quad = 1$$
$$2x_1' - 2x_1'' + x_2' - x_2'' + x_4 = 3$$

and $x_1', x_1'', x_2', x_2'', x_3, x_4 \geq 0.$

Let us take $x_3$ and $x_4$ as the basic variables. Considering other variables as non-basic variables, we get $x_3 = 1,\ x_4 = 3$, which is the initial BFS.

Now we construct the initial simplex table.

| $c_j$ | | | 0 | 0 | 0 | 0 | $-M$ | $-M$ | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1'$ | $y_1''$ | $y_2'$ | $y_2''$ | $y_3$ | $y_4$ | |
| $-M$ | $y_3$ | $x_3 = 1$ | 1 | $-1$ | 1 | $-1$ | 1 | 0 | 1 |
| $-M$ | $y_4$ | $x_4 = 3$ | 2 | $-2$ | 1 | $-1$ | 0 | 1 | 3/2 |
| $z_B = -4M$ | | | $-3M$ | $-3M$ | $-2M$ | $2M$ | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_1'$ is the incoming vector, $y_3$ is the outgoing vector and $y_{11}' = 1$ is the key element. Now we construct the next simplex table.

| $c_j$ | | | 0 | 0 | 0 | 0 | $-M$ | Mini ratio |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1'$ | $y_1''$ | $y_2'$ | $y_2''$ | $y_4$ | |
| 0 | $y_1'$ | $x_1' = 1$ | 1 | $-1$ | 1 | $-1$ | 0 | – |
| $-M$ | $y_4$ | $x_4 = 1$ | 0 | 0 | $-1$ | 1 | 1 | 1 |
| $z = -M$ | | | 0 | 0 | $M$ | $-M$ | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_2''$ is the incoming vector, $y_4$ is the outgoing vector and $y_{22}'' = 1$ is the key element. Now we construct the next simplex table.

| $c_j$ | | | 0 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1'$ | $y_1''$ | $y_2'$ | $y_2''$ | |
| 0 | $y_1'$ | $x_1' = 2$ | 1 | −1 | 0 | 0 | |
| 0 | $y_2''$ | $x_2'' = 1$ | 0 | 0 | −1 | 1 | |
| $z_B = 0$ | | | 0 | 0 | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j$ are non-negative, an optimum solution has been attained. Thus, the solution of the LPP is

$$\begin{aligned} & x_1' = 2,\ x_1'' = 0,\ x_2' = 0,\ x_2'' = 1, \\ \text{i.e.} \quad & x_1 = x_1' - x_1'' = 2 - 0 \qquad = 2 \\ \text{and} \quad & x_2 = x_2' - x_2'' = 0 - 1 \qquad = -1, \end{aligned}$$

which is the solution of the given system of equations.

*Remark* Note that the method is applied in solving an LPP with one or more unrestricted variables with the given objective function.

*Example 10* Using the simplex method, find the inverse of the following matrix:

$$A = \begin{pmatrix} 4 & 2 \\ 1 & 5 \end{pmatrix}.$$

**Solution:** Considering the given matrix as a coefficient matrix, let us take the following system of linear equations:

$$\begin{pmatrix} 4 & 2 \\ 1 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix}, \quad x_1, x_2 \geq 0,$$

where $(6,\ 4)^T$ is a dummy column vector,

$$\begin{aligned} \text{i.e.} \quad & 4x_1 + 2x_2 = 6 \\ & x_1 + 5x_2 \ = 4 \\ \text{and} \quad & x_1,\ x_2 \geq 0. \end{aligned}$$

To solve the system of equations, we introduce two artificial variables $x_3 \geq 0$ and $x_4 \geq 0$ in the first and second equations respectively and a dummy objective function

$z = -Mx_3 - Mx_4$ (assigning a large negative price $-M$ corresponding to each artificial variable) which is to be maximized subject to the constraints as follows:

$$4x_1 + 2x_2 + x_3 = 6$$
$$x_1 + 5x_2 + x_4 = 4$$
$$\text{and } x_1, x_2, x_3, x_4 \geq 0.$$

Let us take $x_3$ and $x_4$ as the basic variables. Considering other variables as non-basic variables, we get $x_3 = 6,\ x_4 = 4$, which is the initial BFS.

Now we construct the initial simplex table.

| $c_j$ | | | 0 | 0 | $-M$ | $-M$ | Mini ratio |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| $-M$ | $y_3$ | $x_3 = 6$ | 4 | 2 | 1 | 0 | 2 |
| $-M$ | $y_4$ | $x_4 = 4$ | 1 | 5 | 0 | 1 | 6/5 |
| $z_B = -10M$ | | | $-5M$ | $-7M$ | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_2$ is the incoming vector, $y_4$ is the outgoing vector and $y_{22} = 5$ is the key element. Now we construct the next simplex table.

| $c_j$ | | | 0 | 0 | $-M$ | $-M$ | Mini ratio |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| $-M$ | $y_3$ | $x_3 = 22/5$ | 18/5 | 0 | 1 | −2/5 | 11/9 |
| 0 | $y_2$ | $x_2 = 4/5$ | 1/5 | 1 | 0 | 1/5 | 4 |
| $z_B = -22/5M$ | | | $-18/5M$ | 0 | 0 | $7/5M$ | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_1$ is the incoming vector, $y_3$ is the outgoing vector and $y_{11} = 18/5$ is the key element. Now we construct the next simplex table.

| $c_j$ | | | 0 | 0 | $-M$ | $-M$ | |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| 0 | $y_1$ | $x_1 = 11/9$ | 1 | 0 | 5/18 | $-1/9$ | |
| 0 | $y_2$ | $x_2 = 5/9$ | 0 | 1 | $-1/18$ | 2/9 | |
| $z_B = 0$ | | | 0 | 0 | $M$ | $M$ | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j$ are non-negative, an optimum solution has been attained.

From the preceding simplex table, it is seen that

$$y_3 = \begin{pmatrix} \frac{5}{18} & -\frac{1}{18} \end{pmatrix}^T \text{ and } y_4 = \begin{pmatrix} -\frac{1}{9} & \frac{2}{9} \end{pmatrix}^T.$$

Since the initial basis consists of the column vectors $y_3$ and $y_4$, and the column vectors $y_1$, $y_2$ corresponding to columns of matrix $A$ reduce to unit vectors, then the inverse of matrix $A$ is given by

$$A^{-1} = \begin{pmatrix} 5/18 & -1/9 \\ -1/18 & 2/9 \end{pmatrix} = \frac{1}{18}\begin{pmatrix} 5 & -2 \\ -1 & 4 \end{pmatrix}.$$

### *3.12.4 Two-Phase Method*

In an earlier section, we discussed the Big M method for solving a particular type of LPP. To solve the same problem, Dantzig, Orden and Wolfe developed an alternative method known as the two-phase method. In this method, the problems are solved in two different phases, Phase I and Phase II.

**Phase I**

The objective of Phase I is either to remove all the artificial variables from the basis or to set some of the said variables (which are present in the basis) at a zero level. For this purpose, an auxiliary LPP is constructed with a new objective function

$$z' = -\sum_i x_{a_i} \; [x_{a_i} \text{ is an artificial variable}]$$

which is to be maximized.

From the mathematical point of view, all artificial variables must be at a zero level at the optimal stage. Hence, the maximum value of $z'$ should be equal to zero. However, this will not be true for all problems. Thus, to find the maximum value of the new objective function $z'$, the following auxiliary LPP is to be solved first:

$$\text{Maximize } z' = -\sum_i x_{a_i}$$
$$\text{subject to all the constraints.}$$

In solving this LPP, when optimality conditions are satisfied in an iteration, the following three cases may arise:

Case I: Max $z' = 0$ and all artificial vectors are not present in the basis.
Case II: Max $z' = 0$, but some of the artificial vectors are present in the basis at a zero level. In this case, some constraints may be redundant.

Case III: Max $z' < 0$. This means that one or more artificial variables are present in the basis at a positive level. In this case, the original LPP has no feasible solution.

**Phase II**

If Phase I is terminated with either Case I or Case II, then we have to start Phase II to find the optimal solution of the original problem. In this phase, the first simplex table is almost identical with the last table of Phase I except that the rows for $c_j$ and $z_j - c_j$ as the values of $c_j$ of the original LPP will be used in the table.

In Case II, Max $z' = 0$ and one or more artificial vectors appear in the basis at zero level. This means that one or more artificial variables are basic, but their values are zero.

In this case, we must take care that these artificial variables never become positive at any iteration of Phase II.

*Example 11* Use the two-phase simplex method to solve the following problem:

$$\begin{aligned}
&\text{Maximize } z = 3x_1 - x_2 \\
&\text{subject to the constraints} \\
&\qquad 2x_1 + x_2 \geq 2 \\
&\qquad x_1 + 3x_2 \leq 2 \\
&\qquad x_1 \leq 4 \\
&\qquad x_1,\ x_2 \geq 0.
\end{aligned}$$

**Solution:**

This is a maximization problem. All $b_i$'s are positive. Now introducing the surplus variable $x_3$ and slack variables $x_4, x_5$, the given problem can be rewritten as follows:

$$\begin{aligned}
&\text{Maximize } z = 3x_1 - x_2 + 0 \cdot x_3 + 0 \cdot x_4 + 0 \cdot x_5 \\
&\text{subject to the constraints} \\
&\qquad 2x_1 + x_2 - x_3 = 2 \\
&\qquad x_1 + 3x_2 + x_4 = 2 \\
&\qquad x_1 + x_5 = 4 \\
&\text{and } x_j \geq 0, j = 1, 2, \ldots, 5.
\end{aligned}$$

The coefficient matrix of the above constraints does not contain a unit matrix. To get a unit matrix, an artificial variable $x_6$ is added to the first equation, and the transformed equations are as follows:

$$2x_1 + x_2 - x_3 + x_6 = 2$$
$$x_1 + 3x_2 + x_4 = 2$$
$$x_1 + x_5 = 4$$

where $x_j \geq 0, \quad j = 1, 2, \ldots, 6.$

**Phase I:** Assigning a profit coefficient $-1$ to the artificial variable and 0 to all other variables, the objective function of the auxiliary LPP to be maximized is given by

$$z' = 0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 + 0 \cdot x_5 - x_6$$

subject to the preceding system of equations.

Let us take $x_4, x_5$ and $x_6$ as the basic variables. Considering other variables as non-basic variables, we get $x_6 = 2, x_4 = 2, x_5 = 4$, which is the initial BFS.

Now we construct the initial simplex table.

| $c_j$ | | | 0 | 0 | 0 | 0 | 0 | $-1$ | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| $-1$ | $y_6$ | $x_6 = 2$ | 2 | 1 | $-1$ | 0 | 0 | 1 | 2/2 = 2 |
| 0 | $y_4$ | $x_4 = 2$ | 1 | 3 | 0 | 1 | 0 | 0 | 2/1 = 2 |
| 0 | $y_5$ | $x_5 = 4$ | 1 | 0 | 0 | 0 | 1 | 0 | 4/1 = 4 |
| $z' = -2$ | | | $-2$ | $-1$ | 1 | 0 | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_1$ is the incoming vector, $y_6$ is the outgoing vector and $y_{11} = 2$ is the key element. Now we construct the next simplex table.

| $c_j$ | | | 0 | 0 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_1$ | $x_1 = 1$ | 1 | 1/2 | $-1/2$ | 0 | 0 | |
| 0 | $y_4$ | $x_4 = 1$ | 0 | 5/2 | 1/2 | 1 | 0 | |
| 0 | $y_5$ | $x_5 = 3$ | 0 | $-1/2$ | 1/2 | 0 | 1 | |
| $z'_B = 0$ | | | 0 | 0 | 0 | 0 | 0 | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j$ are non-negative and Max $z' = 0$. Again no artificial variable appears in the final table of Phase I. We then pass on to Phase II where the objective function of the original problem to be maximized is

$$z = 3x_1 - x_2 + 0 \cdot x_3 + 0 \cdot x_4 + 0 \cdot x_5.$$

Now we construct the initial simplex table of Phase II.

| $c_j$ | | | 3 | −1 | 0 | 0 | 0 | Mini ratio |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 3 | $y_1$ | $x_1 = 1$ | 1 | 1/2 | −1/2 | 0 | 0 | – |
| 0 | $y_4$ | $x_4 = 1$ | 0 | 5/2 | 1/2 | 1 | 0 | 2 |
| 0 | $y_5$ | $x_5 = 3$ | 0 | −1/2 | 1/2 | 0 | 1 | 6 |
| $z_B = 3$ | | | 0 | 5/2 | −3/2 | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_3$ is the incoming vector, $y_4$ is the outgoing vector and $y_{32} = 1/2$ is the key element. Now we construct the next simplex table.

| $c_j$ | | | 3 | −1 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 3 | $y_1$ | $x_1 = 2$ | 1 | 3 | 0 | 1 | 0 | |
| 0 | $y_3$ | $x_3 = 2$ | 0 | 5 | 1 | 2 | 0 | |
| 0 | $y_5$ | $x_5 = 2$ | 0 | −3 | 0 | −1 | 1 | |
| $z_B = 6$ | | | 0 | 10 | 0 | 3 | 0 | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j$ are non-negative. Hence, an optimum solution has been attained, and this solution is given by $x_1 = 2$, $x_2 = 0$ and Max z = 6.

*Example 12* Use the two-phase simplex method to solve the following problem:

$$\text{Maximize } z = 3x_1 + 2x_2$$
$$\text{subject to the constraints}$$
$$2x_1 + x_2 \leq 2$$
$$3x_1 + 4x_2 \geq 12$$
$$\text{and } x_1,\ x_2 \geq 0.$$

**Solution:**

This is a maximization problem. All $b_i$'s are positive. Now introducing the slack variable $x_3$ and surplus variable $x_4$, the given problem can be rewritten as follows:

$$\begin{aligned}&\text{Maximize } z = 3x_1 + 2x_2 + 0\cdot x_3 + 0\cdot x_4\\&\text{subject to}\\&\quad 2x_1 + x_2 + x_3 \quad = 2\\&\quad 3x_1 + 4x_2 - x_4 = 12\\&\quad \text{and } x_1, x_2, x_3, x_4 \geq 0.\end{aligned}$$

The coefficient matrix of these constraints does not contain a unit matrix. To get a unit matrix, an artificial variable $x_5$ is added to the first equation; the transformed equations are as follows:

$$2x_1 + x_2 + x_3 = 2$$
$$3x_1 + 4x_2 - x_4 + x_5 = 12$$

where $x_j \geq 0, \quad j = 1, 2, 3, 4, 5.$

**Phase I:** Assigning a profit coefficient −1 to the artificial variable and 0 to all other variables, the objective function of the auxiliary LPP to be maximized is given by

$$z' = 0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 - x_5$$

subject to the preceding system of equations.

| $c_j$ | | | 0 | 0 | 0 | 0 | −1 | Mini ratio |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_3$ | $x_3 = 2$ | 2 | 1 | 1 | 0 | 0 | 2/1 = 2 |
| −1 | $y_5$ | $x_5 = 12$ | 3 | 4 | 0 | −1 | 1 | 12/4 = 3 |
| $z'_B = -12$ | | | −3 | −4 | 0 | 1 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_2$ is the incoming vector, $y_3$ is the outgoing vector and $y_{21} = 1$ is the key element. Now we construct the next simplex table.

| $c_j$ | | | 0 | 0 | 0 | 0 | −1 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_2$ | $x_2 = 2$ | 2 | 1 | 1 | 0 | 0 | |
| −1 | $y_5$ | $x_5 = 4$ | −5 | 0 | −4 | −1 | 1 | |
| $z'_B = -4$ | | | 5 | 0 | 4 | 1 | 0 | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j \geq 0$. Hence, an optimum basic feasible solution to the auxiliary LPP has been attained. Again, Max $z' < 0$ and the artificial variable $x_5$ is basic with a positive value. Hence, the original LPP does not possess any feasible solution.

*Example 13* Use the two-phase simplex method to solve the following problem:

$$\begin{aligned}&\text{Maximize } z = 5x_1 - 4x_2 + 3x_3\\&\text{subject to the constraints}\\&\quad 2x_1 + x_2 - 6x_3 = 20\\&\quad 6x_1 + 5x_2 + 10x_3 \le 76\\&\quad 8x_1 - 3x_2 + 6x_3 \le 50\\&\quad \text{and } x_1,\ x_2,\ x_3 \ge 0.\end{aligned}$$

**Solution:**

This is a maximization problem. All $b_i$'s are positive. Now introducing the slack variables $x_3$ and $x_4$, the given problem can be rewritten as follows:

$$\begin{aligned}&\text{Maximize } z = 5x_1 - 4x_2 + 3x_3 + 0\cdot x_4 + 0\cdot x_5\\&\text{subject to the constraints}\\&\quad 2x_1 + x_2 - 6x_3 = 20\\&\quad 6x_1 + 5x_2 + 10x_3 + x_4 = 76\\&\quad 8x_1 - 3x_2 + 6x_3 + x_5 = 50\\&\quad \text{and } x_j \ge 0,\quad j = 1, 2, 3, 4, 5.\end{aligned}$$

The coefficient matrix of these constraints does not contain a unit matrix. To get a unit matrix, an artificial variable $x_6$ is added to the first equation, and the transformed equations are as follows:

$$\begin{aligned}2x_1 + x_2 - 6x_3 + x_6 &= 20\\6x_1 + 5x_2 + 10x_3 + x_4 &= 76\\8x_1 - 3x_2 + 6x_3 + x_5 &= 50,\end{aligned}$$

where $x_j \ge 0,\quad j = 1, 2, \ldots, 6$.

**Phase I:** Assigning a profit coefficient $-1$ to the artificial variable and 0 to all other variables, the objective function of the auxiliary LPP to be maximized is given by

$$z' = 0\cdot x_1 + 0\cdot x_2 + 0\cdot x_3 + 0\cdot x_4 + 0\cdot x_5 - x_6$$

subject to the preceding system of equations.

| $c_j$ | | | 0 | 0 | 0 | 0 | 0 | −1 | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| −1 | $y_6$ | $x_6 = 20$ | 2 | 1 | −6 | 0 | 0 | 1 | 10 |
| 0 | $y_4$ | $x_4 = 76$ | 6 | 5 | 10 | 1 | 0 | 0 | 38/3 |

(continued)

(continued)

| $c_j$ | | | 0 | 0 | 0 | 0 | 0 | −1 | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 0 | $y_5$ | $x_5 = 50$ | 8 | −3 | 6 | 0 | 1 | 0 | 25/4 |
| $z'_B = -20$ | | | −2 | −1 | 6 | 0 | 0 | 0 | $z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_1$ is the incoming vector, $y_5$ is the outgoing vector and $y_{31} = 8$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | 0 | 0 | 0 | 0 | 0 | −1 | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| −1 | $y_6$ | $x_6 = 15/2$ | 0 | 7/4 | −15/2 | 0 | −1/4 | 1 | $\frac{30}{7}$ |
| 0 | $y_4$ | $x_4 = 77/2$ | 0 | 29/4 | 11/2 | 1 | −3/4 | 0 | $\frac{154}{29}$ |
| 0 | $y_1$ | $x_1 = 25/4$ | 1 | −3/8 | 3/4 | 0 | 1/8 | 0 | – |
| $z'_B = -15/2$ | | | 0 | −7/4 | 15/2 | 0 | 1/4 | 0 | $z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_2$ is the incoming vector, $y_6$ is the outgoing vector and $y_{12} = 7/4$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | 0 | 0 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_2$ | $x_2 = 30/7$ | 0 | 1 | −30/7 | 0 | −1/7 | |
| 0 | $y_4$ | $x_4 = 52/7$ | 0 | 0 | 256/7 | 1 | 2/7 | |
| 0 | $y_1$ | $x_1 = 55/7$ | 1 | 0 | −6/7 | 0 | 1/14 | |
| $z'_B = 0$ | | | 0 | 0 | 0 | 0 | 0 | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j \geq 0$. Hence, an optimum solution to the auxiliary LPP has been attained. Furthermore, from this table, it is observed that no artificial vector appears in the basis. We then pass on to Phase II, where the objective function of the original problem to be maximized is

$$z = 5x_1 - 4x_2 + 3x_3 + 0 \cdot x_4 + 0 \cdot x_5.$$

**Phase II:**

Now we construct the initial simplex table of Phase II.

| $c_j$ | | | 5 | −4 | 3 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| −4 | $y_2$ | $x_2 = 30/7$ | 0 | 1 | −30/7 | 0 | −1/7 | |
| 0 | $y_4$ | $x_4 = 52/7$ | 0 | 0 | 256/7 | 1 | 2/7 | |
| 0 | $y_1$ | $x_1 = 55/7$ | 1 | 0 | −6/7 | 0 | 1/14 | |
| $z_B = 155/7$ | | | 0 | 0 | 69/7 | 0 | 13/14 | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j \geq 0$. Hence, an optimum basic feasible solution to the given LPP has been attained, and the optimal solution is given by

$$x_1 = 55/7, x_2 = 30/7, x_3 = 0 \,\text{with maximum}\, z = 155/7.$$

*Example 14* Using the simplex method, solve the following system of linear simultaneous equations:

$$\begin{aligned} x_1 + x_2 &= 1 \\ 2x_1 + x_2 &= 3. \end{aligned}$$

**Solution:** Since the non-negativity restrictions of $x_1$ and $x_2$ are not given, these variables are unrestricted in sign. Now to introduce the non-negativity restrictions, let us assume that $x_1 = x_1' - x_1''$ and $x_2 = x_2' - x_2''$, where $x_1', x_1'', x_2', x_2'' \geq 0$.

Now we introduce the artificial variables $x_3 \geq 0$, $x_4 \geq 0$ by assigning profit −1 for each. Then the problem of solving the given system of equations reduces to the following LPP:

$$\begin{aligned} &\text{Maximize}\, z = -x_3 - x_4 \\ &\text{subject to the constraints} \\ &\quad x_1' - x_1'' + x_2' - x_2'' + x_3 = 1 \\ &\quad 2x_1' - 2x_1'' + x_2' - x'' + x_4 = 3 \\ &\quad \text{and}\, x_1', x_1'', x_2', x_2'', x_3, x_4 \geq 0. \end{aligned}$$

| $c_j$ | | | 0 | 0 | 0 | 0 | −1 | −1 | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1'$ | $y_1''$ | $y_2'$ | $y_2''$ | $y_3$ | $y_4$ | |
| −1 | $y_3$ | $x_3 = 1$ | 1 | −1 | 1 | −1 | 1 | 0 | 1 |
| −1 | $y_4$ | $x_4 = 3$ | 2 | −2 | 1 | −1 | 0 | 1 | 3/2 |
| $z_B = -4$ | | | −3 | 3 | −2 | 2 | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_1'$ is the incoming vector, $y_3$ is the outgoing vector and $y_{11}' = 1$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | 0 | 0 | 0 | 0 | −1 | Mini ratio |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1'$ | $y_1''$ | $y_2'$ | $y_2''$ | $y_4$ | |
| 0 | $y_1'$ | $x_1' = 1$ | 1 | −1 | 1 | −1 | 0 | – |
| −1 | $y_4$ | $x_4 = 1$ | 0 | 0 | −1 | 1 | 1 | 1 |
| $z_B = -1$ | | | 0 | 0 | 1 | −1 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \not\geq 0$, the current BFS is not optimal. Here $y_2''$ is the incoming vector, $y_4$ is the outgoing vector and $y_{22}'' = 1$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | 0 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1'$ | $y_1''$ | $y_2'$ | $y_2''$ | |
| 0 | $y_1'$ | $x_1'' = 2$ | 1 | −1 | 0 | 0 | |
| 0 | $y_2''$ | $x_2'' = 1$ | 0 | 0 | −1 | 1 | |
| $z_B = 0$ | | | 0 | 0 | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j$ are non-negative, an optimum solution has been attained. Thus, the solution of the LPP is

$$\begin{aligned} x_1' &= 2, x_2'' = 1, x_1'' = 0, x_2' = 0, \\ \text{i.e.} x_1 &= x_1' - x_1'' = 2 - 0 \quad &= 2 \\ x_2 &= x_2' - x'' = 0 - 1 \quad &= -1, \end{aligned}$$

which is the solution of the given system of equations.

*Remark* Note that the method is applied in solving an LPP with one or more unrestricted variables with the given objective function.

*Example 15* Using the simplex method, find the inverse of the following matrix:

$$A = \begin{pmatrix} 3 & 2 \\ 4 & -1 \end{pmatrix}.$$

**Solution:** Taking the given matrix as a coefficient matrix, let us consider the following system of linear equations:

$$\begin{pmatrix} 3 & 2 \\ 4 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}, \quad x_1, x_2 \geq 0,$$

where $(4, \ 6)^T$ is a dummy column vector,

$$\begin{aligned} \text{i.e.}\quad & 3x_1 + 2x_2 = 4 \\ & 4x_1 - x_2 \; = 6 \\ \text{and}\quad & x_1, x_2 \geq 0. \end{aligned}$$

To solve the system of equations, we introduce two artificial variables $x_3 \geq 0$ and $x_4 \geq 0$ in the first and second equations respectively and a dummy objective function

$$z = -x_3 - x_4$$

which is to be maximized subject to the constraints

$$\begin{aligned} 3x_1 + 2x_2 + x_3 &= 4 \\ 4x_1 - x_2 + x_4 &= 6 \\ \text{and}\, x_1,\; x_2,\; x_3,\; x_4 &\geq 0. \end{aligned}$$

| $c_j$ | | | 0 | 0 | −1 | −1 | Mini ratio |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| −1 | $y_3$ | $x_3 = 4$ | 3 | 2 | 1 | 0 | 4/3 |
| −1 | $y_4$ | $x_4 = 6$ | 4 | −1 | 0 | 1 | 3/2 |
| $z_B = -10$ | | | −7 | −1 | 0 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ngeq 0$, the current BFS is not optimal. Here $y_1$ is the incoming vector, $y_3$ is the outgoing vector and $y_{11} = 3$ is the key element.

Now we construct the next simplex table.

| $c_j$ | | | 0 | 0 | −1 | −1 | |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| 0 | $y_1$ | $x_1 = 4/3$ | 1 | 2/3 | 1/3 | 0 | |
| −1 | $y_4$ | $x_4 = 2/3$ | 0 | −11/3 | −4/3 | 1 | |
| $z_B = -2/3$ | | | 0 | 11/3 | 7/3 | 0 | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j$ are non-negative, an optimal solution has been reached. But to find the inverse of the matrix $A$, we have to convert the vector $y_2$ as a unit vector. For this purpose, we consider $y_2$ as the incoming vector and $y_4$ as the outgoing vector and $y_{22} = -11/3$ as the key element. Now we construct the next simplex table.

| $c_j$ | | | 0 | 0 | −1 | −1 | |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| 0 | $y_1$ | $x_1 = 16/11$ | 1 | 0 | 1/11 | 2/11 | |
| 0 | $y_2$ | $x_2 = -2/11$ | 0 | 1 | 4/11 | −3/11 | |
| $z_B = 0$ | | | 0 | 0 | 1 | 1 | $\leftarrow z_j - c_j$ |

From this simplex table, it is seen that $y_3 = \left(\frac{1}{11}, \frac{4}{11}\right)^T$ and $y_4 = \left(\frac{2}{11}, \frac{-3}{11}\right)^T$, since the initial basis consists of the column vectors $y_3$ and $y_4$ and the column vectors $y_1, y_2$ corresponding to the columns of matrix $A$; therefore, the inverse of matrix $A$ is given by

$$A^{-1} = \begin{pmatrix} \frac{1}{11} & \frac{2}{11} \\ \frac{4}{11} & \frac{-3}{11} \end{pmatrix} = \frac{1}{11}\begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix}.$$

## 3.13 Exercises

1. Solve the following LPPs:

(i) Maximize $z = 10x_1 + x_2 + 2x_3$
subject to $x_1 + x_2 - 2x_3 \le 10$
$4x_1 + x_2 + x_3 \le 20$ and $x_1,\ x_2 \ge 0.$

(ii) Minimize $z = 3x_1 - 2x_2$
subject to $4x_1 + x_2 \le 8$
$2x_1 + 4x_2 \le 20$ and $x_1,\ x_2 \ge 0.$

(iii) Maximize $z = -2x_1 + 5x_2$
subject to $4x_1 - 5x_2 \ge -20$
$x_1 + x_2 \ge 10$
$x_2 \ge 2$ and $x_1,\ x_2 \ge 0.$

(iv) Maximize $z = 5x_1 + 11x_2$
subject to $2x_1 + x_2 \le 4$
$3x_1 + 4x_2 \ge 24$
$2x_1 - 3x_2 \ge 6$ and $x_1,\ x_2 \ge 0.$

(v) Maximize $z = 2x_1 + x_2 + 3x_3$
subject to $x_1 + x_2 + 2x_3 \le 5$
$2x_1 + 3x_2 + 4x_3 = 12$ and $x_1,\ x_2,\ x_3 \ge 0.$

(vi) Maximize $z = 4x_1 + x_2$
subject to $3x_1 + x_2 = 3$
$4x_1 + 3x_2 \ge 6$
$x_1 + 2x_2 \le 3$ and $x_1,\ x_2 \ge 0.$

(vii) Maximize $z = 2x_1 + x_2 - x_3 + 4x_4$
subject to $3x_1 - 5x_2 + x_3 - 2x_4 = 7$
$6x_1 - 10x_2 - x_3 + 5x_4 = 11$ and $x_i \ge 0, i = 1, 2, \ldots, 4.$

(viii) Minimize $z = 2x_1 + x_2$
subject to $3x_1 + x_2 \ge 3$
$4x_1 + 3x_2 \ge 6$
$x_1 + 2x_2 \ge 2$ and $x_1,\ x_2 \ge 0.$

(ix) Minimize $z = 4x_1 + 2x_2$
subject to $3x_1 + x_2 \ge 27$
$x_1 + x_2 \ge 21$
$x_1 + 2x_2 \ge 30$ and $x_1,\ x_2 \ge 0.$

2. Solve the following LPPs:

(i) Maximize $z = x_1 + 5x_2$
subject to $3x_1 + 4x_2 \leq 6$
$x_1 + 3x_2 \geq 3$ and $x_1, x_2 \geq 0$.

(ii) Maximize $z = 2x_1 + x_2 + 3x_3$
subject to $x_1 + x_2 + 2x_3 \leq 5$
$2x_1 + 3x_2 + 4x_3 = 12$ and $x_1, x_2, x_3 \geq 0$.

(iii) Maximize $z = 4x_1 + x_2$
subject to $3x_1 + x_2 = 3$
$4x_1 + 3x_2 \geq 6$
$x_1 + 2x_2 \leq 3$ and $x_1, x_2 \geq 0$.

(iv) Maximize $z = 2x_1 + x_2$
subject to $4x_1 + 3x_2 \leq 12$
$4x_1 + x_2 \leq 8$
$4x_1 - x_2 \leq 8$ and $x_1, x_2 \geq 0$.

(v) Maximize $z = 2x_1 + x_2 + 10x_3$
subject to $x_1 - 2x_3 = 0$
$x_2 + x_3 = 1$ and $x_1, x_2, x_3 \geq 0$.

(vi) Maximize $z = 5x_1 - 2x_2 + 3x_3$
subject to $2x_1 + 2x_2 - x_3 \geq 2$
$3x_1 - 4x_2 \leq 3$
$x_2 + 3x_3 \leq 5$ and $x_1, x_2, x_3 \geq 0$.

(vii) Minimize $z = 4x_1 + 8x_2 + 3x_3$
subject to $x_1 + x_2 \geq 2$
$2x_1 + x_3 \geq 5$ and $x_1, x_2, x_3 \geq 0$.

(viii) Maximize $z = 5x_1 + 3x_2$
subject to $x_1 + x_2 \leq 2$
$5x_1 + 2x_2 \leq 10$
$5x_1 + 2x_2 \leq 10$ and $x_1, x_2 \geq 0$.

3. Solve the following LPPs using the two-phase simplex method:

(i) Minimize $z = x_1 + x_2$
subject to $2x_1 + x_2 \geq 4$
$x_1 + 7x_2 \geq 7$ and $x_1, x_2 \geq 0$.

(ii) Maximize $z = 3x_1 + 2x_2$
subject to $2x_1 + x_2 \leq 2$
$3x_1 + 4x_2 \geq 12$ and $x_1, x_2 \geq 0$.

(iii) Maximize $z = 5x_1 + 8x_2$
subject to $3x_1 + 2x_2 \geq 3$
$x_1 + 4x_2 \geq 4$
$x_1 + x_2 \leq 5$ and $x_1, x_2 \geq 0$.

(iv) Minimize $z = 4x_1 + x_2$
subject to $x_1 + 2x_2 \leq 3$
$4x_1 + 3x_2 \geq 6$
$3x_1 + x_2 = 3$ and $x_1, x_2 \geq 0$.

(v) Minimize $z = 40x_1 + 24x_2$
subject to $2x_1 + 25x_2 \geq 240$
$8x_1 + 5x_2 \geq 720$ and $x_1, x_2 \geq 0$.

4. Solve the following systems of linear simultaneous equations using the simplex method:

(i) $x_1 + x_2 = 1$
$2x_1 + x_2 = 4$.

(ii) $3x_1 + x_2 = 7$
$x_1 + x_2 = 3$.

(iii) $7x_1 - x_2 = 1$
$11x_1 + 2x_2 = 23$.

(iv) $x_1 + 2x_2 + 3x_3 = 14$
$2x_1 - x_2 + 5x_3 = 15$
$-3x_1 + 2x_2 + 4x_3 = 13$.

5. Using the simplex method, find the inverse of the following matrices:

(i) $\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ (ii) $\begin{pmatrix} 3 & 2 \\ 4 & -1 \end{pmatrix}$ (iii) $\begin{pmatrix} 1 & 2 & 3 \\ 3 & -2 & 1 \\ 4 & 2 & 1 \end{pmatrix}$ (iv) $\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}$

(v) $\begin{pmatrix} 3 & 4 \\ -1 & 2 \end{pmatrix}$

# Chapter 4
# Revised Simplex Method

## 4.1 Objective

The objective of this chapter is to discuss:

- The revised simplex methods and their computational procedures
- The relevant information required at each iteration of the revised simplex method
- The use of the revised simplex in comparison to the usual simplex method.

## 4.2 Introduction

The revised simplex method is a modified form of the ordinary simplex method. It is also an efficient method for solving linear programming problems (LPPs). It is efficient in the sense that we need not recompute all the values of $y_j,\ z_j - c_j$, $x_B$ at each iteration. Here, the word 'revised' refers to the procedure of changing or updating the ordinary simplex method.

This method is economical on the computer, as it computes and stores only the relevant information required for testing the optimality condition and for updating the current solution. The basic information required at each iteration in the ordinary simplex method includes:

(i) Net evaluations corresponding to the non-basic variables
(ii) The column vector corresponding to the most negative net evaluation
(iii) The values of the current basic variables
(iv) Selection of an outgoing vector from the basis.

This information can be computed directly from its definitions provided the inverse of the basis matrix is known. To get the inverse of the basis matrix at any

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_4

iteration, we need to compute the same from the previous inverse matrix. For this purpose, the relevant information we must know at each iteration is:

(i) The coefficients of non-basic variables in the objective function
(ii) The components of the vectors corresponding to the non-basic variables.

Another feature of the revised simplex method is to treat the objective function $z = \langle c^{\mathrm{T}}, x \rangle$ as a constraint considering $z$ as an unrestricted variable.

## 4.3 Standard Forms of Revised Simplex Method

There are two standard forms of the revised simplex method.

1. **Standard form I**: In this form, an initial basis matrix is obtained by introducing slack variables and/or surplus variables. Basically, an identity is considered as an initial basis matrix.
2. **Standard form II**: In this form, artificial variables are also required for an identity matrix which is considered as the initial basis matrix. Thus, a two-phase method is used to handle the artificial variables. The method for solving standard form I may also be applied to solve these types of problems.

## 4.4 Standard Form I of Revised Simplex Method

Let us consider the following LPP:

Maximize $z = \langle c^{\mathrm{T}}, x \rangle$ subject to $Ax = b,\ x \geq 0$,

where $c,\ x^{\mathrm{T}} \in R^n,\ b^{\mathrm{T}} \in R^m$ and $A$ is an $m \times n$ real matrix of rank $m$.

In the revised simplex method we consider the objective function $z = \langle c^{\mathrm{T}}, x \rangle$ as one of the constraints and $z$ as a variable. Thus, the problem is now to solve the new system of $(m+1)$ simultaneous linear equations in $(n+1)$ variables $x_1, x_2, \ldots, x_n, z$, for which $z$ is unrestricted in sign and may be as large as possible. The set of constraints can thus be represented as

$$\begin{aligned} &Ax + O \cdot z = b \\ &-\left\langle c^{\mathrm{T}}, x \right\rangle + z = 0 \\ &x \geq 0, z \text{ is unrestricted in sign} \end{aligned}$$

$$\text{or} \quad \begin{pmatrix} A & O \\ -c & 1 \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}, \quad x \geq 0. \tag{4.1}$$

Let $B$ be an initial basis submatrix of $A$ and $x_B = B^{-1}b$ be an initial basic feasible solution (BFS) of the original LPP. Then the initial BFS of the reformulated problem (4.1) is given by

$$\begin{aligned} Bx_B + O \cdot z &= b \\ -\langle c_B^{\mathrm{T}}, x\rangle + z &= 0 \end{aligned}$$

$$\text{or} \quad \begin{pmatrix} B & O \\ -c_B & 1 \end{pmatrix}\begin{pmatrix} x_B \\ z \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

$$\text{or} \quad \widehat{B}\widehat{x}_B = \widehat{b} \text{ where } \widehat{B} = \begin{pmatrix} B & O \\ -c_B & 1 \end{pmatrix}, \widehat{x}_B = \begin{pmatrix} x_B \\ z \end{pmatrix}, \widehat{b} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

$$\text{or} \quad \widehat{x}_B = \widehat{B}^{-1}\widehat{b},$$

which is the initial BFS of (4.1).

We note that $\widehat{x}_B$ is called the BFS whenever $x_B$ is the BFS of the original LPP, even when $z$ is unrestricted in sign.

$$\text{Let} \quad \widehat{B}^{-1} = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$$

$$\therefore \widehat{B}^{-1}\widehat{B} = \begin{pmatrix} I & O \\ O & 1 \end{pmatrix}$$

$$\text{or} \quad \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}\begin{pmatrix} B & O \\ -c_B & 1 \end{pmatrix} = \begin{pmatrix} I & O \\ O & 1 \end{pmatrix}$$

$$\text{or} \quad \begin{pmatrix} A_1B - A_2c_B & A_2 \\ A_3B - A_4c_B & A_4 \end{pmatrix} = \begin{pmatrix} I & O \\ O & 1 \end{pmatrix}$$

$$\begin{aligned} &\therefore A_1B - A_2c_B = I. \\ &A_2 = O \end{aligned} \tag{4.2}$$

$$\begin{aligned} &A_3B - A_4c_B = O \\ &A_4 = 1. \end{aligned} \tag{4.3}$$

From (4.2), we have $A_1B - A_2c_B = I$

$$\text{or} \quad A_1B = I \quad [\because A_2 = O]$$

$$\text{or} \quad A_1 = B^{-1}.$$

From (4.3), $A_3B - A_4c_B = O$

$$\text{or} \quad A_3B - c_B = O \quad [\because A_4 = 1]$$

$$\text{or} \quad A_3B = c_B$$

$$\text{or} \quad A_3 = c_B B^{-1}$$

$$\therefore\ \widehat{B}^{-1} = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} = \begin{pmatrix} B^{-1} & O \\ c_B B^{-1} & 1 \end{pmatrix}.$$

Let us define a new $(m+1) \times n$ matrix

$$\widehat{y} = \widehat{B}^{-1}\widehat{A}, \text{ where } \widehat{A} = \begin{pmatrix} A \\ -c \end{pmatrix}$$

$$\therefore \hat{y} = \begin{pmatrix} B^{-1} & O \\ c_B B^{-1} & 1 \end{pmatrix}\begin{pmatrix} A \\ -c \end{pmatrix} = \begin{pmatrix} B^{-1}A \\ c_B B^{-1}A - c \end{pmatrix}$$

$$= \begin{pmatrix} y \\ c_B y - c \end{pmatrix} \quad \left[\because B^{-1}A = y\right].$$

Now, $c_B y - c$

$$= (c_{B_1}, c_{B_2}, \ldots, c_{B_m}) \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{pmatrix} - (c_1, c_2, \ldots, c_n)$$

$$= \left( \sum_{i=1}^{m} c_{B_i} y_{i1} - c_1, \sum_{i=1}^{m} c_{B_i} y_{i2} - c_2, \ldots, \sum_{i=1}^{m} c_{B_i} y_{\text{in}} - c_n \right)$$

$$= (z_1 - c_1, z_2 - c_2, \ldots, z_n - c_n)$$

$$\therefore\ \widehat{y}_j = \begin{pmatrix} y_j \\ z_j - c_j \end{pmatrix},) \quad j = 1, 2, \ldots, n$$

$$\text{or } (\widehat{y}_1, \widehat{y}_2, \ldots, \widehat{y}_n) = \begin{pmatrix} y_1 & y_2 & \cdots & y_n \\ z_1 - c_1 & z_2 - c_2 & \cdots & z_n - c_n \end{pmatrix}. \tag{4.4}$$

From these expressions, it is clear that the first $m$ components (i.e. the first $m$ rows) of the $j$th column constitute the vector $y_j$ and the $(m+1)$th row is $z_j - c_j$ (the net evaluation).

Hence, the net evaluations are the components of

$$c_B B^{-1}A - c = \left(c_B B^{-1} \ 1\right) \begin{pmatrix} A \\ -c \end{pmatrix}.$$

$$\text{Also,}\quad \hat{x}_B = \begin{pmatrix} x_B \\ z \end{pmatrix} = \hat{B}^{-1}\hat{b} \quad \left[\because \hat{x}_B = \hat{B}^{-1}\hat{b}\right],$$

where the first $m$ components of $\hat{x}_B$ are $x_B$ and the $(m+1)$th component is the corresponding value of the objective function.

## 4.5 Computational Procedure of Standard Form I of Revised Simplex Method

Step 1 Introduce slack and surplus variables, if needed, and reformulate the given LPP in maximization standard form.

Step 2 Obtain an initial BFS with an initial basis $B = I_m$ and obtain the inverse of the new basis matrix

$$\hat{B} = \begin{pmatrix} B & O \\ -c_B & 1 \end{pmatrix} \text{by the formula } \hat{B}^{-1} = \begin{pmatrix} B^{-1} & O \\ c_B B^{-1} & 1 \end{pmatrix}.$$

Step 3 Find $\hat{A}$ and $\hat{b}$ by the formula

$$\hat{A} = \begin{pmatrix} A \\ -c \end{pmatrix}, \quad \hat{b} = \begin{pmatrix} b \\ 0 \end{pmatrix}.$$

Step 4 Compute the net evaluations $z_j - c_j (j = 1, 2, \ldots, n)$ by using the formula

$$(z_j - c_j) = (c_B B^{-1} \;\; 1)\begin{pmatrix} A \\ -c \end{pmatrix}.$$

(a) If all $z_j - c_j$ are non-negative, the current BFS is an optimum one.

(b) If at least one $z_j - c_j$ is negative, determine the most negative of them, say $z_k - c_k$; then the corresponding vector $\hat{y}_k$ enters the basis. Go to Step 5.

Step 5 Compute $\hat{y}_k = \hat{B}^{-1}\hat{a}_k$. In that case,

(a) If all $\hat{y}_{ik} \le 0$, there exists an unbounded optimum solution to the given problem.

(b) If at least one $\hat{y}_{ik} > 0$, find the minimum ratios of the components of $\hat{x}_B$ and the corresponding positive $\hat{y}_{ik}$ and determine the outgoing vector and key or leading element.

Step 6 Write down the results obtained in Step 2 through Step 5 in a tabular form known as a revised simplex table.

Step 7 Update the $B^{-1}$ matrix by suitable row operations to convert the key element to unity and all other elements of the entering column to zero, and then update the current BFS using the formula $\widehat{x}_B = \widehat{B}^{-1}\widehat{b}$.

Step 8 Go to Step 4 and repeat the procedure until an optimum BFS is obtained or there is an indication of an unbounded solution.

*Example 1* Using the revised simplex method, solve the following LPP:

$$\begin{aligned}
&\text{Maximize} \quad z = 6x_1 - 2x_2 + 3x_3\\
&\text{subject to}\\
&\qquad 2x_1 - x_2 + 2x_3 \le 2\\
&\qquad x_1 + 4x_3 \le 4 \;\; \text{and} \;\; x_1, x_2, x_3 \ge 0.
\end{aligned}$$

**Solution**

The given LPP is a maximization problem. Here all $b_i$'s ($i$ =1, 2) are positive.

Now introducing the slack variables $x_4 \ge 0$ and $x_5 \ge 0$ in the first and second constraints respectively, the given LPP is rewritten in standard form as follows:

$$\begin{aligned}
&\text{Maximize z} = 6x_1 - 2x_2 + 3x_3 + 0\cdot x_4 + 0\cdot x_5\\
&\text{subject to}\\
&\qquad 2x_1 - x_2 + 2x_3 + x_4 = 2\\
&\qquad x_1 + 4x_3 + x_5 \qquad = 4 \text{ and } x_j \ge 0, \quad j = 1, 2, \ldots, 5\\
&\text{or Maximize } z = \langle c^{\mathrm{T}}, x\rangle\\
&\text{subject to } Ax = b, \;\; x \ge 0,
\end{aligned}$$

where $c = (6 \quad -2 \quad 3 \quad 0 \quad 0)$,

$$x = (x_1\, x_2\, x_3\, x_4\, x_5)^{\mathrm{T}}, \quad A = \begin{pmatrix} 2 & -1 & 2 & 1 & 0\\ 1 & 0 & 4 & 0 & 1\end{pmatrix}, \quad b = \begin{pmatrix} 2\\ 4\end{pmatrix}.$$

The initial BFS is $x_4 = 2,\; x_5 = 4$ with $B = I_2 = \begin{pmatrix} 1 & 0\\ 0 & 1\end{pmatrix}$ as the initial basis.

Here $c_B = (0,\; 0)$.

$$\therefore\; \widehat{B}^{-1} = \begin{pmatrix} B^{-1} & 0\\ c_B B^{-1} & 1\end{pmatrix} = \begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} \left[\because c_B B^{-1} = (0,\; 0)\begin{pmatrix} 1 & 0\\ 0 & 1\end{pmatrix} = (0,\; 0)\right].$$

Since the net evaluations $z_j - c_j$ are the components of $(c_B B^{-1}\ 1)\begin{pmatrix} A \\ -c \end{pmatrix}$, the net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned}(z_1 - c_1\ z_2 - c_2\ z_3 - c_3) &= \left(c_B B^{-1}\ 1\right)(\hat{a}_1\ \hat{a}_2\ \hat{a}_3) \\ &= (0\quad 0\quad 1)\begin{pmatrix} 2 & -1 & 2 \\ 1 & 0 & 4 \\ -6 & 2 & -3 \end{pmatrix} \\ &\left[\because \hat{A} = \begin{pmatrix} A \\ -c \end{pmatrix} = \begin{pmatrix} 2 & -1 & 2 & 1 & 0 \\ 1 & 0 & 4 & 0 & 1 \\ -6 & 2 & -3 & 0 & 0 \end{pmatrix}\right] \\ &= (-6\quad 2\quad -3).\end{aligned}$$

These calculations indicate that $z_1 - c_1$ is the most negative, and hence $\hat{y}_1$ enters the basis.

$$\text{Now,}\quad \hat{y}_1 = \hat{B}^{-1}\hat{a}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \\ -6 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -6 \end{pmatrix} = \begin{pmatrix} y_1 \\ z_1 - c_1 \end{pmatrix}.$$

$$\text{Also,}\quad \hat{x}_B = \hat{B}^{-1}\hat{b} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix}\left[\because \hat{b} = \begin{pmatrix} b \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix}\right].$$

$$\therefore\ x_B = \begin{pmatrix} 2 \\ 4 \end{pmatrix},\quad z_B = 0.$$

Now we construct the initial revised simplex table.

| $\hat{y}_B$ | $\hat{x}_B$ | $\hat{B}^{-1}$ | | | $\hat{y}_1$ | Mini ratio |
|---|---|---|---|---|---|---|
| $\hat{y}_4$ | 2 | 1 | 0 | 0 | $2^*$ | 1 |
| $\hat{y}_5$ | 4 | 0 | 1 | 0 | 1 | 4 |
| | 0 | 0 | 0 | 1 | −6 | |

The minimum ratio corresponds to the first basic variable, i.e. $x_4$; hence, $\hat{y}_4$ leaves the basis and $\hat{y}_{11} = 2$ is the key element.

**First iteration**:

We introduce $\hat{y}_1$ and drop $\hat{y}_4$.

In this iteration,

$$\widehat{B}^{-1} = \begin{pmatrix} 1/2 & 0 & 0 \\ -1/2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \quad \text{by} \quad \begin{array}{l} R_1' = \frac{1}{2}R_1 \\ R_2' = R_2 - R_1' \\ R_3' = R_3 + 6R_1' \end{array}.$$

Here $c_B = (6 \quad 0)$.

The net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned}(z_2 - c_2 \; z_3 - c_3 \; z_4 - c_4) &= \left(c_B B^{-1} \; 1\right)(\widehat{a}_2 \; \widehat{a}_3 \; \widehat{a}_4) \\ &= (3 \quad 0 \quad 1)\begin{pmatrix} -1 & 2 & 1 \\ 0 & 4 & 0 \\ 2 & -3 & 0 \end{pmatrix} \quad [(c_B B^{-1} \; 1) = (3 \quad 0 \quad 1)] \\ &= (-1 \quad 3 \quad 3).\end{aligned}$$

Clearly, $z_2 - c_2$ is negative. Hence, $\widehat{y}_2$ enters the basis.

Now, $\widehat{y}_2 = \widehat{B}^{-1}\widehat{a}_2 = \begin{pmatrix} 1/2 & 0 & 0 \\ -1/2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 1/2 \\ -1 \end{pmatrix}$

and $\widehat{x}_B = \widehat{B}^{-1}\widehat{b} = \begin{pmatrix} 1/2 & 0 & 0 \\ -1/2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 6 \end{pmatrix}$

$\therefore x_B = (1 \quad 3)^{\mathrm{T}}, \;\; z_B = 6.$

Now we construct the next revised simplex table.

| $\widehat{y}_B$ | $\widehat{x}_B$ | $\widehat{B}^{-1}_{\text{next}}$ | | | $\widehat{y}_2$ | Mini ratio |
|---|---|---|---|---|---|---|
| $\widehat{y}_1$ | 1 | 1/2 | 0 | 0 | $-1/2$ | – |
| $\widehat{y}_5$ | 3 | $-1/2$ | 1 | 0 | 1/2 | 6 |
| | 6 | 3 | 0 | 1 | $-1$ | |

**Second iteration**: We introduce $\widehat{y}_2$ and drop $\widehat{y}_5$.
In this iteration,

$$\widehat{B}^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ 2 & 2 & 1 \end{pmatrix} \quad \text{by} \quad \begin{array}{l} R_2' = 2R_2 \\ R_1' = R_1 + \frac{1}{2}R_2' \\ R_3' = R_3 + R_2' \end{array}.$$

$$\text{Here} \quad c_B = (6 \quad -2)$$

The net evaluations $z_j - c_j$ corresponding to the non-basic variables are given by

$$\begin{aligned}(z_3 - c_3 \; z_4 - c_4 \, z_5 - c_5) &= \left(c_B B^{-1} \;\; 1\right)\left(\widehat{a}_3 \quad \widehat{a}_4 \quad \widehat{a}_5\right) \\ &= (2 \quad 2 \quad 1)\begin{pmatrix} 2 & 1 & 0 \\ 4 & 0 & 1 \\ -3 & 0 & 0 \end{pmatrix} \\ &= (9 \quad 2 \quad 2).\end{aligned}$$

Since all $z_j - c_j > 0$ correspond to non-basic variables, an optimum basic feasible solution is obtained.

The optimal solution is given by

$$\widehat{x}_B = \widehat{B}^{-1}\widehat{b} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ 2 & 2 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ 12 \end{pmatrix}$$

$$\therefore x_B = (4 \quad 6)^{\mathrm{T}} \text{and} \quad z_B = 12.$$

Hence, the optimal solution of the given LPP is $x_1 = 4$, $x_2 = 6$, $x_3 = 0$ and $z_{\max} = 12$.

*Example 2* Using the revised simplex method, solve the following LPP:

$$\begin{aligned}&\text{Maximize} \quad z = x_1 + x_2 + 3x_3 \\ &\text{subject to} \\ &\qquad 3x_1 + 2x_2 + x_3 \le 3 \\ &\qquad 2x_1 + x_2 + 2x_3 \le 2 \text{ and } \quad x_1, x_2, x_3 \ge 0.\end{aligned}$$

**Solution**

This is a maximization problem. Here all $b_i$'s are positive.

Now introducing the slack variables $x_4 \ge 0$ and $x_5 \ge 0$ in the first and second constraints respectively, the given LPP is rewritten in the standard form as follows:

$$\begin{aligned}&\text{Maximize} \quad z = x_1 + x_2 + 3x_3 + 0 \cdot x_4 + 0 \cdot x_5 \\ &\text{subject to} \quad 3x_1 + 2x_2 + x_3 + x_4 = 3 \\ &\qquad\qquad 2x_1 + x_2 + 2x_3 + x_5 = 2 \text{ and } x_j \ge 0, \; j = 1, 2, \ldots, 5\end{aligned}$$

$$\begin{aligned}&\text{or Maximize}\, z = \left\langle c^{\mathrm{T}}, x \right\rangle \\ &\text{subject to } Ax = b, \; x \ge 0,\end{aligned}$$

$$\text{where} \quad c = (1 \quad 1 \quad 3 \quad 0 \quad 0),$$

$$x = (x_1 \;\; x_2 \;\; x_3 \;\; x_4 \;\; x_5)^{\mathrm{T}}, \; A = \begin{pmatrix} 3 & 2 & 1 & 1 & 0 \\ 2 & 1 & 2 & 0 & 1 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Here the initial BFS is $x_4 = 3,\ x_5 = 2$ with $B = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ as the initial basis.

$$\text{Here} \quad c_B = (\,0 \quad 0\,).$$

Now $\widehat{A} = \begin{pmatrix} A \\ -c \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 & 1 & 0 \\ 2 & 1 & 2 & 0 & 1 \\ -1 & -1 & -3 & 0 & 0 \end{pmatrix}$ and $\widehat{b} = \begin{pmatrix} b \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$

$\therefore\ \widehat{B}^{-1} = \begin{pmatrix} B^{-1} & 0 \\ c_B B^{-1} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ $\left[\therefore\ c_B B^{-1} = (\,0 \quad 0\,)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = (\,0 \quad 0\,)\right].$

The net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned}
(z_1 - c_1\ z_2 - c_2\ z_3 - c_3) &= \left(c_B B^{-1}\ 1\right)(\,\widehat{a}_1 \quad \widehat{a}_2 \quad \widehat{a}_3\,) \\
&= (\,0 \quad 0 \quad 1\,)\begin{pmatrix} 3 & 2 & 1 \\ 2 & 1 & 2 \\ -1 & -1 & -3 \end{pmatrix} \\
&= (\,-1 \quad -1 \quad -3\,).
\end{aligned}$$

This result indicates that $z_3 - c_3$ is the most negative. Hence, $\widehat{y}_3$ enter the basis.

$$\text{Now } \widehat{y}_3 = \widehat{B}^{-1}\widehat{a}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} \quad \left[\because\ \widehat{a}_3 = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}\right].$$

$$\text{Also, } \widehat{x}_B = \widehat{B}^{-1}\widehat{b} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}.$$

$$\therefore\ x_B = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \quad \text{and} \quad z_B = 0.$$

Now we construct the initial revised simplex table.

| $\widehat{y}_B$ | $\widehat{x}_B$ | $\widehat{B}^{-1}$ | | | $\widehat{y}_3$ | Mini ratio |
|---|---|---|---|---|---|---|
| $\widehat{y}_4$ | 3 | 1 | 0 | 0 | 1 | 3 |
| $\widehat{y}_5$ | 2 | 0 | 1 | 0 | 2 | 1 |
| | 0 | 0 | 0 | 1 | −3 | |

The minimum ratio corresponds to the second basic variable, i.e. $x_5$; hence, $\widehat{y}_5$ leaves the basis.

**First iteration**:

In this iteration,

$$\widehat{B}^{-1} = \begin{pmatrix} 1 & -1/2 & 0 \\ 0 & 1/2 & 0 \\ 0 & 3/2 & 1 \end{pmatrix} \quad \text{by} \quad \begin{array}{l} R_2' = \frac{R_2}{2} \\ R_1' = R_1 - R_2' \\ R_3' = R_3 + 3R_2' \end{array}.$$

The net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned}
(z_1 - c_1\ z_2 - c_2\ z_5 - c_5) &= \left(c_B B^{-1} 1\right)\left(\widehat{a}_1 \quad \widehat{a}_2 \quad \widehat{a}_5\right) \\
&= \left(0 \quad \tfrac{3}{2} \quad 1\right)\begin{pmatrix} 3 & 2 & 0 \\ 2 & 1 & 1 \\ -1 & -1 & 0 \end{pmatrix} \\
&\left[c_B B^{-1} = (0 \quad 3)\begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix} = \left(0 \quad \tfrac{3}{2}\right)\right. \\
&= \left(2 \quad \tfrac{1}{2} \quad \tfrac{3}{2}\right).
\end{aligned}$$

Since all $z_j - c_j \geq 0$, an optimum basic feasible solution is obtained. The optimal solution is given by

$$\widehat{x}_B = \widehat{B}^{-1}\widehat{b} = \begin{pmatrix} 1 & -1/2 & 0 \\ 0 & 1/2 & 0 \\ 0 & 3/2 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$$

$$\therefore x_B = (2 \quad 1)^{\mathrm{T}} \text{ and } \quad z_B = 3.$$

Hence, the optimal solution is $x_1 = 0,\ \ x_2 = 0,\ \ x_3 = 1$, and the maximum value of $z$ is 3.

## 4.6 Standard Form II of Revised Simplex Method

In this form, if artificial variables appear in an LPP, then the initial basis matrix is obtained by using the vectors' corresponding to slack and artificial variables. In that case, to solve the problem, either Charne's Big M method or the two-phase method may be used. We denote these methods as Charne's Big M revised simplex method and the two-phase revised simplex method. We shall discuss both methods separately.

### 4.6.1 Charne's Big M Revised Simplex Method

In this case, the auxiliary basis matrix

$$\widehat{B} = \begin{pmatrix} B & 0 \\ -c_B & 1 \end{pmatrix}$$

is not a unit basis, and the associated profit vector $c_B$ is not a null vector. Thus,

$$\widehat{B}^{-1} = \begin{pmatrix} B^{-1} & 0 \\ c_B B^{-1} & 1 \end{pmatrix}$$

will not be an identity matrix. This is the only difference between the standard forms I and II of the revised simplex method. All other theoretical developments and computational procedures are the same as in standard form I.

*Example 3:* Using the revised simplex method, solve the following LPP:

$$\begin{array}{ll} \text{Maximize} & z = x_1 + 2x_2 \\ \text{subject to} & 2x_1 + 5x_2 \geq 6 \\ & x_1 + x_2 \geq 2,\ x_1, x_2 \geq 0. \end{array}$$

**Solution**
This is a maximization problem. All the $b_i$'s ($i$ = 1, 2) are positive. Now introducing the surplus variables $x_3 \geq 0$ and $x_4 \geq 0$ in the first and second constraints respectively, the given LPP is written as follows:

$$\begin{array}{ll} \text{Maximize} & z = x_1 + 2x_2 + 0 \cdot x_3 + 0 \cdot x_4 \\ \text{subject to} & 2x_1 + 5x_2 - x_3 = 6 \\ & x_1 + x_2 - \mathrm{x}_4 = 2, \quad x_j \geq 0,\ j = 1, 2, 3, 4. \end{array}$$

The coefficient matrix of the constraints does not contain a unit basis matrix. To obtain a unit matrix, two artificial variables $x_5,\ x_6$ are added to the left-hand side of the first and second constraints. Then the LPP becomes:

$$\begin{array}{ll} \text{Maximize} & z = x_1 + 2x_2 + 0 \cdot x_3 + 0 \cdot x_4 - Mx_5 - Mx_6 \\ \text{subject to} & 2x_1 + 5x_2 - x_3 + x_5 = 6 \\ & x_1 + x_2 - \mathrm{x}_4 + x_6 = 2, \quad x_j \geq 0,\ \ j = 1, 2, \ldots, 6 \\ \text{or Maximize} & z = \langle c^{\mathrm{T}}, x \rangle \\ \text{subject to} & Ax = b, x \geq 0, \end{array}$$

where $c = (1 \quad 2 \quad 0 \quad 0 - M - M),\ x = (x_1\ x_2 \ldots x_6)^{\mathrm{T}}$

$$b = \begin{pmatrix} 6 \\ 2 \end{pmatrix}\ A = \begin{pmatrix} 2 & 5 & -1 & 0 & 1 & 0 \\ 1 & 1 & 0 & -1 & 0 & 1 \end{pmatrix}.$$

The coefficients of the variables constitute the column vectors

$$y_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad y_2 = \begin{pmatrix} 5 \\ 1 \end{pmatrix}, \quad y_3 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad y_4 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad y_5 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad y_6 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and the requirement vector $b = \begin{pmatrix} 6 \\ 2 \end{pmatrix}$.

Here the artificial vectors $y_5$ and $y_6$ constitute the initial basis matrix $B = I_2$, and taking this as the initial basis, we have the initial BFS as $x_B = B^{-1}b = I_2 b = b \geq 0$.

$$\text{Therefore} \quad x_B = (x_{B_1} \quad x_{B_2})^{\mathrm{T}} = (x_5 \quad x_6)^{\mathrm{T}} = (6 \quad 2)^{\mathrm{T}}$$

$$\therefore x_5 = 6, \; x_6 = 2.$$

$$\text{Here} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad c_B = (-M \quad -M).$$

$$\text{Now,} \quad \widehat{A} = \begin{pmatrix} A \\ -c \end{pmatrix} = \begin{pmatrix} 2 & 5 & -1 & 0 & 1 & 0 \\ 1 & 1 & 0 & -1 & 0 & 1 \\ -1 & -2 & 0 & 0 & M & M \end{pmatrix}$$

$$\therefore \widehat{B}^{-1} = \begin{pmatrix} B^{-1} & 0 \\ c_B B^{-1} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -M & -M & 1 \end{pmatrix}$$

$$\left[\because c_B B^{-1} = (-M \quad -M)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = (-M \quad -M).\right.$$

Now, the net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned} (z_1 - c_1 \; z_2 - c_2 \; z_3 - c_3 \; z_4 - c_4) &= \left(c_B B^{-1} \; 1\right)(\widehat{a}_1 \; \widehat{a}_2 \; \widehat{a}_3 \; \widehat{a}_4) \\ &= (-M \quad -M \quad 1)\begin{pmatrix} 2 & 5 & -1 & 0 \\ 1 & 1 & 0 & -1 \\ -1 & -2 & 0 & 0 \end{pmatrix} \\ &= (-3M - 1 \quad -6M - 2 \quad M \quad M). \end{aligned}$$

The result indicates that $z_2 - c_2$ is the most negative, and hence $\widehat{y}_2$ enters the basis.

$$\text{Now, } \hat{y}_2 = \hat{B}^{-1}\hat{a}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -M & -M & 1 \end{pmatrix}\begin{pmatrix} 5 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ -6M-2 \end{pmatrix} \quad \left[\because \hat{a}_2 = \begin{pmatrix} 5 \\ 1 \\ -2 \end{pmatrix}\right].$$

$$\text{Also, } \hat{x}_B = \hat{B}^{-1}\hat{b} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -M & -M & 1 \end{pmatrix}\begin{pmatrix} 6 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ -8M \end{pmatrix}.$$

$$\therefore\ x_B = \begin{pmatrix} 6 \\ 2 \end{pmatrix} \quad \text{and} \quad z_B = -8M.$$

Now, we construct the initial revised simplex table.

| $\hat{y}_B$ | $\hat{x}_B$ | $\hat{B}^{-1}$ | | | $\hat{y}_2$ | Mini ratio |
|---|---|---|---|---|---|---|
| $\hat{y}_5$ | 6 | 1 | 0 | 0 | 5 | 6/5 |
| $\hat{y}_6$ | 2 | 0 | 1 | 0 | 1 | 2 |
| | | $-M$ | $-M$ | 1 | $-6M-2$ | |

The minimum ratio corresponds to the first basic variable, i.e.$x_5$; hence, $\hat{y}_5$ leaves the basis.

**First iteration**:

We introduce $\hat{y}_2$ and drop $\hat{y}_5$.In this case,

$$\hat{B}^{-1} = \begin{pmatrix} 1/5 & 0 & 0 \\ -1/5 & 1 & 0 \\ (M+2)/5 & -M & 1 \end{pmatrix} \quad \text{by} \quad \begin{array}{l} R'_1 = \frac{R_1}{5} \\ R'_2 = R_2 - R'_1 \\ R'_3 = R_3 + (6M+2)R'_1. \end{array}$$

Now, the net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned}(z_1 - c_1\ z_3 - c_3\ z_4 - c_4\ z_5 - c_5) &= \left(c_B B^{-1}\ 1\right)(\hat{a}_1\ \hat{a}_3\ \hat{a}_4\ \hat{a}_5) \\ &= \left(\tfrac{M+2}{5}\quad -M \quad 1\right)\begin{pmatrix} 2 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ -1 & 0 & 0 & M \end{pmatrix} \\ &= \left(-\tfrac{3M+1}{5}\quad -\tfrac{M+2}{5}\quad M \quad \tfrac{6M+2}{5}\right).\end{aligned}$$

Since $z_1 - c_1$ is the most negative, then $\hat{y}_1$ enters the basis.

$$\text{Now, } \hat{y}_1 = \hat{B}^{-1}\hat{a}_1 = \begin{pmatrix} 1/5 & 0 & 0 \\ -1/5 & 1 & 0 \\ (M+2)/5 & -M & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2/5 \\ 3/5 \\ -\frac{3M+1}{5} \end{pmatrix}.$$

$$\text{Again, } \widehat{x}_B = \widehat{B}^{-1}\widehat{b} = \begin{pmatrix} 1/5 & 0 & 0 \\ -1/5 & 1 & 0 \\ (M+2)/5 & -M & 1 \end{pmatrix}\begin{pmatrix} 6 \\ 2 \\ 0 \end{pmatrix} \quad \left[\because \widehat{b} = \begin{pmatrix} b \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 0 \end{pmatrix}\right]$$

$$= \begin{pmatrix} \frac{6}{5} & \frac{4}{5} & \frac{-4M+12}{5} \end{pmatrix}^{\mathrm{T}}$$

$$\therefore x_B = \begin{pmatrix} \frac{6}{5} & \frac{4}{5} \end{pmatrix}^{\mathrm{T}} \text{ and } z_B = \frac{-4M+12}{5}.$$

The next revised simplex table is as follows:

| $\widehat{y}_B$ | $\widehat{x}_B$ | $\widehat{B}^{-1}$ | | | $\widehat{y}_1$ | Mini ratio |
|---|---|---|---|---|---|---|
| $\widehat{y}_2$ | 6/5 | 1/5 | 0 | 0 | 2/5 | 3 |
| $\widehat{y}_6$ | 4/5 | $-1/5$ | 1 | 0 | 3/5 | 4/3 |
| | $\frac{-4M+12}{5}$ | $\frac{M+2}{5}$ | $-M$ | 1 | $-\frac{3M+1}{5}$ | |

**Second iteration**:

We introduce $\widehat{y}_1$ and drop $\widehat{y}_6$.

In this case,

$$\widehat{B}^{-1} = \begin{pmatrix} 1/3 & -2/3 & 0 \\ -1/3 & 5/3 & 0 \\ 1/3 & 1/3 & 1 \end{pmatrix} \quad \text{by} \quad R_2' = \tfrac{5}{3}R_2, R_1' = R_1 - \tfrac{2}{5}R_2', R_3' = R_3 + \tfrac{3M+1}{5}R_2'.$$

Now, the net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned}(z_3 - c_3\ z_4 - c_4\ z_5 - c_5\ z_6 - c_6) &= \left(c_B B^{-1}\ 1\right)\left(\widehat{a}_3 \quad \widehat{a}_4 \quad \widehat{a}_5 \quad \widehat{a}_6\right) \\ &= \left(\tfrac{1}{3} \quad \tfrac{1}{3} \quad 1\right)\begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & M & M \end{pmatrix} \\ &= \left(-\tfrac{1}{3} \quad -\tfrac{1}{3} \quad M + \tfrac{1}{3} \quad M + \tfrac{1}{3}\right).\end{aligned}$$

Clearly, $z_3 - c_3$ and $z_4 - c_4$ are most negative. Any one of them may enter the basis. Arbitrarily, we take $\widehat{y}_3$ as the entering vector.

$$\text{Now, } \widehat{y}_3 = \widehat{B}^{-1}\widehat{a}_3 = \begin{pmatrix} 1/3 & -2/3 & 0 \\ -1/3 & 5/3 & 0 \\ 1/3 & 1/3 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1/3 \\ 1/3 \\ -1/3 \end{pmatrix}.$$

$$\text{Again, } \widehat{x}_B^{-1} = \widehat{B}^{-1}\widehat{b} = \begin{pmatrix} 1/3 & -2/3 & 0 \\ -1/3 & 5/3 & 0 \\ 1/3 & 1/3 & 1 \end{pmatrix}\begin{pmatrix} 6 \\ 2 \\ 0 \end{pmatrix} \quad \left[\because \widehat{b} = \begin{pmatrix} b \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 0 \end{pmatrix}\right]$$

$$= \begin{pmatrix} \frac{2}{3} & \frac{4}{3} & \frac{8}{3} \end{pmatrix}^{\mathrm{T}}$$

$$\therefore x_B = \begin{pmatrix} \frac{2}{3} & \frac{4}{3} \end{pmatrix}^{\mathrm{T}} \quad \text{and} \quad z_B = \frac{8}{3}.$$

The next revised simplex table is as follows:

| $\widehat{y}_B$ | $\widehat{x}_B$ | $\widehat{B}^{-1}_{\text{next}}$ | | | $\widehat{y}_3$ | Mini ratio |
|---|---|---|---|---|---|---|
| $\widehat{y}_2$ | 2/3 | 1/3 | −2/3 | 0 | −1/3 | – |
| $\widehat{y}_1$ | 4/3 | −1/3 | 5/3 | 0 | 1/3 | 4 |
| | 8/3 | 1/3 | 1/3 | 1 | −1/3 | |

The minimum ratio corresponds to the second basic variable, i.e. $x_1$; hence, $\widehat{y}_1$ leaves the basis.

**Third iteration**:

In this case,

$$\widehat{B}^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 5 & 0 \\ 0 & 2 & 1 \end{pmatrix} \quad \text{by} \quad R_2' = 3R_1, R_1' = R_1 + \frac{1}{3}R_2', R_3' = R_3 + \frac{1}{3}R_2'.$$

The net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned}(z_1 - c_1 \;\; z_4 - c_4 \;\; z_5 - c_5 \;\; z_6 - c_6) &= \left(c_B B^{-1} \;\; 1\right)\left(\widehat{a}_1 \quad \widehat{a}_4 \quad \widehat{a}_5 \quad \widehat{a}_6\right) \\ &= (0 \quad 2 \quad 1)\begin{pmatrix} 2 & 0 & 1 & 0 \\ 1 & -1 & 0 & 1 \\ -1 & 0 & M & M \end{pmatrix} = (1 \quad -2 \quad M \quad M).\end{aligned}$$

Clearly, $z_4 - c_4$ is negative. $\therefore \widehat{y}_4$ enters the basis.

$$\text{Now, } \widehat{y}_4 = \widehat{B}^{-1}\widehat{a}_4 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 5 & 0 \\ 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -5 \\ -2 \end{pmatrix}.$$

Since the first two components of the entering vector $\widehat{y}_4$ are negative, there is an indication of an unbounded solution. Hence, the given LPP has an unbounded solution.

### 4.6.2 Two-Phase Revised Simplex Method

In this method, besides the original objective function, an auxiliary objective function $z^*$ is formed by considering the coefficients of artificial variables as $-1$ and the coefficients of the other variables as zero. Then both the objective functions are converted to constraints, considering $z$ and $z^*$ as unrestricted variables.

*Example 4* Use the two-phase revised simplex method to solve the following LPP:

$$\begin{aligned} \text{Maximize} \quad & z = 3x_1 - x_2 \\ \text{subject to} \quad & 2x_1 + x_2 \geq 2 \\ & x_1 + 3x_2 \leq 2 \\ & \text{and}\, x_1,\ x_2 \geq 0. \end{aligned}$$

**Solution**

This is a maximization problem. Here all $b_i$'s are positive. Now introducing the slack variable $x_3$ and surplus variable $x_4$ in the second and first constraints respectively, and then considering the given objective function as a constraint, with $z$ as a variable, we get the reduced set of constraints as follows:

$$\begin{aligned} & 2x_1 + x_2 - x_4 \quad = 2 \\ & x_1 + 3x_2 + x_3 \quad = 2 \\ & -3x_1 + x_2 + z = 0 \\ & \text{and} \quad x_j \geq 0, \ \ j \ = 1, 2, 3, 4. \end{aligned}$$

To get the initial basis matrix $B$ as an identity matrix, introducing one artificial variable $x_5(\geq 0)$ in the first constraint, we get the reduced constraints as

$$\begin{aligned} 2x_1 + x_2 - x_4 + x_5 &= 2 \\ x_1 + 3x_2 + x_3 \qquad &= 2 \\ -3x_1 + x_2 + z \qquad &= 0. \end{aligned}$$

We shall solve this problem using the two-phase revised simplex method.

In Phase I, the auxiliary objective function is

$$\text{Maximize} \quad z^* = -x_5$$

i.e. Maximize $z^* = 2x_1 + x_2 - x_4 - 2$ [substituting $x_5$ from the first constraint].

Here the initial BFS is $x_3 = 2,\ x_5 = 2,\ z = 0$ with $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ as the initial basis, i.e. $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Here $c_B = (\,0 \quad 0 \quad 0\,)$.

Now, $B^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

$$\therefore\; B^{-1} = \begin{pmatrix} B^{-1} & 0 \\ c_B B^{-1} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \text{ Here } b^* = \begin{pmatrix} 2 \\ 2 \\ 0 \\ -2 \end{pmatrix}.$$

The net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned} &(\,z_1 - c_1 \quad z_2 - c_2 \quad z_4 - c_4\,) \\ &= (\,0 \quad 0 \quad 0 \quad 1\,)(\,a_1^* \quad a_2^* \quad a_4^*\,) \left[\because A^* = \begin{pmatrix} a_1^* & a_2^* & a_3^* & a_4^* & a_5^* \\ 2 & 1 & 0 & -1 & 1 \\ 1 & 3 & 1 & 0 & 0 \\ -3 & 1 & 0 & 0 & 0 \\ -2 & -1 & 0 & 1 & 0 \end{pmatrix}\right] \\ &= (\,-2 \quad -1 \quad 1\,). \end{aligned}$$

This indicates that $z_1 - c_1$ is the most negative.

Hence, $y_1^*$ enters the basis.

$$\text{Now,}\quad y_1^* = B^{*-1} a_1^* = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ -3 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -3 \\ -2 \end{pmatrix}.$$

$$\text{Also,}\quad x_B^* = B^{*-1} b^* = (\,2 \quad 2 \quad 0 \quad -2\,)^{\mathrm{T}}.$$

Now we construct the initial revised simplex table.

| $y_B^*$ | $x_B^*$ | $B^{*-1}$ | | | | $y_1^*$ | Mini ratio |
|---|---|---|---|---|---|---|---|
| $y_5^*$ | 2 | 1 | 0 | 0 | 0 | $2^*$ | 2/2 = 1 |
| $y_3^*$ | 2 | 0 | 1 | 0 | 0 | 1 | 2/1 = 2 |
| $z$ | 0 | 0 | 0 | 1 | 0 | −3 | |
| $z^*$ | −2 | 0 | 0 | 0 | 1 | −2 | |

According to the minimum ratio, $y_5^*$ leaves the basis and $y_{11}^* = 1$ is the key element.

**First iteration**:

We introduce $y_1^*$ and drop $y_5^*$.

In this iteration, the reduced basis inverse is given by

$$B^{*-1} = \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ -1/2 & 1 & 0 & 0 \\ 3/2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

The net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned}(\, z_2 - c_2 \quad z_4 - c_4 \quad z_5 - c_5 \,) &= (\, 1 \quad 0 \quad 0 \quad 1 \,)(\, a_2^* \quad a_4^* \quad a_5^* \,) \\ &= (\, 1 \quad 0 \quad 0 \quad 1 \,)\begin{pmatrix} 1 & -1 & 1 \\ 3 & 0 & 0 \\ 1 & 0 & 0 \\ -1 & 1 & 0 \end{pmatrix} = (\, 0 \quad 0 \quad 1 \,).\end{aligned}$$

This indicates that all $z_j - c_j \geq 0$ correspond to non-basic variables. Hence, the current solution is an optimum for Phase I. The optimum solution for this phase is

$$x_B^* = B^{*-1} b^* = \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ -1/2 & 1 & 0 & 0 \\ 3/2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 3 \\ 0 \end{pmatrix}$$

$$\therefore x_B = (\, 1 \quad 1 \quad 3 \,)^{\mathrm{T}} \quad \text{and} \quad z_B = 0,$$

i.e. $x_1 = 1,\ \ x_3 = 1,\ \ z = 3$ and Maximize $z^* = 0$.

Since there is no artificial variable in the current basis, we shall apply Phase II.

**Phase II**:

From the basis inverse $B^{*-1}$ of the optimal obtained in Phase I, we have $\widehat{B}^{-1} = \begin{pmatrix} 1/2 & 0 & 0 \\ -1/2 & 1 & 0 \\ 3/2 & 0 & 1 \end{pmatrix}$ by deleting the fourth row and the fourth column.

$$\text{Here} \quad \widehat{A} = \begin{pmatrix} a_1^* & a_2^* & a_3^* & a_4^* \\ 2 & 1 & 0 & -1 \\ 1 & 3 & 1 & 0 \\ -3 & 1 & 0 & 0 \end{pmatrix}.$$

Now the net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$(\, z_2 - c_2 \quad z_4 - c_4 \,) = (\, 3/2 \quad 0 \quad 1 \,)(\, \widehat{a}_2 \quad \widehat{a}_4 \,) = (\, 3/2 \quad 0 \quad 1 \,)\begin{pmatrix} 1 & -1 \\ 3 & 0 \\ 1 & 0 \end{pmatrix} = (\, 5/2 \quad -3/2 \,).$$

This indicates that $z_4 - c_4$ is negative. Hence, $\widehat{y}_4$ enters the basis.

$$\therefore \widehat{y}_4 = \widehat{B}^{-1}\widehat{a}_4 = \begin{pmatrix} 1/2 & 0 & 0 \\ -1/2 & 1 & 0 \\ 3/2 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 1/2 \\ -3/2 \end{pmatrix}$$

and $\widehat{x}_B = \widehat{B}^{-1}\widehat{b}$, where $\widehat{b} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 & 0 \\ -1/2 & 1 & 0 \\ 3/2 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$.

Now we construct the revised simplex table.

| $\widehat{y}_B$ | $\widehat{x}_B$ | $\widehat{B}^{-1}$ | | | $\widehat{y}_4$ | Mini ratio |
|---|---|---|---|---|---|---|
| $\widehat{y}_1$ | 1 | 1/2 | 0 | 0 | −1/2 | – |
| $\widehat{y}_3$ | 1 | −1/2 | 1 | 0 | 1/2 | 1/(½) = 2 |
| $z$ | 3 | 3/2 | 0 | 1 | −3/2 | |

According to the minimum ratio, $\widehat{y}_3$ leaves the basis and $\widehat{y}_{24} = 1/2$ is the key element.

**First iteration**:

In this iteration, the reduced basis inverse matrix is given by

$$\widehat{B}^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ 0 & 3 & 1 \end{pmatrix}.$$

Now the net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned} (\,z_2 - c_2 \quad z_3 - c_3\,) &= (\,0 \quad 3 \quad 1\,)(\,\widehat{a}_2 \quad \widehat{a}_3\,) \\ &= (\,0 \quad 3 \quad 1\,)\begin{pmatrix} 1 & 0 \\ 3 & 1 \\ 1 & 0 \end{pmatrix} = (\,10 \quad 3\,). \end{aligned}$$

Since all $z_j - c_j > 0$ correspond to non-basic variables, an optimum basic feasible solution has been obtained.

The optimal solution is given by

$$\begin{aligned} &\widehat{x}_B = \widehat{B}^{-1}\widehat{b} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ 0 & 3 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 6 \end{pmatrix} \\ &\therefore \widehat{x}_B = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \quad \text{and} \quad z_B = 6, \\ &\text{i.e. } x_1 = 2,\; x_2 = 0 \quad \text{and} \quad z_{\text{Max}} = 6. \end{aligned}$$

*Example 5* Solve the following LPP by the two-phase revised simplex method:

$$\begin{array}{ll}\text{Maximize} & z = x_1 + 2x_2 + 3x_3 + 4x_4 \\ \text{subject to} & 3x_1 + 2x_2 + 3x_3 - x_4 \le 25 \\ & -2x_1 + x_2 - 2x_3 + x_4 \ge 5 \\ & 2x_1 + 2x_2 + x_3 + x_4 = 20 \quad \text{and} \quad x_1,\ x_2\, x_3,\ x_4 \ge 0.\end{array}$$

**Solution**

This is a maximization problem. Here all $b_i$'s are positive. Now introducing the slack variable $x_5(\ge 0)$ and the surplus variable $x_6(\ge 0)$ in the first and second constraints respectively, the given LPP is rewritten in the standard form as follows:

$$\begin{array}{ll}\text{Maximize} & z = x_1 + 2x_2 + 3x_3 + 4x_4 + 0 \cdot x_5 + 0 \cdot x_6 \\ \text{subject to} & 3x_1 + 2x_2 + 3x_3 - x_4 + x_5 = 25 \\ & -2x_1 + x_2 - 2x_3 + x_4 - x_6 = 5 \\ & 2x_1 + 2x_2 + x_3 + x_4 = 20 \text{ and } x_i \ge 0,\ \ i = 1, 2, \ldots, 6.\end{array}$$

Considering the objective function as a constraint and $z$ as a variable, we get the reduced set of constraints as

$$\begin{aligned} 3x_1 + 2x_2 + 3x_3 - x_4 + x_5 \quad &= 25 \\ -2x_1 + x_2 - 2x_3 + x_4 - x_6 \ &= 5 \\ 2x_1 + 2x_2 + x_3 + x_4 \qquad &= 20 \\ -x_1 - 2x_2 - 3x_3 - 4x_4 + z &= 0. \end{aligned}$$

To get the initial unit basis matrix, introducing two artificial variables $x_7(\ge 0)$ and $x_8(\ge 0)$ in the second and third constraints respectively, we get the reduced constraints as

$$\begin{aligned} 3x_1 + 2x_2 + 3x_3 - x_4 + x_5 \qquad &= 25 \\ -2x_1 + x_2 - 2x_3 + x_4 - x_6 + x_7 &= 5 \\ 2x_1 + 2x_2 + x_3 + x_4 + x_8 \qquad &= 20 \\ -x_1 - 2x_2 - 3x_3 - 4x_4 + z \qquad &= 0. \end{aligned}$$

We shall solve this problem by the two-phase method.

In Phase I, the auxiliary objective function is to

$$\text{Maximize } z^* = -x_7 - x_8,$$

i.e. Maximize $z^* = 3x_2 - x_3 + 2x_4 - x_6 - 25$ [substituting $x_7$ and $x_8$ from the second and third constraints].

Here the initial BFS is $x_5 = 25,\ x_7 = 5,\ x_8 = 20,\ z = 0$ with $B = I_4$ as the initial basis.

$$\text{Here} \quad c_B = (\,0 \quad 0 \quad 0 \quad 0\,).$$

$$\therefore B^{*-1} = \begin{pmatrix} B^{-1} & 0 \\ c_B B^{-1} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The net evaluations $z_j - c_j$ corresponding to non-basic variables are

$$\begin{aligned} &(\,z_1 - c_1 \quad z_2 - c_2 \quad z_3 - c_3 \quad z_4 - c_4 \quad z_6 - c_6\,) \\ &= (\,0 \quad 0 \quad 0 \quad 0 \quad 1\,)[\,a_1^* \quad a_2^* \quad a_3^* \quad a_4^* \quad a_6^*\,] \\ &\left[\because A^* = \begin{pmatrix} 3 & 2 & 3 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -2 & 1 & -2 & 1 & 0 & -1 & 1 & 0 & 0 & 0 \\ 2 & 2 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & -2 & -3 & -4 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -3 & 1 & -2 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}\right] \\ &= (\,0 \quad -3 \quad 1 \quad -2 \quad 1\,). \end{aligned}$$

This indicates that $z_2 - c_2$ is the most negative. Hence, $y_2^*$ enters the basis.

$$\begin{aligned} &\text{Now,} \quad y_2^* = B^{*-1} a_2^* = (\,2 \quad 1 \quad 2 \quad -2 \quad -3\,)^{\mathrm{T}}. \\ &\text{Also,} \quad x_B^* = B^{*-1} b^* = (\,25 \quad 5 \quad 20 \quad 0 \quad -25\,)^{\mathrm{T}}. \\ &\therefore x_B = (\,25 \quad 5 \quad 20 \quad 0\,)^{\mathrm{T}} \text{ and } z_B^* = -25. \end{aligned}$$

Now we construct the initial revised simplex table.

| $y_B^*$ | $x_B^*$ | $B^{*-1}$ | | | | | $y_2^*$ | Mini ratio |
|---|---|---|---|---|---|---|---|---|
| $y_5^*$ | 25 | 1 | 0 | 0 | 0 | 0 | 2 | 25/2 |
| $y_7^*$ | 5 | 0 | 1 | 0 | 0 | 0 | 1 | 5 |
| $y_8^*$ | 20 | 0 | 0 | 1 | 0 | 0 | 2 | 10 |
| $z$ | 0 | 0 | 0 | 0 | 1 | 0 | −2 | – |
| $z^*$ | −25 | 0 | 0 | 0 | 0 | 1 | −3 | – |

According to the minimum ratio, $y_7^*$ leaves the basis and $y_{22}^* = 1$ is the key element.

**First iteration**:

We introduce $y_2^*$ and drop $y_7^*$. In this iteration, the reduced basis inverse is given by

$$B^{*-1} = \begin{pmatrix} 1 & -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 & 1 \end{pmatrix} \quad \text{by} \quad \begin{matrix} R_1' = R_1 - 2R_2' \\ R_3' = R_3 - 2R_2' \\ R_4' = R_4 + 2R_2' \\ R_5' = R_5 + 3R_2' \end{matrix}.$$

Again, the net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned} (z_1 - c_1 \quad z_3 - c_3 \quad z_4 - c_4 \quad z_6 - c_6) &= (0 \quad 3 \quad 0 \quad 0 \quad 1)[a_1^* \quad a_3^* \quad a_4^* \quad a_6^*] \\ &= (0 \quad 3 \quad 0 \quad 0 \quad 1)\begin{pmatrix} 3 & 3 & -1 & 0 \\ -2 & -2 & 1 & -1 \\ 2 & 1 & 1 & 0 \\ -1 & -3 & -4 & 0 \\ 0 & 1 & -2 & 1 \end{pmatrix} = (-6 \quad -5 \quad 1 \quad -2). \end{aligned}$$

This indicates that $z_1 - c_1$ is the most negative. Hence, $y_1^*$ enters the basis.

$$\therefore y_1^* = B^{*-1}a_1^* = \begin{pmatrix} 1 & -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \\ 2 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 7 \\ -2 \\ 6 \\ -5 \\ -6 \end{pmatrix}$$

$$\text{and } x_B^* = B^{*-1}b^* = \begin{pmatrix} 1 & -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 25 \\ 5 \\ 20 \\ 0 \\ -25 \end{pmatrix} = \begin{pmatrix} 15 \\ 5 \\ 10 \\ 10 \\ -10 \end{pmatrix}.$$

Hence, the next revised simplex table is as follows:

| $y_B^*$ | $x_B^*$ | $B^{*-1}$ | | | | | $y_1^*$ | Mini ratio |
|---|---|---|---|---|---|---|---|---|
| $y_5^*$ | 15 | 1 | −2 | 0 | 0 | 0 | 7 | 15/7 |
| $y_2^*$ | 5 | 0 | 1 | 0 | 0 | 0 | −2 | – |

(continued)

(continued)

| $y_B^*$ | $x_B^*$ | $B^{*-1}$ | | | | | $y_1^*$ | Mini ratio |
|---|---|---|---|---|---|---|---|---|
| $y_8^*$ | 10 | 0 | −2 | 1 | 0 | 0 | 6 | 5/3 |
| $z$ | 10 | 0 | 2 | 0 | 1 | 0 | −5 | – |
| $z^*$ | −10 | 0 | 3 | 0 | 0 | 1 | −6 | – |

According to the minimum ratio, $y_8^*$ leaves the basis and $y_{13}^* = 6$ is the key element.

**Second iteration**:

We introduce $y_1^*$ and drop $y_8^*$.

In this iteration, the reduced basis inverse is given by

$$B^{*-1} = \begin{pmatrix} 1 & 1/3 & -7/6 & 0 & 0 \\ 0 & 1/3 & 1/3 & 0 & 0 \\ 0 & -1/3 & 1/6 & 0 & 0 \\ 0 & 1/3 & 5/6 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

Now, the net evaluations $z_j - c_j$ corresponding to non-basic variables are

$$\begin{aligned} (\, z_3 - c_3 \quad z_4 - c_4 \quad z_6 - c_6 \,) &= (\, 0 \quad 1 \quad 1 \quad 0 \quad 1 \,)(\, a_3^* \quad a_4^* \quad a_6^* \,) \\ &= (\, 0 \quad 1 \quad 1 \quad 0 \quad 1 \,)\begin{pmatrix} 3 & -1 & 0 \\ -2 & 1 & -1 \\ 1 & 1 & 0 \\ -3 & -4 & 0 \\ 1 & -2 & 1 \end{pmatrix} = (\, 0 \quad 0 \quad 0 \,). \end{aligned}$$

Since all $z_j - c_j$ corresponding to non-basic variables are zero, the current solution is an optimum one for Phase I. Therefore, the optimum solution for this phase is

$$x_B^* = B^{*-1}b^* = \begin{pmatrix} 1 & 1/3 & -7/6 & 0 & 0 \\ 0 & 1/3 & 1/3 & 0 & 0 \\ 0 & -1/3 & 1/6 & 0 & 0 \\ 0 & 1/3 & 5/6 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 25 \\ 5 \\ 20 \\ 0 \\ -25 \end{pmatrix} = \begin{pmatrix} 10/3 \\ 25/3 \\ 5/3 \\ 55/3 \\ 0 \end{pmatrix}.$$

$$\therefore x_1 = \frac{5}{3}, x_2 = \frac{25}{3}, x_3 = 0, x_4 = 0 \text{ and the Max } z^* = 0.$$

Since there is no artificial variable in the current basis, we shall apply Phase II.

**Phase II**:

From the basis inverse $B^{*-1}$ of the optimal solution obtained in Phase I, we have

$$\widehat{B}^{-1} = \begin{pmatrix} 1 & 1/3 & -7/6 & 0 \\ 0 & 1/3 & 1/3 & 0 \\ 0 & -1/3 & 1/6 & 0 \\ 0 & 1/3 & 5/6 & 1 \end{pmatrix} \text{ by deleting the fifth row and the fifth column.}$$

$$\text{Here} \quad \widehat{A} = \begin{pmatrix} 3 & 2 & 3 & -1 & 1 & 0 \\ -2 & 1 & -2 & 1 & 0 & -1 \\ 2 & 2 & 1 & 1 & 0 & 0 \\ -1 & -2 & -3 & -4 & 0 & 0 \end{pmatrix}.$$

Now the net evaluations $z_j - c_j$ corresponding to non-basic variables are

$$\begin{aligned} &\begin{pmatrix} z_3 - c_3 & z_4 - c_4 & z_6 - c_6 \end{pmatrix} \\ &= \begin{pmatrix} 0 & \frac{1}{3} & \frac{5}{6} & 1 \end{pmatrix}\begin{pmatrix} \widehat{a}_3 & \widehat{a}_4 & \widehat{a}_6 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{3} & \frac{5}{6} & 1 \end{pmatrix}\begin{pmatrix} 3 & -1 & 0 \\ -2 & 1 & -1 \\ 1 & 1 & 0 \\ -3 & -4 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{17}{6} & -\frac{17}{6} & -\frac{1}{3} \end{pmatrix}. \end{aligned}$$

This result indicates that either $\widehat{y}_3$ or $\widehat{y}_4$ can enter the basis. Arbitrarily, we choose $\widehat{y}_3$ as the entering vector.

$$\text{Now, } \widehat{y}_3 = \widehat{B}^{-1}\widehat{a}_3 = \begin{pmatrix} 1 & 1/3 & -7/6 & 0 \\ 0 & 1/3 & 1/3 & 0 \\ 0 & -1/3 & 1/6 & 0 \\ 0 & 1/3 & 5/6 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \\ 1 \\ -3 \end{pmatrix} = \begin{pmatrix} 7/6 \\ -1/3 \\ 5/6 \\ -17/6 \end{pmatrix}$$

$$\text{and} \quad \widehat{x}_B = \widehat{B}^{-1}\widehat{b} = \begin{pmatrix} 10/3 & 25/3 & 5/3 & 55/3 \end{pmatrix}^{\mathrm{T}} \therefore z_B = \frac{55}{3}.$$

Hence, the next revised simplex table is as follows:

| $\widehat{y}_B$ | $\widehat{x}_B$ | $\widehat{B}^{-1}$ | | | | $\widehat{y}_3$ | Mini ratio |
|---|---|---|---|---|---|---|---|
| $\widehat{y}_5$ | 10/3 | 1 | 1/3 | −7/6 | 0 | 7/6 | 20/7 |
| $\widehat{y}_2$ | 25/3 | 0 | 1/3 | 1/3 | 0 | −1/3 | – |
| $\widehat{y}_1$ | 5/3 | 0 | −1/3 | 1/6 | 0 | $5/6^*$ | 2 |
| $z$ | 55/3 | 0 | 1/3 | 5/6 | 1 | −17/6 | – |

According to the minimum ratio, $\widehat{y}_1$ leaves the basis and $\widehat{y}_{33} = 5/6$ is the key element.

**First iteration**:

In this iteration, the reduced basis matrix is given by

$$\widehat{B}^{-1} = \begin{pmatrix} 1 & 4/5 & -7/5 & 0 \\ 0 & 1/5 & 2/5 & 0 \\ 0 & -2/5 & 1/5 & 0 \\ 0 & -4/5 & 7/5 & 1 \end{pmatrix}.$$

Now, the net evaluations $z_j - c_j$ corresponding to non-basic variables are

$$\begin{aligned} (z_1 - c_1 \quad z_4 - c_4 \quad z_6 - c_6) &= \begin{pmatrix} 0 & -\frac{4}{5} & \frac{7}{5} & 1 \end{pmatrix} (\widehat{a}_1 \quad \widehat{a}_4 \quad \widehat{a}_6) \\ &= \begin{pmatrix} 0 & -\frac{4}{5} & \frac{7}{5} & 1 \end{pmatrix} \begin{pmatrix} 3 & -1 & 0 \\ -2 & 1 & -1 \\ 2 & 1 & 0 \\ -1 & -4 & 0 \end{pmatrix} = \begin{pmatrix} \frac{17}{5} & -\frac{17}{5} & \frac{4}{5} \end{pmatrix}. \end{aligned}$$

This indicates that $\widehat{y}_4$ enters the basis.

$$\therefore \widehat{y}_4 = \widehat{B}^{-1}\widehat{a}_4 = \begin{pmatrix} 1 & 4/5 & -7/5 & 0 \\ 0 & 1/5 & 2/5 & 0 \\ 0 & -2/5 & 1/5 & 0 \\ 0 & -4/5 & 7/5 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 1 \\ -4 \end{pmatrix} = \begin{pmatrix} -8/5 \\ 3/5 \\ -1/5 \\ -17/5 \end{pmatrix}$$

$$\text{and } \widehat{x}_B = \widehat{B}^{-1}\widehat{b} = \begin{pmatrix} 1 & 4/5 & -7/5 & 0 \\ 0 & 1/5 & 2/5 & 0 \\ 0 & -2/5 & 1/5 & 0 \\ 0 & -4/5 & 7/5 & 1 \end{pmatrix} \begin{pmatrix} 25 \\ 5 \\ 20 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 9 \\ 2 \\ 24 \end{pmatrix}.$$

$$\therefore z_B = (1 \quad 9 \quad 2)^{\mathrm{T}}, \;\; z_B = 24.$$

Hence, the next revised simplex table is as follows:

| $\widehat{y}_B$ | $\widehat{x}_B$ | $\widehat{B}^{-1}$ | | | | $\widehat{y}_4$ | Mini ratio |
|---|---|---|---|---|---|---|---|
| $\widehat{y}_5$ | 1 | 1 | 4/5 | −7/5 | 0 | − 8/5 | – |
| $\widehat{y}_2$ | 9 | 0 | 1/5 | 2/5 | 0 | 3/5* | 15 |
| $\widehat{y}_3$ | 2 | 0 | −2/5 | 1/5 | 0 | −1/5 | – |
| $z$ | 24 | 0 | −4/5 | 7/5 | 1 | −17/5 | |

According to the minimum ratio, $\widehat{y}_2$ leaves the basis and $\widehat{y}_{24} = 3/5$ is the key element.

**Second iteration**:

In this iteration, the reduced basis matrix is given by

$$\widehat{B}^{-1} = \begin{pmatrix} 1 & 4/3 & -1/3 & 0 \\ 0 & 1/3 & 2/3 & 0 \\ 0 & -1/3 & 1/3 & 0 \\ 0 & 1/3 & 11/3 & 1 \end{pmatrix}.$$

Now the net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned} &( z_1 - c_1 \quad z_2 - c_2 \quad z_6 - c_6 ) \\ &= \begin{pmatrix} 0 & \frac{1}{3} & \frac{11}{3} & 1 \end{pmatrix} ( \widehat{a}_1 \quad \widehat{a}_2 \quad \widehat{a}_6 ) \\ &= \begin{pmatrix} 0 & \frac{1}{3} & \frac{11}{3} & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 & 0 \\ -2 & 1 & -1 \\ 2 & 2 & 0 \\ -1 & -2 & 0 \end{pmatrix} = \begin{pmatrix} \frac{17}{3} & \frac{17}{3} & -\frac{1}{3} \end{pmatrix}. \end{aligned}$$

This indicates that $\widehat{y}_6$ enters the basis.

$$\therefore \widehat{y}_6 = \widehat{B}^{-1}\widehat{a}_6 = \begin{pmatrix} 1 & 4/3 & -1/3 & 0 \\ 0 & 1/3 & 2/3 & 0 \\ 0 & -1/3 & 1/3 & 0 \\ 0 & 1/3 & 11/3 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -4/3 \\ -1/3 \\ 1/3 \\ -1/3 \end{pmatrix}$$

$$\text{and } \widehat{x}_B = \widehat{B}^{-1}\widehat{B} = \begin{pmatrix} 1 & 4/3 & -1/3 & 0 \\ 0 & 1/3 & 2/3 & 0 \\ 0 & -1/3 & 1/3 & 0 \\ 0 & 1/3 & 11/3 & 1 \end{pmatrix} \begin{pmatrix} 25 \\ 5 \\ 20 \\ 0 \end{pmatrix} = \begin{pmatrix} 25 \\ 15 \\ 5 \\ 75 \end{pmatrix}.$$

The next revised simplex table is as follows:

| $\widehat{y}_B$ | $\widehat{x}_B$ | $\widehat{B}^{-1}$ | | | | $\widehat{y}_6$ | Mini ratio |
|---|---|---|---|---|---|---|---|
| $\widehat{y}_5$ | 25 | 1 | 4/3 | −1/3 | 0 | −4/3 | – |
| $\widehat{y}_4$ | 15 | 0 | 1/3 | 2/3 | 0 | −1/3 | – |
| $\widehat{y}_3$ | 5 | 0 | −1/3 | 1/3 | 0 | 1/3* | 15 |
| $z$ | 75 | 0 | 1/3 | 11/3 | 1 | −1/3 | |

According to the minimum ratio, $\widehat{y}_3$ leaves the basis and $\widehat{y}_{36} = \frac{1}{3}$ is the key element.

**Third iteration**:

In this iteration, the reduced basis matrix is given by

$$\widehat{B}^{-1} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 4 & 1 \end{pmatrix}.$$

Now the net evaluations $z_j - c_j$ corresponding to non-basic variables are given by

$$\begin{aligned} (\, z_1 - c_1 \quad z_2 - c_2 \quad z_3 - c_3 \,) &= (\, 0 \quad 0 \quad 4 \quad 1 \,)(\, \widehat{a}_1 \quad \widehat{a}_2 \quad \widehat{a}_3 \,) \\ &= (\, 0 \quad 0 \quad 4 \quad 1 \,) \begin{pmatrix} 3 & 2 & 3 \\ -2 & 1 & -2 \\ 2 & 2 & 1 \\ -1 & -2 & -3 \end{pmatrix} = (\, 7 \quad 6 \quad 1 \,). \end{aligned}$$

Since the net evaluations $z_j - c_j$ corresponding to non-basic variables are greater than zero, an optimum basic feasible solution has been obtained.

The optimal solution is given by

$$\widehat{x}_B = \widehat{B}^{-1}\widehat{b} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 4 & 1 \end{pmatrix} \begin{pmatrix} 25 \\ 5 \\ 20 \\ 0 \end{pmatrix} = \begin{pmatrix} 45 \\ 20 \\ 15 \\ 80 \end{pmatrix}.$$

$$\therefore \widehat{x}_B = \begin{pmatrix} 45 \\ 20 \\ 15 \end{pmatrix} \quad \text{and} \quad z_B = 80,$$

i.e. $x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 20$ and $z_{\text{Max}} = 80$.

## 4.7 Advantages of Revised Simplex Method

The revised simplex method has the following advantages:

(i) This method automatically generates the inverse of the current basis matrix and the new basic feasible solution as well.
(ii) It provides more information with less computational effort.
(iii) It requires fewer computations than that of the ordinary simplex method.
(iv) Fewer entries need to be made in each table of the revised simplex method.

## 4.8 Difference Between Revised Simplex Method and Simplex Method

Although the simplex method is a straightforward algebraic procedure, it is not an efficient computational procedure for solving LPPs on a digital computer. The ordinary simplex method requires the storing of all the elements of the entire table in the memory of the computer, which may not be feasible for a very large problem. Also, at each iteration the ordinary simplex method requires many computations that are not needed in the subsequent iteration. In fact, it is not necessary to compute the entire table during each iteration. The only pieces of information needed at each iteration are: (i) net evaluations $z_j - c_j$, (ii) entering vector, (iii) current basic variables and their values. In the revised simplex method the preceding information is directly obtained very efficiently, avoiding unnecessary calculations as compared to the ordinary simplex method. Thus, the revised simplex method computes and stores only those pieces of information which are needed in the current iteration.

## 4.9 Exercises

Solve the following LPPs using the revised simplex method:

1. Maximize $z = x_1 + x_2$
   such that $x_1 + 2x_2 \le 2$
   $4x_1 + x_2 \le 4$ and $x_1,\ x_2 \ge 0.$
2. Maximize $z = 5x_1 + 3x_2$
   such that $3x_1 + 5x_2 \le 15$
   $3x_1 + 2x_2 \le 10$ and $x_1,\ x_2 \ge 0.$
3. Maximize $z = x_1 + x_2 + 3x_3$
   such that $3x_1 + 2x_2 + x_3 \le 3$
   $2x_1 + x_2 + 2x_3 \le 2$ and $x_1,\ x_2,\ x_3 \ge 0.$
4. Maximize $z = x_1 + x_2$
   such that $3x_1 + 3x_2 \le 6$
   $x_1 + 4x_2 \le 4$ and $x_1,\ x_2 \ge 0.$
5. Maximize $z = 3x_1 + 2x_2 + 5x_3$
   such that $x_1 + 2x_2 + x_3 \le 430$
   $3x_1 + 2x_3 \le 460$
   $x_1 + 4x_2 \le 420$ and $x_1,\ x_2,\ x_3 \ge 0.$
6. Maximize $z = x_1 + 2x_2$
   such that $x_1 + 2x_2 \le 3$
   $x_1 + 3x_2 \le 1$ and $x_1,\ x_2 \ge 0.$
7. Maximize $z = x_1 + 2x_2$
   such that $x_1 + x_2 \le 3$
   $x_1 + 2x_2 \le 5$
   $3x_1 + x_2 \le 6$ and $x_1,\ x_2 \ge 0.$

8. Maximize $z = 5x_1 + 3x_2$
such that $4x_1 + 5x_2 \geq 10$
$5x_1 + 2x_2 \leq 10$
$3x_1 + 8x_2 \leq 12$ and $x_1,\ x_2 \geq 0.$
9. Minimize $z = 3x_1 + x_2$
such that $x_1 + x_2 \geq 1$
$2x_1 + x_2 \geq 0$ and $x_1,\ x_2 \geq 0.$
10. Minimize $z = 4x_1 + 2x_2 + 3x_3$
such that $2x_1 + 4x_3 \geq 5$
$2x_1 + 4x_2 + x_3 \geq 4$ and $x_1,\ x_2,\ x_3 \geq 0.$
11. Maximize $z = 4x_1 + 3x_2$
such that $3x_1 + 4x_2 \leq 12$
$3x_1 + 3x_2 \leq 10$
$2x_1 + x_2 \leq 4$ $x_1,\ x_2 \geq 0.$
12. Maximize $z = 2x_1 + 3x_2$
such that $-x_1 + x_2 \geq 0$
$x_1 \leq 4$ and $x_1,\ x_2 \geq 0.$
13. Minimize $z = 2x_1 + x_2$
such that $3x_1 + x_2 \leq 3$
$4x_1 + 3x_2 \geq 6$
$x_1 + 2x_2 \leq 3$ and $x_1,\ x_2 \geq 0.$
14. Maximize $z = x_1 + x_2$
such that $2x_1 + 5x_2 \leq 6$
$x_1 + x_2 \geq 2$ and $x_1,\ x_2 \geq 0.$

# Chapter 5
# Dual Simplex Method

## 5.1 Objective

The objective of this chapter is to discuss an advanced technique, called the dual simplex method, for solving linear programming problems with $\geq$ type constraints.

## 5.2 Introduction

In Chap. 3, the simplex method was discussed for solving linear programming problems (LPPs). In this method, a negative net evaluation (for a maximization primal problem) indicates that the current solution is not optimum.

Consider the $j$th dual constraint of a primal maximization problem, namely

$$\sum_{i=1}^{m} a_{ij}w_i \geq c_j,\ j = 1, 2, \ldots, n$$

where $a_{ij}$ is the constraint coefficient of the $j$th primal variable $x_j$ in the $i$th primal constraint, and $c_j$ is the profit/cost associated with $x_j$ in the primal objective function. Then clearly the net evaluation $z_j - c_j$ corresponding to $x_j$ is $\sum_{i=1}^{m} a_{ij}w_i - c_j$. Thus, a negative value of $z_j - c_j$ indicates that the corresponding dual constraint is not satisfied. This leads to the conclusion that the dual is infeasible when the primal is non-optimum and vice versa. This gives the following result.

While the primal problem starts as feasible but non-optimum and continues to be feasible until the optimum is reached, the dual problem starts as infeasible but better than optimum and continues to be infeasible until the true optimal solution is reached. In other words, while the primal problem is seeking optimality, the dual problem is automatically seeking feasibility.

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_5

Hence, the solution will start optimum (or actually better than optimum) and infeasible and will remain infeasible until the true optimum is reached, at which the solution becomes feasible. Such a procedure is called the dual simplex method.

The dual simplex method is very similar to the ordinary simplex method. In these methods, the only difference is in the criterion used for selecting the entering and outgoing vectors. In this context, we note that in the dual simplex method the outgoing vector is determined first and then the entering vector. This is just the reverse of what is done in the ordinary simplex method.

The dual simplex method gives an algorithm in which we start with a basic optimal solution of the primal problem with all $z_j - c_j \geq 0$, but it is not a feasible solution, as the values of some basic variables are negative. In each iteration, the numbers of negative basic variables are decreased while maintaining the optimality. An optimal solution is reached in a finite number of steps.

In this method, inclusions of artificial variables are not required. Hence, this method greatly reduces the amount of computation.

The dual simplex method has a wide range of applications. Some are as follows:

(i) Parametric programming problem
(ii) Integer linear programming problem
(iii) Post-optimality analysis when the requirement vector is changed or when new constraints are added.

**Theorem** *For a basic solution (not necessarily feasible) with all* $z_j - c_j \geq 0\,\forall j$, *the value of the objective function corresponding to this basic solution is optimal or better than optimal.*

*Proof* Let the primal LPP be as follows:

$$\begin{aligned}&\text{Maximize } z = \langle c^{\mathrm{T}}, x\rangle\\&\text{subject to} Ax = b, x \geq 0,\\&\text{where } c, x^{\mathrm{T}} \in \mathbb{R}^n, b^{\mathrm{T}} \in \mathbb{R}^m \text{ and } A \text{ is an m} \times \text{n real matrix.}\end{aligned}$$

Let $B$ be the current basis of the primal LPP and $x_B$ be the corresponding basic solution.

$$\therefore Bx_B = b \quad \text{or} \quad x_B = B^{-1}b.$$

Let $z_j - c_j \geq 0, \;\; \forall j = 1, 2, \ldots, n$ and one or more basic variables has negative value. The dual of the preceding LPP is given by the following:

$$\begin{aligned}&\text{Minimize} \quad z' = \langle b^{\mathrm{T}}, w\rangle\\&\text{subject to} \quad A^{\mathrm{T}}w \geq c^{\mathrm{T}}, w \text{ is unrestricted.}\end{aligned}$$

The value of $z$ corresponding to the basic solution $x_B$ is given by

$$
\begin{aligned}
(z_B) &= c_B x_B = [c_B x_B]^{\mathrm{T}} \quad [\because c_B x_B \text{ is (1,1) matrix as } c_B \text{ is } 1\times m \text{ and } x_B \text{ is } m\times 1 \text{ matrix}]\\
&= x_B^{\mathrm{T}} c_B^{\mathrm{T}}\\
&= \left(B^{-1}b\right)^{\mathrm{T}} c_B^{\mathrm{T}} \qquad \left[\because x_B = B^{-1}b\right]\\
&= b^{\mathrm{T}}\left(B^{-1}\right)^{\mathrm{T}} c_B^{\mathrm{T}} = b^{\mathrm{T}}\left(c_B B^{-1}\right)^{\mathrm{T}}\\
&= b^{\mathrm{T}} w^*, \text{where } w^* = \left(c_B B^{-1}\right)^{\mathrm{T}}.
\end{aligned}
$$

Therefore $z_B = \langle b^{\mathrm{T}}, w^* \rangle$.

It should be noted that $w^* \in \mathbb{R}^m$.

Since all $z_j - c_j \geq 0$, we have

$$
\begin{aligned}
&\begin{pmatrix} c_B B^{-1} & 1 \end{pmatrix}\begin{pmatrix} A \\ -c \end{pmatrix} \geq 0\\
&\text{or } \left(c_B B^{-1}\right)A - c \geq 0\\
&\text{or } \left(\left(c_B B^{-1}\right)A\right)^{\mathrm{T}} \geq c^{\mathrm{T}}, \text{or } A^{\mathrm{T}}\left(c_B B^{-1}\right)^{\mathrm{T}} \geq c^{\mathrm{T}} \text{ or } A^{\mathrm{T}} w' \geq c^{\mathrm{T}}
\end{aligned}
$$

This implies that $w'$ is a basic feasible solution of the dual LPP.

Now $\langle b^{\mathrm{T}}, w' \rangle$ is the value of the dual objective function for the solution $w'$.

$$z'_{\mathrm{Min}} \leq \left\langle b^{\mathrm{T}}, w' \right\rangle.$$

But we have $\langle b^{\mathrm{T}}, w' \rangle = z_B$.

Therefore, $z'_{\mathrm{Min}} \leq z_B$.

According to the duality theorem, we have $\therefore z'_{\mathrm{Min}} \leq z_{\mathrm{Max}}$.

Hence, $z_{\mathrm{Max}} \leq z_B$ or $z_B \geq z_{\mathrm{Max}}$,

i.e. the value of the objective function corresponding to the basic solution $x_B = B^{-1}b$ is equal to or greater than the optimal value of the objective function.

This proves the theorem.

**Theorem** *Let* $x_{B_r}$ *be a negative basic variable in a dual simplex table, let all net evaluations* $z_j - c_j$ *be non-negative and let the primal LPP be of maximization type. If* $y_{rj} \geq 0$ *for all non-basic variables* $x_j$*, then there does not exist any feasible solution to the primal LPP.*

*Proof* Let the primal LPP be as follows:

$$
\begin{aligned}
&\text{Maximize } z = \left\langle c^{\mathrm{T}}, x \right\rangle\\
&\text{subject to } Ax = b, x \geq 0,\\
&\text{where } c, x^{\mathrm{T}} \in \mathbb{R}^n, b^{\mathrm{T}} \in \mathbb{R}^m \text{ and } A \text{ is an m} \times \text{n real matrix.}
\end{aligned}
$$

Let $x_B = B^{-1}b$ be a basic solution of this primal LPP.

The dual of this LPP is given by

$$\begin{aligned}&\text{Minimize } z' = \langle b^{\mathrm{T}}, w\rangle\\ &\text{subject to } A^{\mathrm{T}}w \geq c^{\mathrm{T}},\\ &\text{where } w \text{ is unrestricted and } w^{\mathrm{T}} \in \mathbb{R}^m.\end{aligned}$$

Since all $z_j - c_j \geq 0$, we have

$$\begin{pmatrix} c_B B^{-1} & 1 \end{pmatrix} \begin{pmatrix} A \\ -c \end{pmatrix} \geq 0$$

or $A^{\mathrm{T}}w_0 \geq c^{\mathrm{T}}$, where $w_0 = (c_B B^{-1})^{\mathrm{T}}$.

Hence, $w_0$ is a feasible solution of the dual LPP.

Let $\eta$ be any positive real number and $w_1 = w_0 + \eta\beta_r$, where $\beta_r$ is the $r$th column of $(B^{-1})^{\mathrm{T}}$.

$$\text{Now} \quad A^{\mathrm{T}}w_1 = A^{\mathrm{T}}(w_0 + \eta\beta_r) = A^{\mathrm{T}}w_0 + \eta A^{\mathrm{T}}\beta_r. \tag{5.1}$$

We know that $y_j = B^{-1}a_j$ for all $j$.

Hence, $Y = B^{-1}A$ or $Y^{\mathrm{T}} = (B^{-1}A)^{\mathrm{T}} = A^{\mathrm{T}}(B^{-1})^{\mathrm{T}}$

$$\therefore Y^{\mathrm{T}} = A^{\mathrm{T}}[\beta_1\, \beta_2\, \ldots\, \beta_m] = \left[A^{\mathrm{T}}\beta_1\, A^{\mathrm{T}}\beta_2\, \ldots\, A^{\mathrm{T}}\beta_m\right]$$

$$\text{or} \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & & & \\ \vdots & & & \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} = \left[A^{\mathrm{T}}\beta_1\, A^{\mathrm{T}}\beta_2\, \ldots\, A^{\mathrm{T}}\beta_m\right]$$

$$\text{or} \begin{bmatrix} y_{11} & y_{21} & \cdots & y_{m1} \\ y_{12} & y_{22} & \cdots & y_{m2} \\ \vdots & & & \\ \vdots & & & \\ y_{1n} & y_{2n} & \cdots & y_{mn} \end{bmatrix} = \left[A^{\mathrm{T}}\beta_1\, A^{\mathrm{T}}\beta_2\, \ldots\, A^{\mathrm{T}}\beta_m\right]$$

$$\therefore \begin{bmatrix} y_{r1} \\ y_{r2} \\ \vdots \\ \vdots \\ y_{rn} \end{bmatrix} = A^{\mathrm{T}}\beta_r \text{ for } r = 1, 2, \ldots, m$$

Since all $y_{rj} \geq 0$, we have $A^T \beta_r \geq 0$.

As $\eta \geq 0$ we have $\eta A^T \beta_r \geq 0$.

Hence from (5.1), we have

$$\begin{aligned} &A^{\mathrm{T}} w_0 \geq A^{\mathrm{T}} w_1 \\ &\text{or } A^{\mathrm{T}} w_0 \geq c^{\mathrm{T}} \text{as } A^{\mathrm{T}} w_1 \geq c^{\mathrm{T}}. \end{aligned}$$

Hence, $w_0$ is a feasible solution to the dual LPP for all $\alpha$.

The value of the dual objective function $z'$ for the feasible solution $w_0$ is given by

$$z'|_{w_0} = \langle b^{\mathrm{T}}, w_0 \rangle = \langle b^{\mathrm{T}}, w_1 + \eta \beta_r \rangle = \langle b^{\mathrm{T}}, w_1 \rangle + \eta \langle b^{\mathrm{T}}, \beta_r \rangle. \tag{5.2}$$

As, $x_B = B^{-1} b, x_B^{\mathrm{T}} = (B^{-1} b)^{\mathrm{T}} \text{or } x_B^{\mathrm{T}} = b^{\mathrm{T}} (B^{-1})^{\mathrm{T}}$
or $[\, x_{B_1} \quad x_{B_2} \quad \ldots \quad x_{B_r} \quad \ldots \quad x_{B_m} \,] = b^{\mathrm{T}} [\, \beta_1 \quad \beta_2 \quad \ldots \quad \beta_r \quad \ldots \quad \beta_m \,]$.

Hence, $x_{B_r} = \langle b^{\mathrm{T}}, \beta_r \rangle$for all $r$.

Thus, from (5.2), $z'|_{w_0} = \langle b^{\mathrm{T}}, w_1 \rangle + \eta x_{B_r}$.

Letting $\eta \to \infty$, we have $z'|_{w_0} \to -\infty$ [$\because x_{B_r} < 0$ and $\eta > 0$].

This means that the dual LPP has an unbounded solution. Therefore, the primal LPP has no feasible solution.

**Criteria for incoming and outgoing vectors in dual simplex method**

In the simplex method, if the basic variable $x_{B_r}$ is replaced by the non-basic variable $x_k$, then we have

$$\hat{z} = z - \frac{x_{Br}}{y_{rk}} (z_k - c_k) \tag{5.3}$$

$$\text{and } \hat{z}_j - \hat{c}_j = z_j - c_j - \frac{y_{rj}}{y_{rk}} (z_k - c_k). \tag{5.4}$$

Now to remove negative basic variables, first the most negative basic variable $x_{B_r}$ is selected as the outgoing basic variable from all basic variables. So the corresponding vector $y_{B_r}$ is the outgoing vector.

We know that if for some basic solution (not necessarily feasible) all components of net evaluations are non-negative, the value of the objective function to this basic solution is either optimal or better than optimal. So, we have to decrease the value of the objective function to obtain the value of $z_{\mathrm{Max}}$.

Since $x_{B_r} < 0$ and $z_k - c_k \geq 0$, in order to have $\hat{z} \leq z$, we must choose an entering variable $x_k$ such that $y_{rk} < 0$. In this case, the new values of the net evaluations are given by

$$\hat{z}_j - \hat{c}_j = z_j - c_j - \frac{y_{rj}}{y_{rk}}(z_k - c_k).$$

$$\text{Again,} \quad \hat{z}_j - \hat{c}_j \geq 0 \ \forall j.$$

Hence, we can write

$$\begin{aligned} & z_j - c_j \geq \frac{y_{rj}}{y_{rk}}(z_k - c_k), \\ \text{i.e. } & \frac{z_j - c_j}{y_{rj}} \leq \frac{z_k - c_k}{y_{rk}} \quad \text{for} \quad y_{rj} < 0 \\ \text{or } & \frac{z_k - c_k}{y_{rk}} = \operatorname*{Max}_j \left\{ \frac{z_j - c_j}{y_{rj}}; \ y_{rj} < 0 \right\}. \end{aligned}$$

From this criterion, the incoming vector and key element can be found easily. Here the incoming vector is $y_k$ and $y_{rk}$ is the key element.

## 5.3 Dual Simplex Algorithm

The iterative procedures for the dual simplex method are as follows:

Step 1 If the given LPP is in the minimization form, then convert it into a maximization problem using the mini-max principle.

Step 2 Convert $\geq$ type constraints of the given LPP, if any, into $\leq$ type, multiplying both sides of the corresponding constraints by $(-1)$.

Step 3 Introduce slack variables in the constraints of the given LPP and obtain an initial basic solution. Put this solution in the starting dual simplex table.

Step 4 Test the nature of the net evaluations $z_j - c_j$ in the starting simplex table.

- (i) If all $z_j - c_j$ and $x_{B_j}$ are non-negative for all $i$ and $j$, then an optimum basic feasible solution has been obtained.
- (ii) If all $z_j - c_j$ are non-negative and at least one basic variable, say $x_{B_i}$, is negative, then go to Step 5.
- (iii) If at least one $z_j - c_j$ is negative, then the dual simplex method is not applicable for the given problem.

Step 5 Select the most negative of $x_{B_i}$, say $x_{B_r}$. The corresponding basis vector $y_{B_r}$ leaves the basis.

Step 6 Test the nature of all $y_{rj}, j = 1, 2, \ldots, n$, i.e. the elements of the $r$th row.

- (i) If all $y_{rj}$ are non-negative, there does not exist any feasible solution to the given LPP.

(ii) If at least one $y_{rj}$ is negative, then compute

$$\operatorname*{Max}_{j}\left\{\frac{z_j - c_j}{y_{rj}} : y_{rj} < 0\right\}, \quad j = 1, 2, \ldots, n,$$

If this is $\frac{z_k - c_k}{y_{rk}}$, then $y_k$ enters the basis in place of $y_{B_r}$ and $y_{rk}$ is the key element.

Step 7 With $y_{rk}$ as the key element, form the next table. Using elementary row operations, convert the key element to unity and all other elements of the key column to zero to get the improved solution.

Step 8 Repeat Steps 4 through 7 until either an optimum basic feasible solution is obtained or there is an indication of no feasible solution.

*Example 1* Using the dual simplex method, solve the following LPP:

$$\begin{aligned} \text{Minimize} \quad & z = x_1 + x_2 \\ \text{subject to} \quad & \\ & 2x_1 + x_2 \geq 2 \\ & -x_1 - x_2 \geq 1 \text{ and } x_1, x_2 \geq 0. \end{aligned}$$

**Solution**

The given LPP can be written as follows:

$$\begin{aligned} \text{Minimize} \quad & z = x_1 + x_2 \\ \text{subject to} \quad & \\ & -2x_1 - x_2 \leq -2 \\ & x_1 + x_2 \leq -1 \text{ and } x_1, x_2 \geq 0. \end{aligned}$$

Here the problem is a minimization problem.

Let $z' = -z$. Then $\operatorname{Min} z = -\operatorname{Max}(-z) = -\operatorname{Max} z'$ using the mini-max principle.

Hence, the problem is to maximize $z' = -x_1 - x_2$.

Now introducing two slack variables $x_3$ and $x_4$ in the first and second constraints respectively in the given LPP, we get the following converted equations:

$$\begin{aligned} -2x_1 - x_2 + x_3 &= -2 \\ x_1 + x_2 + x_4 &= -1 \text{ and } x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

Then the readjusted objective function $z'$ is given by

$$z' = -x_1 - x_2 + 0 \cdot x_3 + 0 \cdot x_4.$$

Now taking $x_1 = x_2 = 0$, we get $x_3 = -2$, $x_4 = -1$, which is an initial basic but infeasible solution to the given problem.

Now we construct the starting dual simplex table as follows:

| $c_j$ | | | $-1$ | $-1$ | $0$ | $0$ | |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| $0$ | $y_3$ | $x_3 = -2$ | $-2$ | $-1$ | $1$ | $0$ | |
| $0$ | $y_4$ | $x_4 = -1$ | $1$ | $1$ | $0$ | $1$ | |
| | $z'_B = 0$ | | $1$ | $1$ | $0$ | $0$ | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \geq 0$ and $x_{B_1} = -2$, $x_{B_2} = -1$, the current basic solution is optimum or more than optimum but infeasible.

$$\text{Now } x_{B_r} = \text{Min}\{x_{B_i}; x_{B_i} < 0\} = \text{Min}\{-2, -1\} = -2 = x_{B_1}$$
$$\therefore r = 1$$

Hence, the vector $y_{B_1}$, i.e. $y_3$, leaves the basis.

$$\begin{aligned} \text{Again,} \quad \text{Max}\left\{\frac{z_j - c_j}{y_{rj}}, y_{rj} < 0\right\} &= \text{Max}\left\{\frac{z_j - c_j}{y_{1j}}, y_{1j} < 0\right\} \\ = \text{Max}\left\{\frac{z_1 - c_1}{y_{11}}, \frac{z_2 - c_2}{y_{12}}\right\} &= \text{Max}\left\{\frac{1}{-2}, \frac{1}{-1}\right\} = -\frac{1}{2}, \end{aligned}$$

which occurs for $j = 1$. Then the vector $y_1$ enters the basis, and $y_{rj}$ i.e. $y_{11} = -2$ is the key element.

Now we construct the next table.

| | $c_j$ | | $-1$ | $-1$ | $0$ | $0$ | |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| $-1$ | $y_1$ | $x_1 = 1$ | $1$ | $\frac{1}{2}$ | $-\frac{1}{2}$ | $0$ | |
| $0$ | $y_4$ | $x_4 = -2$ | $0$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $1$ | |
| | $z'_B = -1$ | | $0$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $0$ | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \geq 0$ and $x_{B_2} = -2$ is negative, the current basic solution is an optimum or more than optimum but infeasible solution.

Since only $x_{B_2} = -2$ is negative, then $r = 2$ and the corresponding basis vector $y_{B_2}$ i.e. $y_4$ leaves the basis.

But since all $y_{2j} (j = 1, 2, 3, 4)$ are non-negative, no other vector enters the basis. Thus, there does not exist any feasible solution to the given LPP.

*Example 2* Using the dual simplex method, solve the following LPP:

Minimize $z = 2x_1 + x_2 + x_3$
subject to

$$4x_1 + 6x_2 + 3x_3 \le 8$$
$$x_1 - 9x_2 + x_3 \le -3$$
$$-2x_1 - 3x_2 + 5x_3 \le -4 \quad \text{and} \quad x_1,\, x_2,\, x_3 \ge 0.$$

**Solution**
Here the problem is a minimization problem.

Let $z' = -z$. Then $\text{Min}\, z = -\text{Max}(-z) = -\text{Max}\, z'$ using the mini-max principle.

Hence, the problem is to maximize $z' = -2x_1 - x_2 - x_3$.

Now introducing three slack variables $x_4 \ge 0$, $x_5 \ge 0$, $x_6 \ge 0$, one to each constraint respectively, we get the following converted equations:

$$4x_1 + 6x_2 + 3x_3 + x_4 = 8$$
$$x_1 - 9x_2 + x_3 + x_5 = -3$$
$$-2x_1 - 3x_2 + 5x_3 + x_6 = -4 \quad \text{and} \quad x_j \ge 0, j = 1, 2, \ldots, 6$$

Then the readjusted objective function $z'$ is given by

$$z' = -2x_1 - x_2 - x_3 + 0 \cdot x_4 + 0 \cdot x_5 + 0 \cdot x_6.$$

Now taking $x_1 = x_2 = x_3 = 0$, we get $x_8 = 8$, $x_5 = -3$, $x_6 = -4$, which is an initial basic but infeasible solution to the given problem.

Now we construct the starting dual simplex table.

| | $c_j$ | | $-2$ | $-1$ | $-1$ | $0$ | $0$ | $0$ | |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| $0$ | $y_4$ | $x_4 = 8$ | $4$ | $6$ | $3$ | $1$ | $0$ | $0$ | |
| $0$ | $y_5$ | $x_5 = -3$ | $1$ | $-9$ | $1$ | $0$ | $1$ | $0$ | |
| $0$ | $y_6$ | $x_6 = -4$ | $-2$ | $-3$ | $5$ | $0$ | $0$ | $1$ | |
| $z'_B = 0$ | | | $2$ | $1$ | $1$ | $0$ | $0$ | $0$ | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \ge 0$ and $x_{B_2} = -3$, $x_{B_3} = -4$, the current solution is optimum or more than optimum but infeasible.

Since $x_{B_r} = \text{Min}[x_{B_i};\, x_{B_i} < 0] = \text{Min}[x_{B_2},\, x_{B_3}] = \text{Min}[-3, -4] = -4 = x_{B_3}$,
$\therefore r = 3.$

Hence, the vector $y_{B_3}$, i.e. $y_6$, leaves the basis.

$$\text{Again,}\quad \text{Max}\left\{\frac{z_j - c_j}{y_{rj}};\, y_{rj} < 0\right\} = \text{Max}\left\{\frac{z_j - c_j}{y_{3j}};\, y_{3j} < 0\right\} = \text{Max}\left\{\frac{z_1 - c_1}{y_{31}}, \frac{z_2 - c_2}{y_{32}}\right\}$$
$$= \text{Max}\left\{\frac{2}{-2}, \frac{1}{-3}\right\} = -\frac{1}{3}, \text{which occurs for } j = 2.$$

Then the vector $y_2$ enters the basis, and $y_{rj}$, i.e. $y_{32} = -3$, is the key element.

Now we construct the next table.

| | $c_j$ | | $-2$ | $-1$ | $-1$ | $0$ | $0$ | $0$ | |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| $0$ | $y_4$ | $x_4 = 0$ | $0$ | $0$ | $13$ | $1$ | $0$ | $2$ | |
| $0$ | $y_5$ | $x_5 = 9$ | $7$ | $0$ | $-14$ | $0$ | $1$ | $-3$ | |
| $-1$ | $y_2$ | $x_2 = \frac{4}{3}$ | $\frac{2}{3}$ | $1$ | $-\frac{5}{3}$ | $0$ | $0$ | $-\frac{1}{3}$ | |
| $z'_B = -\frac{4}{3}$ | | | $\frac{4}{3}$ | $0$ | $\frac{8}{3}$ | $0$ | $0$ | $\frac{1}{3}$ | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \geq 0 \;\forall j = 1, 2, \ldots,$ 6 and all $x_{B_i} \geq 0$, an optimal basic feasible solution has been obtained. Hence, the required optimal basic feasible solution to the given LPP is $x_1 = 0,\; x_2 = \frac{4}{3},\; x_3 = 0$ with Min $z = -\text{Max}\, z' = -\left(-\frac{4}{3}\right) = \frac{4}{3}$.

*Example 3* Use the dual simplex method to solve the following LPP:

$$\begin{aligned} &\text{Maximize} \quad z = -3x_1 - x_2 \\ &\text{subject to} \\ &\qquad\qquad x_1 + x_2 \geq 1 \\ &\qquad\qquad 2x_1 + 3x_2 \geq 2 \;\text{ and } x_1,\; x_2 \geq 0. \end{aligned}$$

**Solution**

The given LPP can be written as follows:

$$\begin{aligned} &\text{Maximize} \quad z = -3x_1 - x_2 \\ &\text{subject to} \\ &\qquad\qquad -x_1 - x_2 \leq -1 \\ &\qquad\qquad -2x_1 - 3x_2 \leq -2 \;\text{ and }\; x_1, x_2 \geq 0. \end{aligned}$$

Now introducing the slack variables $x_3 \geq 0$ and $x_4 \geq 0$ in the constraint of the given LPP, we get the following reduced LPP:

$$\begin{aligned} &\text{Maximize} \quad z = -3x_1 - x_2 + 0 \cdot x_3 + 0 \cdot x_4 \\ &\text{subject to} \\ &\qquad\qquad -x_1 - x_2 + x_3 = -1 \\ &\qquad\qquad -2x_1 - 3x_2 + x_4 = -2 \quad \text{and} \quad x_1, x_2, x_3, x_4 \geq 0. \end{aligned}$$

Now setting $x_1 = x_2 = 0$, we get $x_3 = -1,\ x_4 = -2$, which is an initial basic but non-feasible solution to the given problem.

The starting dual simplex table is as follows:

| | $c_j$ | | $-3$ | $-1$ | $0$ | $0$ | |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| $0$ | $y_3$ | $x_3 = -1$ | $-1$ | $-1$ | $1$ | $0$ | |
| $0$ | $y_4$ | $x_4 = -2$ | $-2$ | $-3$ | $0$ | $1$ | |
| $z_B = 0$ | | | $3$ | $1$ | $0$ | $0$ | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \geq 0$ and $x_{B_1} = -1,\ x_{B_2} = -2$, the current basic solution is optimum or more than optimum but infeasible.

Here the most negative basic variable is $x_{B_2} = -2$. Hence $r = 2$, and the corresponding basis vector $y_4$ leaves the basis.

For the vector which enters the basis we select

$$\text{Max}\left\{\frac{z_j - c_j}{y_{2j}};\ y_{ij} < 0\right\} = \text{Max}\left\{\frac{z_1 - c_1}{y_{21}}, \frac{z_2 - c_2}{y_{22}}\right\}$$
$$= \text{Max}\left\{\frac{3}{-2}, \frac{1}{-3}\right\} = \text{Max}\left\{-\frac{3}{2}, -\frac{1}{3}\right\} = -\frac{1}{3},\ \text{which occurs for } j = 2.$$

Then $y_2$ enters the basis, and $y_{2j}$, i.e. $y_{22} = -3$, is the key element.

Now we construct the next table.

| | $c_j$ | | $-3$ | $-1$ | $0$ | $0$ | |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| $0$ | $y_3$ | $x_3 = -\frac{1}{3}$ | $-\frac{1}{3}$ | $0$ | $1$ | $-\frac{1}{3}$ | |
| $-1$ | $y_2$ | $x_2 = \frac{2}{3}$ | $\frac{2}{3}$ | $1$ | $0$ | $-\frac{1}{3}$ | |
| $z_B = -\frac{2}{3}$ | | | $\frac{7}{3}$ | $0$ | $0$ | $\frac{1}{3}$ | $\leftarrow z_j - c_j$ |

Since all $z_j - c_j \geq 0$ and $x_{B_1} = -\frac{1}{3}$, the current basic solution is optimum or more than optimum but infeasible.

Since $x_{B_1} = -\frac{1}{3}$ is negative, hence $r = 1$ and the corresponding basis vector $y_3$ leaves the basis.

Now, $\text{Max}\left\{\frac{z_j - c_j}{y_{1j}}; y_{1j} < 0\right\}$

$$= \text{Max}\left\{\frac{z_1 - c_1}{y_{11}}, \frac{z_4 - c_4}{y_{14}}\right\} = \text{Max}\left\{\frac{7/3}{-1/3}, \frac{1/3}{-1/3}\right\} = \text{Max}\{-7, -1\} = -1, \text{which occurs for } j = 4.$$

Then $y_4$ enters the basis and $y_{14} = -\frac{1}{3}$ is the key element.

Now we construct the third table.

| | $c_j$ | | $-3$ | $-1$ | $0$ | $0$ | |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| $0$ | $y_4$ | $x_4 = 1$ | $1$ | $0$ | $-3$ | $1$ | |
| $-1$ | $y_2$ | $x_2 = 1$ | $1$ | $1$ | $-1$ | $0$ | |
| $z_B = -1$ | | | $2$ | $0$ | $1$ | $0$ | $\leftarrow z_j - c_j$ |

Here all $z_j - c_j \geq 0\ \forall j = 1, 2, 3, 4$ and all $x_{B_i}$'s are non-negative. This shows that an optimal basic feasible solution has been obtained. Thus, the required optimal basic feasible solution is $x_1 = 0,\ x_2 = 1$ and $z_{\text{Max}} = -1$.

## 5.4 Difference Between Simplex Method and Dual Simplex Method

The dual simplex method is similar to the regular simplex method, except that in the latter the starting initial basic solution is feasible but not optimum, while in the former it is infeasible but optimum or better than optimum. The simplex method works towards optimality, while the dual simplex method works towards feasibility.

In the consecutive iterations of the simplex method for solving a maximization problem, the value of the objective function gradually increases and finally reaches to its optimal value. In each iteration, the solution is basic feasible but non-optimal. On the other hand, in the consecutive iterations of the dual simplex method for solving the same problem, the value of the objective function gradually decreases and finally reaches to its optimal value. In each iteration, the solution is basic non-feasible but optimal (or better than optimal).

## 5.5 Modified Dual Simplex Method

If the initial table of the dual simplex method contains some negative basic variables and some of the net evaluations are negative, then the dual simplex method is not applicable. In such situations, the dual simplex method is to be modified to form an equivalent LPP in which some basic variables are negative but all net evaluations are non-negative. Hence, the standard dual simplex method can be applied to that equivalent LPP.

One of the modified methods is the artificial constraint method. In this method, the variables corresponding to the negative net evaluations and the variable corresponding to the most negative component of net evaluations are considered. Let $z_p - c_p$ be the most negative net evaluation corresponding to the variable $x_p$. In this method we have to consider the artificial constraint

$$\sum x_j \leq M,$$

where $\sum$ is extended over all $j$'s for which $z_j - c_j < 0$, and $M$ is a sufficiently large positive number. Adding the slack variable $x_M$ to this constraint, we have

$$\sum x_j + x_M = M.$$

From this equation, we obtain $x_p = M - \left(x_M + \sum_{j \neq p} x_j\right)$.

Then substituting this $x_p$ in the original objective function and all the constraints, we get a reduced problem. This reduced problem together with the artificial constraint is equivalent to the given problem. For this equivalent LPP, all $z_j - c_j \geq 0$. Thus, to solve this problem, the dual simplex method can now be applied.

*Example 4* Solve the following problem by the modified dual simplex method:

$$\begin{array}{ll} \text{Minimize} & z = -2x_1 - x_2 - x_3 \\ \text{subject to} & \\ & 4x_1 + 6x_2 + 3x_3 \leq 8 \\ & -x_1 + 9x_2 - x_3 \geq 3 \\ & 2x_1 + 3x_2 - 5x_3 \geq 4 \\ & \text{and } x_1, x_2, x_3 \geq 0. \end{array}$$

**Solution**

The given minimization problem is written as a maximization problem as follows:

$$\begin{array}{ll} \text{Maximize} & z' = 2x_1 + x_2 + x_3 + 0 \cdot x_4 + 0 \cdot x_5 + 0 \cdot x_6 \\ \text{subject to} & \\ & 4x_1 + 6x_2 + 3x_3 + x_4 \quad = 8 \\ & x_1 - 9x_2 + x_3 + x_5 \quad = -3 \\ & -2x_1 - 3x_2 + 5x_3 + x_6 = -4 \\ & \text{and } x_1, x_2, x_3, x_4, x_5, x_6 \geq 0. \end{array}$$

Now we construct the initial dual simplex table.

| | | $c_j$ | 2 | 1 | 1 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 0 | $y_4$ | $x_4 = 8$ | 4 | 6 | 3 | 1 | 0 | 0 | |
| 0 | $y_5$ | $x_5 = -3$ | 1 | −9 | 1 | 0 | 1 | 0 | |
| 0 | $y_6$ | $x_6 = -4$ | −2 | −3 | 5 | 0 | 0 | 1 | |
| $z'_B = 0$ | | | −2 | −1 | −1 | 0 | 0 | 0 | $\leftarrow z_j - c_j$ |

Here there are negative net evaluations, so the standard dual simplex method is not applicable.

From this simplex table, it is seen that the most negative evaluation is $z_1 - c_1 = -2$ corresponding to the variablex $x_1$.

Hence, the artificial constraint is

$x_1 + x_2 + x_3 \leq M$, where $M$ is a very large positive number.

Adding the slack variable $x_M$, we have

$$x_1 + x_2 + x_3 + x_M = M,$$

from which we have $x_1 = M - x_2 - x_3 - x_M$.

Substituting the expression for $x_1$ in the given LPP, we have:

$$\begin{array}{ll} \text{Maximize} & z'' = 2(M - x_2 - x_3 - x_M) + x_2 + x_3 + 0 \cdot x_4 + 0 \cdot x_5 + 0 \cdot x_6 \\ \text{subject to} & \\ & 4(M - x_2 - x_3 - x_M) + 6x_2 + 3x_3 + x_4 \quad = 8 \\ & M - x_2 - x_3 - x_M - 9x_2 + x_3 + \qquad = -3 \\ & -2(M - x_2 - x_3 - x_M) - 3x_2 + 5x_3 + x_6 = -4 \\ & x_1 + x_2 + x_3 + x_M \qquad = M \\ & \text{and } x_1, x_2, x_3, x_4, x_5, x_6, x_M \geq 0 \end{array}$$

$$\begin{array}{ll} \text{or Maximize} & z'' = -2x_M - x_2 - x_3 + 0 \cdot x_4 + 0 \cdot x_5 + 0 \cdot x_6 + 2M \\ \text{subject to} & \\ & -4x_M + 2x_2 - x_3 + x_4 = 8 - 4M \\ & -x_M - 10x_2 + x_5 \qquad = -3 - M \\ & 2x_M - x_2 + 7x_3 + x_6 \quad = -4 + 2M \\ & x_M + x_1 + x_2 + x_3 \qquad = M \\ & \text{and } x_1, x_2, x_3, x_4, x_5, x_6, x_M \geq 0 \end{array}$$

Now we construct the dual simplex table.

| | | $c_j$ | $-2$ | 0 | $-1$ | $-1$ | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_M$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 0 | $y_4$ | $x_4 = 8 - 4M$ | $-4$ | 0 | 2 | $-1$ | 1 | 0 | 0 | |
| 0 | $y_5$ | $x_5 = -3 - M$ | $-1$ | 0 | $-10$ | 0 | 0 | 1 | 0 | |
| 0 | $y_6$ | $x_6 = -4 + 2M$ | 2 | 0 | $-1$ | 7 | 0 | 0 | 1 | |
| 0 | $y_1$ | $x_1 = M$ | 1 | 1 | 1 | 1 | 0 | 0 | 0 | |
| $z''_B = 2M$ | | | 2 | 1 | 1 | 1 | 0 | 0 | 0 | $\leftarrow z_j - c_j$ |

Now, $\quad x_{B_r} = \text{Min}\{x_{B_i}; x_{B_i} < 0\} = \text{Min}\{8 - 4M, -3 - M\} = 8 - 4M = x_{B_1}$.

Hence, $r = 1$ and the corresponding vector $y_{B_1}$, i.e. $y_4$ leaves the basis.

Again, $\text{Max}\left\{\frac{z_j-c_j}{y_{rj}}; y_{rj}<0\right\} = \text{Max}\left\{\frac{z_j-c_j}{y_{1j}}; y_{1j}<0\right\} = \text{Max}\left\{\frac{2}{-4}, \frac{1}{-1}\right\} = \frac{2}{-4}$, which occurs for the vector $y_M$. Then the vector $y_M$ enters the basis and $y_{M1} = -4$ is the key element.
Now we construct the next dual simplex table.

| | | $c_j$ | $-2$ | 0 | $-1$ | $-1$ | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_M$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| $-2$ | $y_M$ | $x_M = -2+M$ | 1 | 0 | $-1/2$ | 1/4 | $-1/4$ | 0 | 0 | |
| 0 | $y_5$ | $x_5 = -5$ | 0 | 0 | $-21/2$ | 1/4 | $-1/4$ | 1 | 0 | |
| 0 | $y_6$ | $x_6 = 0$ | 0 | 0 | 0 | 13/2 | 1/2 | 0 | 1 | |
| 0 | $y_1$ | $x_1 = 2$ | 0 | 1 | 3/2 | 3/4 | ¼ | 0 | 0 | |
| $z''_B = 4$ | | | 0 | 0 | 2 | 1/2 | 1/2 | 0 | 0 | $\leftarrow z_j - c_j$ |

Since only $x_{B_2}$ is negative hence $r = 2$ and the corresponding basis vector $y_5$ leaves the basis.

Again, $\text{Max}\left\{\frac{z_j-c_j}{y_{2j}}; y_{2j}<0\right\} = \text{Max}\left\{\frac{2}{-21/2}, \frac{1/2}{1/4}\right\} = \frac{4}{-21}$, which occurs for the vector $y_2$. Then the vector $y_2$ enters the basis and $y_{22} = -21/2$ is the key element.

Now we construct the next dual simplex table.

| | | $c_j$ | $-2$ | 0 | $-1$ | $-1$ | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_M$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| $-2$ | $y_M$ | $x_M = M - 37/21$ | 1 | 0 | 0 | 5/21 | $-5/21$ | $-1/21$ | 0 | |
| $-1$ | $y_2$ | $x_2 = 10/21$ | 0 | 0 | 1 | $-1/42$ | 1/42 | $-2/21$ | 0 | |
| 0 | $y_6$ | $x_6 = 0$ | 0 | 0 | 0 | 13/2 | 1/2 | 0 | 1 | |
| 0 | $y_1$ | $x_1 = 9/7$ | 0 | 1 | 0 | 11/14 | 3/14 | 1/7 | 0 | |
| $z''_B = 64/21$ | | | 0 | 0 | 0 | 13/42 | 19/42 | 4/21 | 0 | $\leftarrow z_j - c_j$ |

Here all the basic variables are non-negative. So an optimal basic feasible solution has been obtained. Therefore, the required optimal basic feasible solution is given by

$$x_1 = 9/7, x_2 = 10/21, x_3 = 0 \text{ and } z'_{\text{Max}} = (-2M + 64/21) + 2M = 64/21.$$

Therefore, $z_{\text{Min}} = -z'_{\text{Max}} = -64/21$.

## 5.6 Exercises

Use the dual simplex method to solve the following LPPs:

1. Minimize $z = -3x_1 - x_2$
   subject to
   $$x_1 + x_2 \geq 1$$
   $$2x_1 + 3x_2 \geq 2 \quad \text{and } x_1,\ x_2 \geq 0.$$

2. Minimize $z = 2x_1 + x_2$
   subject to
   $$3x_1 + x_2 \geq 3$$
   $$4x_1 + 3x_2 \geq 6$$
   $$x_1 + 2x_2 \leq 3 \text{ and } x_1,\ x_2 \geq 0.$$

3. Maximize $z = -2x_1 - 3x_2 - x_3$
   subject to
   $$2x_1 + x_2 + 2x_3 \geq 3$$
   $$3x_1 + 2x_2 + x_3 \geq 4 \text{ and } x_1,\ x_2,\ x_3 \geq 0.$$

4. Minimize $z = x_1 + 2x_2 + 3x_3$
   subject to
   $$2x_1 - x_2 + x_3 \geq 4$$
   $$x_1 + x_2 + 2x_3 \leq 8$$
   $$x_2 - x_3 \geq 2 \text{ and } x_1,\ x_2,\ x_3 \geq 0.$$

5. Minimize $z = -2x_1 - x_2 - x_3$
   subject to
   $$4x_1 + 6x_2 + 3x_3 \leq 8$$
   $$x_1 - 9x_2 + x_3 \leq -3$$
   $$-2x_1 - 3x_2 + 5x_3 \leq -4 \text{ and } x_1,\ x_2,\ x_3 \geq 0.$$

6. Minimize $z = 10x_1 + 6x_2 + 2x_3$
   subject to
   $$-x_1 + x_2 + x_3 \geq 1$$
   $$3x_1 + x_2 - x_3 \geq 2 \text{ and } x_1,\ x_2,\ x_3 \geq 0.$$

7. Minimize $z = x_1 + x_2$
   subject to
   $$2x_1 + x_2 \geq 2$$
   $$-x_1 - x_2 \geq 1 \text{ and } x_1,\ x_2 \geq 0.$$

8. Maximize $z = -3x_1 - 2x_2$
subject to $x_1 + x_2 \geq 1$
$x_1 + x_2 \leq 7$
$x_1 + 2x_2 \geq 10$
$x_2 \leq 3$ and $x_1,\ x_2 \geq 0.$

9. Maximize $z = 3x_1 + 2x_2$
subject to

$2x_1 + x_2 \leq 5$
$x_1 + x_2 \leq 3$ and $x_1,\ x_2 \geq 0.$

10. Maximize $z = 2x_1 - 3x_2 - 2x_3$
subject to

$x_1 - 2x_2 - 3x_3 = 8$
$2x_2 + x_3 \leq 10$
$x_2 - 2x_3 \geq 4$ and $x_1,\ x_2,\ x_3 \geq 0.$

11. Minimize $z = 3x_1 + x_2$
subject to

$x_1 + x_2 \geq 1$
$2x_1 + 3x_2 \geq 2$ and $x_1,\ x_2 \geq 0.$

12. Minimize $z = 6x_1 + 7x_2 + 3x_3 + 5x_4$
subject to

$5x_1 + 6x_2 - 3x_3 + 4x_4 \geq 12$
$x_2 + 5x_3 - 6x_4 \geq 10$
$2x_1 + 5x_2 + x_3 + x_4 \geq 8$ and $x_1,\ x_2,\ x_3,\ x_4 \geq 0.$

13. Minimize $z = 5x_1 + 6x_2$
subject to

$x_1 + x_2 \geq 2$
$4x_1 + x_2 \geq 4$ and $x_1,\ x_2 \geq 0.$

14. Minimize $z = 4x_1 + 2x_2$
subject to

$3x_1 + x_2 \geq 27$
$x_1 + x_2 \geq 21$
$x_1 + 2x_2 \geq 30$ and $x_1,\ x_2 \geq 0.$

# Chapter 6
# Bounded Variable Technique

## 6.1 Objective

The objective of this chapter is to discuss the modified simplex method to solve LPPs whose variables are restricted with lower and upper bounds.

## 6.2 Introduction

In the theory of linear programming, it is well known that the decision variables of a linear programming problem (LPP) satisfy non-negativity conditions/restrictions only. However, in real-life situations, some (or all) of the variables are bounded by lower and upper limits. These bounds of the variables are expressed as bound constraints. This means that, in an LPP, the bound constraints of some (or all) variables are given in addition to the general constraints. In such cases, the standard form of the LPP will be as follows:

Maximize $z = \sum_{j=1}^{n} c_j x_j$

subject to the constraints $Ax = b,\ l \leq x \leq u$,

$$\begin{aligned} \text{where}\quad & x = [x_1, x_2, \ldots, x_n]^{\mathrm{T}}, \quad b = [b_1, b_2, \ldots, b_n]^{\mathrm{T}}, \quad l = [l_1, l_2, \ldots, l_n]^{\mathrm{T}}, \\ & u = [u_1, u_2, \ldots, u_n]^{\mathrm{T}}, \quad A = [a_{ij}]_{\mathrm{mxn}}, \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n. \end{aligned}$$

Here $A$ is an $m \times n$ real matrix, and $l$ and $u$ denote the lower and upper bounds of $x$ respectively. In cases of unbounded variables, these limits are considered as 0 and $\infty$ respectively.

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_6

The bound constraints $l \leq x \leq u$ can be written separately as $x \geq l$ and $x \leq u$. Now introducing surplus variables $s'$ and slack variables $s''$ in the preceding bound constraints, we get the equality constraints as follows:

$$x - s' = l \text{ and } x + s'' = u, \quad \text{where} \quad x, s', s'' \geq 0.$$

The lower bound constraints can be written as $x = l + s'$. Thus, $x$ can be eliminated from all the remaining constraints by substituting $x = l + s'$.

On the other hand, the upper bound constraints can be written $x = u - s''$. This does not serve the purpose, since there is no guarantee that $x$ will be non-negative. This difficulty is overcome by using a special technique called the bounded variable simplex method. In this method, the optimality condition for a solution will be the same, as the constraints $x + s'' = u$ require modification in the feasibility condition of decision variables due to the following reasons:

(i) A basic variable should become a non-basic variable at its upper bound (in the usual simplex method, all non-basic variables are at zero level).
(ii) When a non-basic variable becomes a basic variable, its value should not exceed its upper bound without affecting the non-negativity and upper bound conditions of all existing basic variables.

## 6.3 The Computational Procedure

The iterative procedure for finding an optimum basic feasible solution (BFS) to an LPP with bounded variables is as follows:

Step 1: Check whether the objective function of the given LPP is to be maximized or minimized. If it is to be minimized, then convert it into a problem of maximizing by using

$$\text{Minimum}\, z = -\text{Maximum}(-z).$$

Step 2: Check whether all $b_i$ ($i$ = 1, 2,…, $m$) are non-negative. If any one of $b_i$ is negative, then multiply the corresponding inequality of the constraints by (−1) so as to get all $bi$ ($i$ = 1, 2,…, $m$) as non-negative.

Step 3: Convert all the inequality constraints (except bound constraints) into equations by introducing slack and/or surplus variables in the constraints. Set the coefficients of these variables in the objective function equal to zero. We note that the lower and upper bounds of slack and surplus variables are assumed as 0 and $\infty$ respectively.

Step 4: If the lower bound $l_j$ ($j$ = 1, 2, …, $n$) of any variable $x_j$ is at a positive level (i.e. if $l_j \leq x_j$, $l_j > 0$), set $x_j = l_j + x_j$, $x_j \geq 0$ in the reduced problem obtained from Step 3.

Step 5: Obtain an initial BFS.

Step 6: Compute the net evaluations $z_j - c_j$ ($j = 1, 2, \ldots, n$) and examine their signs. Two cases may arise:

(i) If all $z_j - c_j \geq 0$, then an optimum BFS has been attained.
(ii) If at least one $z_j - c_j < 0$, choose the most negative of $z_j - c_j$. Let it be $z_r - c_r$ for some $j = r$.

Step 7: Compute $\theta = \text{Min}\{\theta_1, \theta_2, u_r\}$,

$$\text{where} \quad \theta_1 = \underset{i}{\text{Min}}\left\{\frac{x_{B_i}}{y_{ir}}, y_{ir} > 0\right\},$$
$$\theta_2 = \underset{i}{\text{Min}}\left\{\frac{u(x_{B_i}) - x_{B_i}}{-y_{ir}}, y_{ir} < 0\right\},$$

where the $x_{Bi}$'s correspond to the current BFS, and $u_r$ is the upper bound for the variable $x_r$.

Note that $\theta_2 = \infty$ if all $y_{ir} > 0$ and $\theta_1 = \infty$ if all $y_{ir} < 0$.

In finding the value of $\theta$, three cases may arise:

(i) If $\theta = \theta_1$ which corresponds to $x_{Bk}$, then $y_r$ enters the basis and $y_k$ leaves it.
(ii) If $\theta = \theta_2$ which corresponds to $x_{Bk}$, then $y_r$ enters the basis and $y_k$ leaves it. Due to this, the value of one or more basic variable will be negative in the next iteration. In that case, the values of basic variables are to be updated. This can be done by putting the non-basic variables $x_r$ (corresponding to the most negative $z_j - c_j$) at zero level at the upper bound by substituting $x_r = u_r - x'_r (0 \leq x'_r \leq u_r)$, as the upper bound of $x_r$ is $u_r$. In this case,

$$(x_{B_i})_{\text{new}} = (x_{B_i})_{\text{old}} - y_{ir} u_r$$

(iii) If $\theta = u_r, x_r$ is substituted at its upper bound and it remains non-basic; i.e. it can be put at zero level by using

$$x_r = u_r - x'_r, 0 \leq x'_r \leq u_r.$$

In this situation, the basic variables are to be updated by the formula

$$(x_{B_i})_{\text{new}} = (x_{B_i})_{\text{old}} - y_{ir} u_r.$$

Then Go to Step 6.

*Example 1* Using the bounded variable technique, solve the following LPP:

Maximize $z = 3x_1 + 5x_2 + 2x_3$

subject to $x_1 + 2x_2 + 2x_3 \leq 14, 2x_1 + 4x_2 + 3x_3 \leq 23, 0 \leq x_1 \leq 4, 2 \leq x_2 \leq 5,$ $0 \leq x_3 \leq 3.$

**Solution** By introducing the stack variables $x_4(\geq 0)$ and $x_5(\geq 0)$, in the first and second constraints respectively, we get the reduced problem as

$$
\begin{aligned}
\text{Maximize } z &= 3x_1 + 5x_2 + 2x_3 + 0.x_4 + 0.x_5 \\
x_1 + 2x_2 + 2x_3 + x_4 &= 14 \\
2x_1 + 4x_2 + 3x_3 + x_5 &= 23 \\
\text{and } 0 \leq x_1 \leq 4, 2 \leq x_2 &\leq 5, 0 \leq x_3 \leq 3, x_4, x_5 \geq 0.
\end{aligned}
$$

Since there are no upper bounds specified for the variables $x_4$ and $x_5$, we arbitrarily assume that all have an upper bound at $\propto$, i.e. $u_4 = u_5 = \propto$.

Since $x_2$ has a positive lower bound, let us take $x_2 = x_2' + 2$. Then $2 \leq x_2 \leq 5$ implies $0 \leq x_2' \leq 3$. Then the given LPP reduces to

$$
\begin{aligned}
&\text{Maximize } z = 3x_1 + 5(x_2' + 2) + 2x_3 + 0.x_4 + 0.x_5 \\
&x_1 + 2(x_2' + 2) + 2x_3 + x_4 = 14 \text{ or, } x_1 + 2x_2' + 2x_3 + x_4 = 10 \\
&2x_1 + 4(x_2' + 2) + 3x_3 + x_5 = 23 \text{ or, } 2x_1 + 4x_2' + 3x_3 + x_5 = 15 \\
&\text{and } 0 \leq x_1 \leq 4, \quad 0 \leq x_2' \leq 3, 0 \leq x_3 \leq 3, x_4, x_5 \geq 0
\end{aligned}
$$

Here the initial BFS is $x_4 = 10, x_5 = 15$ with $B = I_2$ as initial basis.

Now we construct the starting simplex table:

| | | $c_j$ | 3 | 5 | 2 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2'$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_4$ | $x_4 = 10$ | 1 | 2 | 2 | 1 | 0 | |
| 0 | $y_5$ | $x_5 = 15$ | 2 | 4 | 3 | 0 | 1 | |
| $z_B = 10$ | | | –3 | –5 | –2 | 0 | 0 | $\leftarrow \Delta_j$ |
| $u_j$ | | | 4 | 3 | 3 | $\infty$ | $\infty$ | |

Since $\Delta_2$ is most negative, $y_2'$ enters the basis.

Now, $\theta_1 = \underset{i}{\text{Min}}\left\{\frac{x_{B_i}}{y_{i2}'}, y_{i2}' > 0\right\} = \text{Min}\left\{\frac{10}{2}, \frac{15}{4}\right\} = \frac{15}{4}$ (corresponds to $x_5$)

$$\theta_2 = \underset{i}{\text{Min}}\left\{\frac{u(x_{B_i}) - x_{B_i}}{-y'_{i2}}; y'_{i2} < 0\right\} = \infty \text{ [As } y'_{i2} > 0 \quad \text{for} \quad i = 1, 2]$$

and $u'_2 = 3$

Hence, $\theta = \text{Min}\{\theta_1, \theta_2, u'_2\} = \text{Min}\{\frac{15}{4}, \infty, 3\} = 3$, which implies that $\theta = u'_2$.

Therefore, $x'_2$ is substituted at its upper bound and it remains non-basic.

Thus, $x'_2 = u'_2 - x''_2 = 3 - x''_2$, where $0 \le x''_2 \le 3$.

Due to this substitution, the values of the basic variables will be updated, and these updated values are

$$(x_{B_1})_{\text{new}} = (x_{B_1})_{\text{old}} - y'_{12}u'_2 = 10 - 2 \times 3 = 4$$
$$(x_{B_2})_{\text{new}} = (x_{B_2})_{\text{old}} - y'_{22}u'_2 = 15 - 4 \times 3 = 3.$$

Therefore, by using the updated values of the basic and non-basic variables, the initial simplex table becomes

| | | $c_j$ | 3 | –5 | 2 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y''_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_4$ | $x_4 = 4$ | 1 | −2 | 2 | 1 | 0 | |
| 0 | $y_5$ | $x_5 = 3$ | 2 | −4 | 3 | 0 | 1 | |
| $z_B = 15 + 10 = 25$ | | | −3 | 5 | −2 | 0 | 0 | $\leftarrow \Delta_j$ |
| | | $u_j$ | 4 | 3 | 3 | $\infty$ | $\infty$ | |

Since $\Delta_1$, i.e. $z_1 - c_1$, is most negative; hence $y_1$ enters the basis.

Now, $\theta_1 = \text{Min}\{\frac{4}{1}, \frac{3}{2}\} = \frac{3}{2}$ (corresponds to $x_5$)

$\theta_2 = \infty$, since all the elements of $y_1$ are positive and $u_1 = 4$

$$\therefore \theta = \text{Min}\{\theta_1, \theta_2, u_1\} = \text{Min}\left\{\frac{3}{2}, \infty, 4\right\} = \frac{3}{2}, \text{which implies } \theta = \theta_1.$$

Hence $y_{12} = 2$ is the key element and $y_5$ leaves the basis.

Now, we construct the next table.

| | | $c_j$ | 3 | −5 | 2 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y''_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_4$ | $x_4 = 5/2$ | 0 | 0 | 1/2 | 1 | −1/2 | |
| 3 | $y_1$ | $x_1 = 3/2$ | 1 | −2 | 3/2 | 0 | 1/2 | |
| $z_B = (25 + 3 \times 3/2) = 59/2$ | | | 0 | −1 | 5/2 | 0 | 3/2 | $\leftarrow \Delta_j$ |
| | | $u_j$ | 4 | 3 | 3 | $\infty$ | $\infty$ | |

Since $\Delta_j$ corresponding to $y_2''$ is negative, then $y_2''$ enters the basis.

Now, $\theta_1 = \infty$, since all the elements of $y_2''$ are either negative or zero.

$$\theta_2 = \text{Min}\left\{\frac{4-\frac{3}{2}}{-(-2)}\right\} = \frac{5}{4}, \text{which corresponds to } x_1.$$
$$u_2'' = 3.$$

Hence, $\theta = \text{Min}\{\theta_1, \theta_2, u_2''\} = \frac{5}{4}$, which implies $\theta = \theta_2$.

Hence, $y_{22}'' = -2$ is the key element and $y_1$ leaves the basis.

Now we construct the next table.

| | | $c_j$ | 3 | −5 | 2 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2''$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_4$ | $x_4 = 5/2$ | 0 | 0 | 1/2 | 1 | $-1/2$ | |
| −5 | $y_2''$ | $x_2'' = -3/4$ | $-1/2$ | 1 | −3/4 | 0 | $-1/4$ | |
| $z_B = (25 + 15/4) = 115/4$ | | | $-1/2$ | 0 | 7/4 | 0 | 5/4 | $\leftarrow \Delta_j$ |
| | | $u_j$ | 4 | 3 | 3 | $\infty$ | $\infty$ | |

Since $\Delta_1$ is only negative, then $y_1$ enters the basis. But from the preceding table, it is seen that the value of one basic variable is negative. So, we have to update the basic variables. This can be done by putting the non-basic variable $x_1$ at zero level at its upper bound by substituting $x_1 = 4 - x_1' (0 \le x_1' \le 4)$, as the upper bound of $x_1$ is 4.

Due to this substitution, the values of the basic variables will be updated, and these updated values are

$$(x_{B_1})_{\text{new}} = (x_{B_1})_{\text{old}} - y_{11}u_1 = \frac{5}{2} - 0 \times 4 = \frac{5}{2}$$
$$(x_{B_2})_{\text{new}} = (x_{B_2})_{\text{old}} - y_{21}u_1 = -\frac{3}{4} - (-\frac{1}{2}) \times 4 = \frac{5}{4}.$$

Thus using the updated values of the basic variables, we construct the simplex table.

| | | $c_j$ | −3 | −5 | 2 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1'$ | $y_2''$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_4$ | $x_4 = \frac{5}{2}$ | 0 | 0 | $\frac{1}{2}$ | 1 | $-\frac{1}{2}$ | |
| −5 | $y_2''$ | $x_2'' = \frac{5}{4}$ | $\frac{1}{2}$ | 1 | $-\frac{3}{4}$ | 0 | $-\frac{1}{4}$ | |
| $z_B = 25 + 3 \times 4 - \frac{25}{4} = \frac{123}{4}$ | | | $\frac{1}{2}$ | 0 | $\frac{7}{4}$ | 0 | $\frac{5}{4}$ | $\leftarrow \Delta_j$ |

Since all $\Delta_j \ge 0$, an optimal solution has been attained. The solution is given by $x_1' = 0, x_2'' = \frac{5}{4}, x_3 = 0$ with Maximum $z = \frac{123}{4}$.

Now, since $x_1 = 4 - x_1', x_1' = 0$ implies $x_1 = 4$.

Again, since $x_2' = 3 - x_2'', x_2'' = \frac{5}{4}$ implies $x_2' = 3 - \frac{5}{4} = \frac{7}{4}$.

Also, $x_2 = x_2' + 2 = \frac{7}{4} + 2 = \frac{15}{4}$.

Hence, the optimal BFS is $x_1 = 4, x_2 = \frac{15}{4}, x_3 = 0$ and Maximum $z = \frac{123}{4}$.

**Example 2** Solve the LPP by using the bounded variable method.

$$\begin{aligned} \text{Minimize } & z = 6x_1 - 2x_2 - 3x_3 \\ \text{subject to } & 2x_1 + 4x_2 + 2x_3 \le 8 \\ & x_1 - 2x_2 + 3x_3 \le 7 \end{aligned}$$

$$0 \le x_1 \le 2,\ 0 \le x_2 \le 2 \text{ and } 0 \le x_3 \le 1.$$

**Solution** The given LPP is a minimization problem.

Let $z' = -z$; then Min $z = -\text{Max}\ (-z) = -\text{Max}\ z^*$.
Hence, the problem is to Maximize $z^* = -6x_1 + 2x_2 + 3x_3$
subject to the given constraints.

Introducing slack variables $x_4$ and $x_5$, the given LPP reduces to

$$\begin{aligned} \text{Maximize } z^* &= -6x_1 + 2x_2 + 3x_3 \\ \text{subject to } 2x_1 + 4x_2 + 2x_3 + x_4 &= 8 \\ x_1 - 2x_2 + 3x_3 + x_5 &= 7 \end{aligned}$$

$$0 \le x_1, x_2 \le 2, 0 \le x_3 \le 1 \text{ and } x_4, x_5 \ge 0.$$

Since there are no specified upper bounds for the variables $x_4$, $x_5$, we may arbitrarily assume that $u_4 = \infty, u_5 = \infty$.

Here the initial BFS is $x_4 = 8$, $x_5 = 7$, with $B = I_2$ as the initial basis. Now, we construct the initial simplex table as follows:

| | | $c_j$ | $-6$ | 2 | 3 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 0 | $y_4$ | 8 | 2 | 4 | 2 | 1 | 0 | |
| 0 | $y_5$ | 7 | 1 | $-2$ | 3 | 0 | 1 | |
| $z_B = 0$ | | | 6 | $-2$ | $-3$ | 0 | 0 | $\leftarrow \Delta_j$ |
| | | $u_j$ | 2 | 2 | 1 | $\infty$ | $\infty$ | |

Since $\Delta_3$ is most negative, $y_3$ enters the basis.

Now $\theta_1 = \text{Min}\left\{\frac{8}{2}, \frac{7}{3}\right\} = \frac{7}{3}$, which corresponds to $y_5$,

$$\theta_2 = \infty \text{ since all } y_{i3} > 0 \text{ for i} = 1{,}2.$$
$$u_3 = 1.$$

$\theta = \text{Min}\{\theta_1, \theta_2, u_3\} = \text{Min}\left\{\frac{7}{3}, \infty, 1\right\} = 1$, which corresponds to $\theta = u_3$. Therefore, $x_3$ is substituted at its upper bound and it remains non-basic, thus

$$x_3 = u_3 - x_3' = 1 - x_3', \text{ where } \quad 0 \le x_3' \le 1.$$

Due to this substitution, the basic variables will be updated, and these updated values are

$$(x_{B_1})_{\text{new}} = (x_{B_1})_{\text{old}} - y_{13}u_3 = 8 - 2 \times 1 = 6$$
$$(x_{B_2})_{\text{new}} = (x_{B_2})_{\text{old}} - y_{23}u_3 = 7 - 3 \times 1 = 4.$$

The initial table is as follows:

| | | $c_j$ | $-6$ | 2 | -3 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3'$ | $y_4$ | $y_5$ | |
| 0 | $y_4$ | 6 | 2 | 4 | $-2$ | 1 | 0 | |
| 0 | $y_5$ | 4 | 1 | $-2$ | $-3$ | 0 | 1 | |
| $z_B^* = 3$ | | | 6 | $-2$ | 3 | 0 | 0 | $\leftarrow \Delta_j$ |
| | | $u_j$ | 2 | 2 | 1 | $\infty$ | $\infty$ | |

Since, $\Delta_2$ is negative only then $y_2$ enters the basis.

Now, $\theta_1 = \text{Min}\left\{\frac{6}{4}\right\} = \frac{3}{2}$, which corresponds to $y_4$,

$$\theta_2 = \text{Min}\{\infty\} = \infty \text{ and } u_2 = 2.$$

$\therefore \theta = \text{Min}\left\{\frac{3}{2}, \infty, 2\right\} = \frac{3}{2}$, which corresponds to $\theta = \theta_1$.

Hence, $y_{12}$ = 4 is the key element, and $y_4$ leaves the basis. Now we construct the next table:

| | | $c_j$ | $-6$ | 2 | $-3$ | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3'$ | $y_4$ | $y_5$ | |
| 2 | $y_2$ | $\frac{3}{2}$ | $\frac{1}{2}$ | 1 | $-\frac{1}{2}$ | $\frac{1}{4}$ | 0 | |
| 0 | $y_5$ | 7 | 2 | 0 | $-4$ | | 1 | |

(continued)

(continued)

| | | $c_j$ | $-6$ | 2 | $-3$ | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | $\frac{1}{2}$ | | |
| $z_B^* = 6$ | | | 7 | 0 | 2 | $\frac{1}{2}$ | 0 | $\leftarrow \Delta_j$ |

Since all $\Delta_j \geq 0$, the optimal solution has been reached and the optimal solution is given by

$$x_1 = 0, x_2 = \frac{3}{2} \text{ and } x_3' = 0,$$
$$\text{i.e. } x_1 = 0, x_2 = \frac{3}{2} \text{ and } x_3 = 1 \text{ as } x_3 = 1 - x_3'$$
$$\text{and Minimum } z = 6.0 - 2.\frac{3}{2} - 3.1 = -6.$$

## 6.4 Exercises

Solve the following LPP:

1. Maximize $z = 3x_1 + 2x_2$
   subject to the constraints

$$x_1 - 3x_2 \leq 3, x_1 - 2x_2 \leq 4, 2x_1 + x_2 \leq 20$$
$$x_1 + 3x_2 \leq 30, -x_1 + x_2 \leq 6$$
$$\text{and} \quad 0 \leq x_1 \leq 8,\ 0 \leq x_2 \leq 6.$$

2. Maximize $z = 3x_1 + 5x_2 + 2x_3$
   subject to the constraints

$$x_1 + 2x_2 + 2x_3 \leq 14,\ 2x_1 + 4x_2 + 3x_3 \leq 23$$
$$\text{and} \quad 0 \leq x_1 \leq 4,\ 2 \leq x_2 \leq 5,\ 0 \leq x_3 \leq 3.$$

3. Maximize $z = x_2 + 3x_3$
   subject to the constraints

$$x_1 + x_2 + x_3 \leq 10,\ x_1 - 2x_3 \leq 0,\ 2x_2 - x_3 \leq 10$$
$$\text{and} \quad 0 \leq x_1 \leq 8,\ 0 \leq x_2 \leq 4,\ x_3 \geq 0.$$

4. Maximize $z = 4x_1 + 4x_2 + 3x_3$
subject to the constraints

$$\begin{gathered} -x_1 + 2x_2 + 3x_3 \le 15,\ -x_2 + x_3 \le 4, \\ 2x_1 + x_2 - x_3 \le 6,\ x_1 - x_2 + 2x_3 \le 10 \\ \text{and} \quad 0 \le x_1 \le 8,\ x_2 \ge 0, 0 \le x_3 \le 4. \end{gathered}$$

5. Maximize $z = 3x_1 + x_2 + x_3 + 7x_4$
subject to the constraints

$$2x_1 + 3x_2 - x_3 + 4x_4 \le 40,\ -2x_1 + 2x_2 + 5x_3 - x_4 \le 35$$

$$x_1 + x_2 - 2x_3 + 3x_4 \le 100$$

$$\text{and} \quad x_1 \ge 2,\ x_2 \ge 1, x_3 \ge 3, x_4 \ge 4.$$

6. Maximize $z = 4x_1 + 10x_2 + 9x_3 + 11x_4$
subject to the constraints

$$\begin{aligned} &2x_1 + 2x_2 + 2x_3 + 2x_4 \le 5, \\ &48x_1 + 80x_2 + 160x_3 + 240x_4 \le 257 \end{aligned}$$

$$\text{and} \quad 0 \le x_j \le 1,\ j = 1, 2, 3, 4.$$

7. Maximize $z = 6x_1 - 2x_2 - 3x_3$
subject to the constraints

$$\begin{gathered} 2x_1 + 4x_2 + 2x_3 \le 8,\ x_1 - 2x_2 + 3x_3 \le 7 \\ \text{and} \quad 0 \le x_1 \le 2,\ 0 \le x_2 \le 2,\ 0 \le x_3 \le 1. \end{gathered}$$

# Chapter 7
# Post-optimality Analysis in LPPs

## 7.1 Objectives

The objectives of this chapter are:

- To discuss the effect of discrete changes of the cost/profit vector and requirement vector on the optimal solution of a linear programming problem (LPP)
- To find the feasibility and optimality conditions for the discrete changes of the elements of the coefficient matrix of the LPP
- To study the effect of structural changes (i.e. inclusion and deletion of a decision variable and constraint) in an LPP.

## 7.2 Introduction

When solving a real-life linear programming problem (LPP), after formulation and obtaining an optimal solution of the said LPP, sometimes we are interested in studying the effect of changes (discrete as well as continuous) of different parameters of the problem on the current optimal solution. This situation arises in the following two cases:

**Case I**: An error (wrong value used for cost/profit coefficient, omission of variables or constraints, etc.) has been observed in the problem formulation after obtaining the optimal solution.
**Case II**: The system is analysed to study the effect of changes in different parameters.

For Case I, the error can be removed in two different ways:

(i) Solve the new LPP after replacing the correct values of the components which are used wrongly.

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_7

(ii) Develop some procedure so that the current optimum solution can be used to find the correct solution of the problem.

For Case I, to reduce the computational effort, one should choose option (ii). On the other hand, for Case II, option (ii) of Case I is used, as there is no other option. An analysis of such a post-optimal problem can thus be termed a post-optimality analysis.

The changing parameters may be either discrete or continuous. The study of the effect of discrete parameter changes on the optimal solution is called sensitivity analysis, and that of continuous changes is termed parametric programming.

The changes or variations in an LPP which are studied in the post-optimality analysis are:

(i) Changes in the cost coefficients $c_j$ of the objective function, i.e. a change in the cost or profit vector
(ii) Changes in the requirement vector $b$
(iii) Changes in the coefficient of the constraints
(iv) Structural changes due to addition or deletion of some variables
(v) Structural changes due to addition or deletion of some linear constraints.

After these changes have been effected in an LPP, we may have any of the following cases:

(i) The optimum solution remains unaltered; i.e. the basic variables and their values are essentially unchanged.
(ii) The basic variables remain the same, but their values may change.
(iii) The basic solution is changed entirely.

## 7.3 Discrete Changes in the Profit Vector

Let $x_B$ be the optimal basic feasible solution of the following LPP:

$$\text{Maximize } z = \langle c^T, x \rangle$$

subject to $Ax = b$ and $x \geq 0$,

where c, $x^T \in R^n, b^T \in R^m$ and $A$ is an $m \times n$ real matrix of rank $m$.

Let $\Delta c_k$ be the amount which is added to the $k$th component $c_k$ of $c = (c_1, c_2, \ldots, c_n)$, so that the new value of the $k$th component becomes $c_k^* = c_k + \Delta c_k$.

Since $x_B = B^{-1}b$ is independent of $c$, then for any change of $c_k, k = 1, 2, \ldots, n$, the value of $x_B$ will not be changed.

Hence, the solution $x_B$ remains basic feasible for the current problem.

Here, two cases may arise:

**Case I**: $c_k$ is not in $c_B$.
**Case II**: $c_k$ is in $c_B$.

**Case I**: When $c_k$ is not in $c_B$, $c_k$ is not the coefficient of the basic variable in the objective function. In that case, the net evaluation corresponding to the $k$th profit coefficient is given by

$$z_k^* - c_k^* = z_k - (c_k + \Delta c_k), \quad \forall c_k \notin c_B$$

since $z_k$ is not affected for any change in $c_k$.

Also, we note that the net evaluations corresponding to the profit coefficients other than $c_k$ remain unchanged.

Thus, the current solution $x_B$ remains optimum for the new problem

$$\begin{aligned} &\text{if} \quad z_k - c_k - \Delta c_k \geq 0 \quad && \forall c_k \notin c_B, \\ &\text{i.e. if } \Delta c_k \leq z_k - c_k && \forall c_k \notin c_B. \end{aligned}$$

**Case II**: When $c_k$ is in $c_B$, $c_k$ is the coefficient of the basic variable in the objective function.

In this case, the net evaluations are given by

$$\begin{aligned} z_j - c_j &= \sum_{i=1}^{m} c_{B_i} y_{ij} - c_j \\ &= \sum_{\substack{i=1 \\ i \neq \lambda}}^{m} c_{B_i} y_{ij} + c_{B_\lambda} y_{\lambda j} - c_j, \textit{where } c_k = c_{B_\lambda}. \end{aligned}$$

If $c_k$ is replaced by $c_k + \Delta c_k = c_k^*$, then $z_j^* - c_j^*$

$$\begin{aligned} &= \sum_{\substack{i=1 \\ i \neq \lambda}}^{m} c_{B_i} y_{ij} + (c_{B_\lambda} + \Delta c_{B_\lambda}) y_{\lambda j} - c_j (j \neq k) \\ &= \sum_{i=1}^{m} c_{B_i} y_{ij} + \Delta c_{B_\lambda} y_{\lambda j} - c_j \\ &= (z_j - c_j) + \Delta c_{B_\lambda} y_{\lambda j}. \end{aligned}$$

Thus, the new basic feasible solution $x_B^*$ will remain optimum so long as $\Delta c_k$ satisfies

$$\begin{aligned} & z_j - c_j + \Delta c_{B_\lambda} y_{\lambda j} \geq 0 \\ \text{or} \quad & \Delta c_{B_\lambda} y_{\lambda j} \geq -(z_j - c_j). \end{aligned}$$

When $y_{\lambda j} > 0$, $\Delta c_{B_\lambda} \geq -\frac{z_j - c_j}{y_{\lambda j}}$,

$$\text{i.e. } \Delta c_{B_\lambda} \geq \text{Max}\left\{\frac{-(z_j - c_j)}{y_{\lambda j}}, y_{\lambda j} > 0\right\}.$$

Again, when $y_{\lambda j} < 0$, $\Delta c_{B_\lambda} \leq \frac{-(z_j - c_j)}{y_{\lambda j}}$, i.e. $\Delta c_{B_\lambda} \leq \text{Min}\left\{\frac{-(z_j - c_j)}{y_{\lambda j}}, y_{\lambda j} < 0\right\}$.

Hence, $\text{Max}\left\{\frac{-(z_j - c_j)}{y_{\lambda j}}, y_{\lambda j} > 0\right\} \leq \Delta c_{B_\lambda} = \Delta c_k \leq \text{Min}\left\{\frac{-(z_j - c_j)}{y_{\lambda j}}, y_{\lambda j} < 0\right\}$ $(j \neq k)$.

*Example 1* We are given the following LPP:

$$\text{Max} \quad z = -x_1 + 2x_2 - x_3$$

subject to

$$\begin{aligned} 3x_1 + x_2 - x_3 &\leq 10 \\ -x_1 + 4x_2 + x_3 &\geq 6 \\ x_2 + x_3 &\leq 4 \\ \text{and} \quad x_j \geq 0, \; j &= 1, 2, 3. \end{aligned}$$

(a) Determine the effect of discrete changes of those components of the profit vector which correspond to the basic variables.
(b) Discuss the effect of discrete changes in those components of the profit vector which do not correspond to basic variables.

**Solution** Introducing the slack variables $x_4 \geq 0$, $x_6 \geq 0$, surplus variable $x_5 \geq 0$ and an artificial variable $x_7 \geq 0$ in the constraints of the given LPP and then solving the resulting problem by the simplex method, the following optimum simplex table is obtained:

| $c_j$ | | | $-1$ | 2 | $-1$ | 0 | 0 | 0 | $-M$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | |
| 0 | $y_4$ | $x_4 = 6$ | 3 | 0 | $-2$ | 1 | 0 | $-1$ | 0 | |
| 2 | $y_2$ | $x_2 = 4$ | 0 | 1 | 1 | 0 | 0 | 1 | 0 | |
| 0 | $y_5$ | $x_5 = 10$ | 1 | 0 | 3 | 0 | 1 | 4 | $-1$ | |
| $z_B = 8$ | | | 1 | 0 | 3 | 0 | 0 | 2 | $M$ | $\leftarrow \Delta_j$ |

(a) Here the basic variables are $x_4, x_2, x_5$. The corresponding profit coefficients are $c_4, c_2, c_5$.

We know that an optimum basic feasible solution will maintain its optimality if a change $\Delta c_k$ of $c_k \in c_B$ satisfies the following inequalities:

$$\text{Max}\left\{\frac{-(z_j - c_j)}{y_{\lambda j}}, y_{\lambda j} > 0\right\} \leq \Delta c_{B_\lambda} = \Delta c_k \leq \text{Min}\left\{\frac{-(z_j - c_j)}{y_{\lambda j}}, y_{\lambda j} < 0\right\} \quad (j \neq k).$$

For $k = 2$, i.e. $\lambda = 2$, we have

$$\text{Max}\left\{\frac{-(z_j - c_j)}{y_{2j}}, y_{2j} > 0\right\} \leq \Delta c_{B_2} = \Delta c_2 \leq \text{Min}\left\{\frac{-(z_j - c_j)}{y_{2j}}, y_{2j} < 0\right\},$$

i.e. $\text{Max}\left\{\frac{-3}{1}, \frac{-2}{1}\right\} \leq \Delta c_2$,
i.e. $-2 \leq \Delta c_2$ or $\Delta c_2 \geq -2$.

(b) Here the non-basic variables are $x_1, x_3, x_6$. Among these variables, only $x_1$ and $x_3$ are the decision variables of the given LPP. The corresponding profit coefficients are $c_1$ and $c_3$.

We know that the change $\Delta c_k$ in $c_k \in c_B$ must satisfy the upper limit $\Delta c_k \leq (z_k - c_k)$ in order to maintain the optimality of the optimum basic feasible solution.

Thus, for $k = 1, \Delta c_1 \leq z_1 - c_1$,
i.e. $\Delta c_1 \leq 1$.
For $k = 3, \Delta c_3 \leq z_3 - c_3$, i.e. $\Delta c_3 \leq 3$.

**Alternative method**

(a) Here the basic variables are $x_4$, $x_2$, $x_5$. Among these variables, only $x_2$ is the given decision variable of the given LPP. The corresponding profit coefficient is $c_2 = 2$.

Let $\Delta c_2$ be the amount by which $c_2$ is changed. Hence, the new value of $c_2$ is $c_2^* = c_2 + \Delta c_2 = 2 + \Delta c_2$.

Now for the change of $c_2 = 2$ to $c_2^*$, we have the following corresponding changed simplex table:

| $c_j$ | | | $-1$ | $2+\Delta c_2$ | $-1$ | $0$ | $0$ | $0$ | $-M$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | |
| $0$ | $y_4$ | $x_4 = 6$ | $3$ | $0$ | $-2$ | $1$ | $0$ | $-1$ | $0$ | |
| $2+\Delta c_2$ | $y_2$ | $x_2 = 4$ | $0$ | $1$ | $1$ | $0$ | $1$ | $1$ | $0$ | |
| $0$ | $y_5$ | $x_5 = 10$ | $1$ | $0$ | $3$ | $0$ | $0$ | $4$ | $-1$ | |
| | | | $1$ | $0$ | $3+\Delta c_2$ | $0$ | $0$ | $2+\Delta c_2$ | $M$ | $\leftarrow \Delta_j$ |

The current solution $x_B$ remains optimum for the new problem
if $3 + \Delta c_2 \geq 0$ and $2 + \Delta c_2 \geq 0$,
i.e. if $\Delta c_2 \geq -3$ and $\Delta c_2 \geq -2$,
i.e. if $\Delta c_2 \geq \text{Max}\{-3, -2\}$,
i.e. if $\Delta c_2 \geq -2$.

(b) Here the non-basic variables are $x_1, x_3, x_6$. Among these, only $x_1$ and $x_3$ are the given decision variables. The corresponding profit coefficients are $c_1$ and $c_3$. Let $\Delta c_1$ be the amount by which $c_1$ is changed. Hence, the new value of $c_1$ is $c_1^* = c_1 + \Delta c_1 = -1 + \Delta c_1$.

Now for the change of $c_1 = -1$ to $c_1^*$, we have the following corresponding changed simplex table:

| $c_j$ | | | $-1+\Delta c_1$ | 2 | $-1$ | 0 | 0 | 0 | $-M$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | |
| 0 | $y_4$ | $x_4 = 6$ | 3 | 0 | $-2$ | 1 | 0 | $-1$ | 0 | |
| 2 | $y_2$ | $x_2 = 4$ | 0 | 1 | 1 | 0 | 0 | 1 | 0 | |
| 0 | $y_5$ | $x_5 = 10$ | 1 | 0 | 3 | 0 | 1 | 4 | $-1$ | |
| | | | $1-\Delta c_1$ | 0 | 3 | 0 | 0 | 2 | $M$ | $\leftarrow \Delta_j$ |

The current solution remains optimum for the new problem if $1 - \Delta c_1 \geq 0$, i.e. if $\Delta c_1 \leq 1$.

Similarly, the range of $\Delta c_3$ can be obtained.

*Example 2* Given the LPP

$$\begin{aligned} &\text{Max} \quad z = 15x_1 + 45x_2 \\ &\text{subject to} \\ &\qquad x_1 + 16x_2 \leq 240 \\ &\qquad 5x_1 + 2x_2 \leq 162 \\ &\qquad x_2 \leq 50 \text{ and } x_1, x_2 \geq 0, \end{aligned}$$

if Maximize $z = \sum_j c_j x_j$, $j = 1, 2$ and $c_2$ is kept fixed at 45, determine how much $c_1$ can be changed without affecting the optimal solution.

**Solution** By introducing the slack variables $x_3 \geq 0, x_4 \geq 0$ and $x_5 \geq 0$ in the constraints of the given LPP and then solving the resulting problem by the simplex method, the following optimum simplex table is obtained:

| $c_j$ | | | 15 | 45 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 45 | $y_2$ | $x_2 = 173/13$ | 0 | 1 | 5/78 | +1/78 | 0 | |
| 15 | $y_1$ | $x_1 = 352/13$ | 1 | 0 | −1/39 | 8/39 | 0 | |
| 0 | $y_5$ | $x_5 = 477/13$ | 0 | 0 | -5/78 | 1/78 | 1 | |
| $z_B = 1005$ | | | 0 | 0 | 5/2 | 5/2 | 0 | $\leftarrow \Delta_j$ |

From this optimum table, it is seen that $y_1$, $y_2$, $y_5$ are the basic vectors. Let $c_1$ be changed to $c_1 + \Delta c_1$.

Since $c_2$ is kept fixed and $c_1$ can vary, an optimum basic feasible solution will maintain its optimality if $\Delta c_1$ in $c_1 \in c_B$ satisfies

$$\text{Max}\left\{\frac{-(z_j-c_j)}{y_{2j}}, y_{2j}>0\right\} \le \Delta c_{B_2} = \Delta c_1 \le \text{Min}\left\{\frac{-(z_j-c_j)}{y_{2j}}, y_{2j}<0\right\}, \quad \text{here } \lambda = 2$$

$$\text{or} \quad \text{Max}\left\{\frac{-(z_4-c_4)}{y_{24}}\right\} \le \Delta c_1 \le \text{Min}\left\{\frac{-(z_3-c_3)}{y_{23}}\right\}$$

$$\text{or} \quad \text{Max}\left\{\frac{-5/2}{8/39}\right\} \le \Delta c_1 \le \text{Min}\left\{\frac{-5/2}{-1/39}\right\}$$

$$\text{or} \quad -\frac{195}{16} \le \Delta c_1 \le \frac{195}{2}.$$

*Example 3* We are given the following LPP:

$$\text{Max}\, z = 3x_1 + 5x_2$$

subject to $3x_1 + 2x_2 \le 18$

$$x_1 \le 2$$
$$x_2 \le 6 \text{ and } x_1,\ x_2 \ge 0.$$

Discuss the effect on the optimality of the solution, where the objective function is changed to $3x_1 + x_2$.

**Solution** After introducing the slack variables $x_3 \ge 0$, $x_4 \ge 0$ and $x_5 \ge 0$ in the constraints of the preceding LPP and then solving the reformulated LPP by the simplex method, the optimum simplex table is as follows:

| $c_j$ | | | 3 | 5 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 3 | $y_1$ | $x_1 = 2$ | 1 | 0 | 1/3 | 0 | −2/3 | |
| 0 | $y_4$ | $x_4 = 0$ | 0 | 0 | −2/3 | 1 | 4/3 | |
| 5 | $y_2$ | $x_2 = 6$ | 0 | 1 | 0 | 0 | 1 | |
| $z_B = 36$ | | | 0 | 0 | 1 | 0 | 3 | $\leftarrow \Delta_j$ |

Since the objective function is now changed to $z = 3x_1 + x_2$, the new $c_B$ becomes (3 0 1).

Hence, the new net evaluations corresponding to the variables $x_3,\ x_4,\ x_5$ become

$$\Delta_3 = 3 \times \frac{1}{3} + 0 \times \frac{-2}{3} + 1 \times 0 - 0 \qquad = 1$$
$$\Delta_4 = 3 \times 0 + 0 \times 1 + 1 \times 0 - 0 \qquad = 0$$
$$\Delta_5 = 3 \times \left(-\frac{2}{3}\right) + 0 \times \frac{4}{3} + 1 \times 1 - 0 = -1.$$

This indicates that $y_5$ enters the basis in the next iteration.

The new simplex table is the following:

| $c_j$ | | | 3 | 1 | 0 | 0 | 0 | Mini ratio |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 3 | $y_1$ | $x_1 = 2$ | 1 | 0 | 1/3 | 0 | −2/3 | – |
| 0 | $y_4$ | $x_4 = 0$ | 0 | 0 | −2/3 | 1 | 4/3 | 0/(4/3) = 0 |
| 1 | $y_2$ | $x_2 = 6$ | 0 | 1 | 0 | 0 | 1 | 6/1 = 6 |
| $z_B = 12$ | | | 0 | 0 | 1 | 0 | −1 | $\leftarrow \Delta_j$ |

Here $y_4$ is the outgoing vector, and $y_{25} = \frac{4}{3}$ is the key element.

Now we construct the next table:

| $c_j$ | | | 3 | 1 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 3 | $y_1$ | $x_1 = 2$ | 1 | 0 | 0 | 1/2 | 0 | |
| 0 | $y_5$ | $x_5 = 0$ | 0 | 0 | −1/2 | 3/4 | 1 | |
| 1 | $y_2$ | $x_2 = 6$ | 0 | 1 | 1/2 | −3/4 | 0 | |
| $z_B = 12$ | | | 0 | 0 | 1/2 | 3/4 | 0 | $\leftarrow \Delta_j$ |

Here all $\Delta_j \geq 0$. Hence, the solution is optimal, and it is $x_1 = 2$, $x_2 = 6$ and Max $z = 12$.

*Example 4* We are given the LPP

$$\begin{aligned} \text{Maximize } & z = 6x_1 - 2x_2 + 3x_3 \\ \text{subject to } & 2x_1 - x_2 + 2x_3 \leq 2 \\ & x_1 + 4x_3 \leq 4 \\ & \text{and } x_1, x_2, x_3 \geq 0. \end{aligned}$$

Find the ranges of the profit components when (i) $c_1$ and $c_2$ are the two profit components changed at the same time, (ii) all three profit components $c_1$, $c_2$ and $c_3$ are changed at a time to keep the optimal solution the same.

**Solution** Introducing the slack variables $x_4 \geq 0$ and $x_5 \geq 0$ in the constraints of the given LPP and then solving the resulting problem by the simplex method, the following optimal simplex table is obtained:

| | $c_j$ | | 6 | −2 | 3 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| 6 | $y_1$ | $x_1 = 4$ | 1 | 0 | 4 | 0 | 1 | |
| -2 | $y_2$ | $x_2 = 6$ | 0 | 1 | 6 | −1 | 2 | |
| | | | 0 | 0 | 9 | 2 | 2 | $\leftarrow \Delta_j$ |

(i) When two profit components $c_1$ and $c_2$ are changed at a time, for the change of $c_1 = 6$ and $c_2 = -2$ to $c_1^*$ and $c_2^*$, the modified simplex table for the new problem is as follows:

| $c_j$ | | | $c_2^*$ | $c_2^*$ | 3 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| $c_1^*$ | $y_1$ | $x_1 = 4$ | 1 | 0 | 4 | 0 | 1 | |
| $c_2^*$ | $y_2$ | $x_2 = 6$ | 0 | 1 | 6 | −1 | 2 | |
| | | | 0 | 0 | $4c_1^* + 6c_2^* - 3$ | $-c_2^*$ | $c_1^* + 2c_2^*$ | $\leftarrow \Delta_j$ |

The current solution remains optimum for the new problem if $z_j - c_j \geq 0$ for all $j$,
i.e. if $4c_1^* + 6c_2^* - 3 \geq 0$, $-c_2^* \geq 0$ and $c_1^* + 2c_2^* \geq 0$,
i.e. if $c_2^* \geq \frac{3-4c_1^*}{6}$, $c_2^* \leq 0$ and $c_2^* \geq -\frac{c_1^*}{2}$,
i.e. if $\text{Max}\left\{\frac{3-4c_1^*}{6}, \frac{-c_1^*}{2}\right\} \leq c_2^* \leq 0$ and $c_1^*$ is any real number.

(ii) When all the three components $c_1$, $c_2$ and $c_3$ are changed at a time, for the change of $c_1 = 6$, $c_2 = -2$ and $c_3 = 3$ to $c_1^*$, $c_2^*$ and $c_3^*$, the modified simplex table for the new problem is as follows:

| $c_j$ | | | $c_2^*$ | $c_2^*$ | $c_3^*$ | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | |
| $c_1^*$ | $y_1$ | $x_1 = 4$ | 1 | 0 | 4 | 0 | 1 | |
| $c_2^*$ | $y_2$ | $x_2 = 6$ | 0 | 1 | 6 | −1 | 2 | |
| | | | 0 | 0 | $4c_1^* + 6c_2^* - c_3^*$ | $-c_2^*$ | $c_1^* + 2c_2^*$ | $\leftarrow \Delta_j$ |

The current solution remains optimum for the new problem if $z_j - c_j \geq 0$ for all $j$,
i.e. if $4c_1^* + 6c_2^* - c_3^* \geq 0$, $-c_2^* \geq 0$ and $c_1^* + 2c_2^* \geq 0$,
i.e. if $c_2^* \geq \frac{c_3^*-4c_1^*}{6}$, $c_2^* \leq 0$ and $c_2^* \geq -\frac{c_1^*}{2}$,
i.e. if $\text{Max}\left\{\frac{c_3^*-4c_1^*}{6}, \frac{-c_1^*}{2}\right\} \leq c_2^* \leq 0$ and $c_1^*$, $c_3^*$ are any real numbers.

## 7.4 Discrete Changes in the Requirement Vector

Let $x_B$ be the optimum basic feasible solution of the following LPP:

$$\text{Maximize}\, z = \langle c^T, x \rangle$$
$$\text{subject to}\, Ax = b \;\text{ and }\; x \geq 0$$

where $c$, $x^T \in R^n, b^T \in R^m$ and $A$ is an $m \times n$ real matrix of rank $m$.

Let $\Delta b_k$ be the amount which is added to $b_k$, and the new value of $b_k$ is $b'_k = b_k + \Delta b_k, k = 1, 2, \ldots, m$.

Let the new basic solution be $x'_B$.

Then $x'_B = B^{-1}b'$, where $b' = \begin{bmatrix} b_1 \\ b_2 \\ \ldots \\ b_k + \Delta b_k \\ \ldots \\ b_m \end{bmatrix}$

$$= B^{-1}\begin{bmatrix} b_1 \\ b_2 \\ \ldots \\ \ldots \\ b_m \end{bmatrix} + B^{-1}\begin{bmatrix} 0 \\ 0 \\ \ldots \\ \Delta b_k \\ \ldots \\ 0 \end{bmatrix}$$

$$= x_B + \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1k} & \ldots & b_{1m} \\ b_{21} & b_{22} & \ldots & b_{2k} & \ldots & b_{2m} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ b_{m1} & b_{m2} & \ldots & b_{mk} & \ldots & b_{mm} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ \ldots \\ \Delta b_k \\ \ldots \\ 0 \end{bmatrix}$$

$$\text{or} \quad \begin{bmatrix} x'_{B_1} \\ x'_{B_2} \\ \ldots \\ x'_{B_i} \\ \ldots \\ x'_{B_m} \end{bmatrix} = \begin{bmatrix} x_{B_1} \\ x_{B_2} \\ \ldots \\ x_{B_i} \\ \ldots \\ x_{B_m} \end{bmatrix} + \begin{bmatrix} b_{1k}\,\Delta b_k \\ b_{2k}\,\Delta b_k \\ \ldots \\ b_{ik}\,\Delta b_k \\ \ldots \\ b_{mm}\,\Delta b_k \end{bmatrix}.$$

$\therefore x'_{B_i} = x_{B_i} + b_{ik}\Delta b_k$, where $b_{ik}$ is the ($i$, $k$)th element.

For maintaining the feasibility of the basic solution, we must have

$$\begin{aligned} x'_B \geq 0 \text{ i.e. } x'_{B_i} &\geq 0 \ \forall i = 1, 2, \ldots, m \\ \text{i.e. } x_{B_i} + b_{ik}\Delta b_k &\geq 0 \ \forall i = 1, 2, \ldots, m \\ \text{i.e. } b_{ik}\Delta b_k &\geq -x_{B_i}. \end{aligned}$$

When $b_{ik} > 0, \Delta b_k \geq -\frac{x_{B_i}}{b_{ik}}, \forall i = 1, 2, \ldots, m,$

i.e. $\text{Max}\left\{\frac{-x_{B_i}}{b_{ik}}, b_{ik} > 0\right\} \leq \Delta b_k.$

Similarly, for $b_{ik} < 0, \Delta b_k \leq \text{Min}\left\{\frac{-x_{B_i}}{b_{ik}}, b_{ik} < 0\right\} \forall i = 1, 2, \ldots, m.$

Hence, $\text{Max}\left\{\frac{-x_{B_i}}{b_{ik}}, b_{ik} > 0\right\} \leq \Delta b_k \leq \text{Min}\left\{\frac{-x_{B_i}}{b_{ik}}, b_{ik} > 0\right\}, \forall i = 1, 2, \ldots, m.$

We observe that $x'_B$ remains optimal since the optimality condition $\Delta_j \geq 0$ remains unaffected by the change of $b_k$.

*Example 5* We are given the following LPP:

$$\begin{aligned} &\text{Maximize } z = -x_1 + 2x_2 - x_3 \\ &\text{subject to } 3x_1 + x_2 - x_3 \leq 10 \\ &\qquad -x_1 + 4x_2 + x_3 \geq 6 \\ &\qquad x_2 + x_3 \leq 4 \text{ and } x_j \geq 0, \quad j = 1, 2, 3. \end{aligned}$$

(a) Determine the ranges for discrete changes in the components $b_2$ and $b_3$ of the requirement vector so as to maintain the optimality of the current optimum solution.

**Solution** Introducing slack variables $x_4$ and $x_6$, surplus variable $x_5$ and artificial variable $x_7$, the given LPP can be written as

$$\text{Max } z = -x_1 + 2x_2 - x_3 + 0.x_4 + 0.x_5 + 0.x_6 - Mx_7$$

$$\begin{aligned} &\text{subject to } 3x_1 + x_2 - x_3 + x_4 = 10 \\ &\qquad -x_1 + 4x_2 + x_3 - x_5 + x_7 = 6 \\ &\qquad x_2 + x_3 + x_6 = 4. \end{aligned}$$

The initial simplex table is the following:

| $c_j$ | | | $-1$ | $2$ | $-1$ | $0$ | $0$ | $0$ | $-M$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | |
| 0 | $y_4$ | $x_4 = 10$ | 3 | 1 | -1 | 1 | 0 | 0 | 0 | |
| $-M$ | $y_7$ | $x_7 = 6$ | $-1$ | 4 | 1 | 0 | $-1$ | 0 | 1 | |

(continued)

(continued)

| $c_j$ | | | $-1$ | $2$ | $-1$ | $0$ | $0$ | $0$ | $-M$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | |
| 0 | $y_6$ | $x_6 = 4$ | 0 | 1 | 1 | 0 | 0 | 1 | 0 | |
| $z_B = -6M$ | | | $M-1$ | $-4M-2$ | $-M+1$ | 0 | $M$ | 0 | 0 | $\leftarrow \Delta_j$ |

The optimum table is as follows:

| $c_j$ | | | $-1$ | $2$ | $-1$ | $0$ | $0$ | $0$ | $-M$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | |
| 0 | $y_4$ | $x_4 = 6$ | 3 | 0 | $-2$ | 1 | 0 | $-1$ | 0 | |
| 2 | $y_2$ | $x_2 = 4$ | 0 | 1 | 1 | 0 | 0 | 1 | 0 | |
| 0 | $y_5$ | $x_5 = 10$ | 1 | 0 | 3 | 0 | 1 | 4 | $-1$ | |
| $z_B = 8$ | | | 1 | 0 | 3 | 0 | 0 | 2 | $M$ | $\leftarrow \Delta_j$ |

From the optimum simplex table, we have

$$x_B = [6, 4, 10], \quad b = [10, 6, 4]$$

$$B^{-1} = [y_4, y_7, y_6] = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & -1 & 4 \end{pmatrix}.$$

The individual effects of changes in $b_2$ and $b_3$ where $b = [b_1, b_2, b_3]$ without violating the optimality of the basic feasible solution are given by

$$\text{Max}\left\{\frac{-x_{B_i}}{b_{ik}}, b_{ik} > 0\right\} \le \Delta b_k \le \text{Min}\left\{\frac{-x_{B_i}}{b_{ik}}, b_{ik} < 0\right\}, k = 2, 3.$$

When $k = 2$, then

$$\text{Max}\left\{\frac{-x_{B_i}}{b_{i2}}, b_{i2} > 0\right\} \le \Delta b_2 \le \text{Min}\left\{\frac{-x_{B_i}}{b_{i2}}, b_{i2} < 0\right\}$$

$$\text{or} \quad \Delta b_2 \le \text{Min}\left\{-\frac{x_{B_3}}{b_{32}}\right\}$$

$$\text{or} \quad \Delta b_2 \le \text{Min}\left\{\frac{-10}{-1}\right\} \quad \text{or} \quad \Delta b_2 \le 10.$$

Again, when $k = 3$, then

$$\begin{aligned}
&\text{Max}\left\{\frac{-x_{B_i}}{b_{i3}}, b_{i3} > 0\right\} \le \Delta b_3 \le \text{Min}\left\{\frac{-x_{B_i}}{b_{i3}}, b_{i3} < 0\right\} \\
\text{or}\quad &\text{Max}\left\{\frac{-x_{B_2}}{b_{23}}, \frac{-x_{B_3}}{b_{33}}\right\} \le \Delta b_3 \le \text{Min}\left\{\frac{-x_{B_i}}{b_{13}}\right\} \\
\text{or}\quad &\text{Max}\left\{\frac{-4}{1}, \frac{-10}{4}\right\} \le \Delta b_3 \le \text{Min}\left\{\frac{-6}{-1}\right\} \\
\text{or}\quad &\text{Max}\{-4, -5/2\} \le \Delta b_3 \le 6 \\
\text{or}\quad &-5/2 \le \Delta b_3 \le 6.
\end{aligned}$$

*Example 6* Consider the LPP

$$\begin{aligned}
&\text{Maximize } z = 2x_1 + x_2 + 4x_3 - x_4 \\
&\quad \text{subject to} \\
&\qquad x_1 + 2x_2 + x_3 - 3x_4 \le 8 \\
&\qquad -x_2 + x_3 + 2x_4 \le 0 \\
&\qquad 2x_1 + 7x_2 - 5x_3 - 10x_4 \le 21 \\
&\qquad x_1, x_2, x_3, x_4 \ge 0
\end{aligned}$$

The optimal simplex table of this LPP is

| | | | 2 | 1 | 4 | −1 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ |
| 2 | $y_1$ | 8 | 1 | 0 | 3 | 1 | 1 | 2 | 0 |
| 1 | $y_2$ | 0 | 0 | 1 | −1 | −2 | 0 | −1 | 0 |
| 0 | $y_7$ | 5 | 0 | 0 | −4 | 2 | −2 | 3 | 1 |
| $z_B = 16$ | | | 0 | 0 | 1 | 1 | 2 | 3 | 0 |

When $b$ is changed to $[3 \quad -2 \quad 4]^{\mathrm{T}}$, make the necessary corrections in the optimal simplex table and solve the resulting problem.

**Solution** From the given optimal simplex table, we have

$$x_B = [8 \quad 0 \quad 5]^{\mathrm{T}}, \quad b = [8 \quad 0 \quad 21]^{\mathrm{T}}$$

$$B^{-1} = [y_5 \quad y_6 \quad y_7] = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & 0 \\ -2 & 3 & 1 \end{bmatrix}.$$

When $b$ is changed to $[3 \quad -2 \quad 4]^{\mathrm{T}}$, the new values of the current basic variables are given by

$$\begin{bmatrix} x_1 \\ x_2 \\ x_7 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & 0 \\ -2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ -8 \end{bmatrix}.$$

Since the values of $x_1$ and $x_7$ are negative, the current basic solution becomes infeasible. Hence, we have to apply the dual simplex method to obtain the basic feasible solution.

Now we construct the next dual simplex table.

| $c_j$ | | | 2 | 1 | 4 | −1 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ |
| 2 | $y_1$ | −1 | 1 | 0 | 3 | 1 | 1 | 2 | 0 |
| 1 | $y_2$ | 2 | 0 | 1 | −1 | −2 | 0 | −1 | 0 |
| 0 | $y_7$ | −8 | 0 | 0 | −4 | 2 | −2 | 3 | 1 |
| $z_B = 0$ | | | 0 | 0 | 1 | 1 | 2 | 3 | 0 |

Now $x_{B_r} = \min\{x_{B_i}, x_{B_i} < 0\} = \min\{-1, -8\} = -8 = x_{B_3}$.
$\therefore r = 3$.
Hence, the vector $y_{B_3}$, i.e. $y_7$, has the basis.

$$\begin{aligned} &\text{Again, Max}\left\{\frac{z_j - c_j}{y_{r_j}};\ y_{r_j} < 0\right\} \\ &= \text{Max}\left\{\frac{z_j - c_j}{y_{3j}};\ y_{3j} < 0\right\} = \text{Max}\left\{\frac{z_3 - c_3}{y_{33}}, \frac{z_5 - c_5}{y_{35}}\right\} \\ &= \text{Max}\left\{\frac{1}{-4}, \frac{2}{-2}\right\} = -\frac{1}{4}, \text{ which occurs for } j = 3. \end{aligned}$$

Hence, the vector $y_3$ enters the basis, and $y_{rj}$, i.e. $y_{33} = -4$, is the key element. Now we construct the next dual simplex table.

| $c_j$ | | | 2 | 1 | 4 | −1 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ |
| 2 | $y_1$ | −7 | 1 | 0 | 0 | $\frac{5}{2}$ | $-\frac{1}{2}$ | $\frac{17}{4}$ | $\frac{3}{4}$ |
| 1 | $y_2$ | 4 | 0 | 1 | 0 | $-\frac{5}{2}$ | $\frac{1}{2}$ | $-\frac{7}{4}$ | $-\frac{1}{4}$ |
| 4 | $y_3$ | 2 | 0 | 0 | 1 | $-\frac{1}{2}$ | $\frac{1}{2}$ | $-\frac{3}{4}$ | $-\frac{1}{4}$ |
| $z_B = -2$ | | | 0 | 0 | 0 | 3/2 | 3/2 | 15/4 | 1/4 |

| $c_j$ | | | 2 | 1 | 4 | −1 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ |
| 0 | $y_5$ | 14 | −2 | 0 | 0 | −5 | 1 | $-\frac{17}{2}$ | $-\frac{3}{2}$ |
| 1 | $y_2$ | −3 | 1 | 1 | 0 | 0 | 0 | $\frac{5}{2}$ | $\frac{1}{2}$ |
| 4 | $y_3$ | −5 | 1 | 0 | 1 | 2 | 0 | $\frac{7}{2}$ | $\frac{1}{2}$ |
| $z_B = -23$ | | | 3 | 0 | 0 | 9 | 0 | $\frac{33}{2}$ | $\frac{5}{2}$ |

This table shows that $y_3$ leaves the basis, but since all the elements of the departing row are non-negative, by the dual simplex method, the given problem does not possess any feasible solution.

## 7.5 Discrete Changes in the Coefficient Matrix

Here, we shall discuss the post-optimality problems which involve discrete changes in the element of the coefficient matrix $A$.

Let us consider the following LPP:

$$\begin{aligned}&\text{Maximize}\, z = \langle c^T, x\rangle\\&\text{subject to}\, Ax = b \,\text{and}\, x \geq 0,\end{aligned}$$

where $c$, $x^T \in R^n, b^T \in R^m$ and $A$ is an $m \times n$ real matrix of rank $m$.

Let $x_B$ be the optimum basic feasible solution of that LPP. Again, let us assume that the element $a_{rk}$ in the $r$th row and $k$th column of matrix $A$ is changed to $a'_{rk}(= a_{rk} + \Delta a_{rk})$.

Two possibilities may arise:

Case I: $a_k \notin B$
Case II: $a_k \in B$, where $B$ is the basis matrix.
**Case I**: In this case, the post-optimal change does not have any effect on the feasibility of the current optimal solution $x_B = B^{-1}b$. Thus, the only effect, if any, may be on the optimality of the current solution.

Let $B^{-1} = (\beta_1, \beta_2, \ldots, \beta_m)$ and $a_k = (a_{1k}, a_{2k}, \ldots, a_{mk})^{\mathrm{T}}$.

$$\begin{aligned}\text{Therefore, } a'_k &= (a_{1k}, a_{2k}, \ldots, a_{rk} + \Delta a_{rk}, \ldots, a_{mk})^{\mathrm{T}}\\&= (a_{1k}, a_{2k}, \ldots, a_{rk}, \ldots, a_{mk})^{\mathrm{T}} + (0, 0, \ldots, \Delta a_{rk}, \ldots, 0)^{\mathrm{T}}\\&= a_k + (0, 0, \ldots, \Delta a_{rk}, \ldots, 0)^{\mathrm{T}}\end{aligned}$$

Now, to preserve the optimality of the optimal solution of the original LPP, we must have $z_k^* - c_k \geq 0$.

$$
\begin{aligned}
\text{where } z_k^* &= \left\langle c_B^T, y_k^* \right\rangle = \left\langle c_B^T, B^{-1}a_k^* \right\rangle \\
&= \left\langle c_B^T, B^{-1}\left[a_k + (0,0,\ldots,\Delta a_{rk},\ldots,0)^{\mathrm{T}}\right] \right\rangle \\
&= \left\langle c_B^T, B^{-1}a_k \right\rangle + \left\langle c_B^T, B^{-1}(0,0,\ldots,\Delta a_{rk},\ldots,0)^{\mathrm{T}} \right\rangle \\
&= z_k + \left\langle c_B^T, (\beta_1,\beta_2,\ldots\beta_r,\ldots,\beta_m)(0,0,\ldots,\Delta a_{rk},\ldots,0)^{\mathrm{T}} \right\rangle \\
&= z_k + \left\langle c_B^T, \beta_r\ \Delta a_{rk} \right\rangle \\
\therefore z_k^* - c_k &= z_k - c_k + \left\langle c_B^T, \beta_r\ \Delta a_{rk} \right\rangle = z_k - c_k + \left\langle c_B^T, \beta_r \right\rangle \Delta a_{rk}.
\end{aligned}
$$

Now, from the optimality condition, we have

$$
\begin{aligned}
& z_k^* - c_k \geq 0 \\
& \Rightarrow z_k - c_k + \left\langle c_B^T, \beta_r \right\rangle \Delta a_{rk} \geq 0 \\
& \Rightarrow \left\langle c_B^T, \beta_r \right\rangle \Delta a_{rk} \geq -(z_k - c_k) \\
& \Rightarrow \begin{bmatrix}
\Delta a_{rk} \geq \frac{-(z_k - c_k)}{\left\langle c_B^T, \beta_r \right\rangle} & \text{for } \left\langle c_B^T, \beta_r \right\rangle > 0 \\
\Delta a_{rk} \leq \frac{-(z_k - c_k)}{\left\langle c_B^T, \beta_r \right\rangle} & \text{for } \left\langle c_B^T, \beta_r \right\rangle < 0 \\
\Delta a_{rk} \text{ is unrestrieted} & \text{for } \left\langle c_B^T, \beta_r \right\rangle = 0
\end{bmatrix}
\end{aligned}
$$

**Case II**: In this case, $a_k \in B$, i.e. $a_k$ is a basis vector. So, the feasibility of the current optimum solution may also be destroyed. Therefore, in this case, we have to find a range for discrete changes $\Delta a_{rk}$ in $a_{rk}$ so as to maintain the feasibility as well as the optimality of the solution. Now, we shall find the conditions for maintaining feasibility of the solution.

Let $B = (B_1, B_2, \ldots, B_m)$, $B_j \in A$; $j = 1, 2, \ldots, m$ and $B^{-1} = (\beta_1, \beta_2, \ldots, \beta_m)$, where $\beta_j$ is the $j$th column vector of $B^{-1}$.

Since $a_k \in B$, any change in $a_{rk}$ will affect only the $p$th column of $B$. Then the basis matrix $B^*$ becomes $B^* = \left(B_1, B_2, \ldots, B_p^*, \ldots, B_m\right)$,

where $B_p^* = (B_{1p}, B_{2p}, \ldots, B_{rp} + \Delta a_{rp}, \ldots, B_{mp})^T$.

Since $B_p^* \in B^*$, it can be expressed as the linear combination for vectors $B_1, B_2, \ldots, B_p, \ldots, B_m$ in $B$; therefore, $B_p^* = \lambda_1 B_1 + \lambda_2 B_2 + \cdots + \lambda_m B_m = B\lambda$ where $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)^{\mathrm{T}}$.

$$\therefore \lambda = B^{-1}B_p^* = B^{-1}[B_p + (0,0,\ldots,\Delta a_{rp},\ldots,0)^{\mathrm{T}}]$$
$$= B^{-1}B_p^+ \beta_r \Delta a_{rk}$$
$$= e_p + \beta_r \Delta a_{rk} \quad \begin{bmatrix} \text{since } B^{-1}B_p = e_p \\ \text{as } Be_p = \big(B_1, B_2, \ldots, B_p, \ldots, B_m\big)(0,0,\ldots,1,\ldots,0) = B_p \end{bmatrix},$$

i.e. $(\lambda_1, \lambda_2, \ldots, \lambda_p, \ldots, \lambda_m)^{\mathrm{T}} = (0,0,\ldots,1,\ldots,0)^{\mathrm{T}}(\beta_{1r}, \beta_{2r}, \ldots, \beta_{pr}, \ldots, \beta_{mr})^{\mathrm{T}}\Delta a_{rk}$.

Now, equating the $p$th element and $i$th ($i \neq p$) element of both sides, we have $\lambda_p = 1 + \beta_{pr}\Delta a_{rk}$ and $\lambda_{\mathrm{i}} = \beta_{ir}\Delta a_{rk}$ for $i \neq p$.

Now, $B^{*-1}$ exists only if $B^*$ is the non-singular (i.e., $|B^*| \neq 0$).

Therefore, we must have $\lambda_p \neq 0$; i.e. $1 + \beta_{pr}\Delta a_{rk} \neq 0$.

Now, we shall introduce a new matrix $E$ which differs from an identity matrix $I_m$ in the $p$th column only.

The $p$th column of matrix $E$ can be obtained as follows:

$$B_p^* = \lambda_1 B_1 + \lambda_2 B_2 + \cdots + \lambda_p B_p + \cdots + \lambda_m B_m$$
$$\text{or} \quad B_p = -\frac{\lambda_1}{\lambda_p}B_1 - \frac{\lambda_2}{\lambda_p}B_2 - \ldots + \frac{1}{\lambda_p}B_p^* \ldots - \frac{\lambda_m}{\lambda_p}B_m$$
$$= \big(B_1, B_2, \ldots, B_P^*, \ldots, B_m\big)\left(\frac{-\lambda_1}{\lambda_p}, \frac{-\lambda_2}{\lambda_p}, \ldots + \frac{1}{\lambda_p}, \ldots, \frac{-\lambda_m}{\lambda_p}\right)^T = B^*\eta,$$

where $\eta = \big(-\lambda_1/\lambda_p, -\lambda_2/\lambda_p, \ldots, 1/\lambda_p, -\lambda_m/\lambda_p\big)^T$ = the $p$th column of $E$.

$$\text{Hence, } E = \begin{pmatrix} 1 & 0 & \ldots & \ldots & \ldots & -\lambda_1/\lambda_p & \ldots & \ldots & \ldots & 0 \\ 0 & 1 & \ldots & \ldots & \ldots & -\lambda_2/\lambda_p & \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & \ldots & 1/\lambda_p & \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & \ldots & -\lambda_m/\lambda_p & \ldots & \ldots & \ldots & 1 \end{pmatrix}.$$

Therefore, $B^*E = B^*(e_1, e_2, \ldots, \eta, \ldots, e_m) = (B^*e_1, B^*e_2, \ldots, B^*\eta, \ldots, B^*e_m)$

$$= \big(B_1, B_2, \ldots, B_p, \ldots, B_m\big)$$
$$\Big[\text{as } B^*e_1 = \Big(B_1, B_2, B_p^*, \ldots, B_m\Big)(1,0,\ldots,0)^{\mathrm{T}} = B_1$$
$$\text{and } B^*\eta = B_p\Big].$$

$$\begin{aligned}
&\therefore B^*E = B\\
&\Rightarrow (B^*E)B^{-1} = BB^{-1}\\
&\Rightarrow B^*\left(EB^{-1}\right) = I\\
&\Rightarrow (B^*)^{-1} = EB^{-1.}
\end{aligned}$$

Hence, the new solution becomes

$$\begin{aligned}
x_B^* &= (B^*)^{-1}b = (EB^{-1})b\\
&= E(B^{-1}b)\\
&= Ex_B \quad \left[\because x_B = B^{-1}b\right]
\end{aligned}$$

$$\begin{aligned}
&\Rightarrow x_B^* = Ex_B\\
&\Rightarrow \left(x'_{B1}, x'_{B2}, \ldots, x'_{Bp}, \ldots, x'_{Bm}\right) = E\left(x_{B1}, x_{B2}, \ldots, x_{Bp}, \ldots, x_{Bm}\right)^T\\
&\Rightarrow x'_{Bi} = \begin{cases} x_{Bi} - \frac{\lambda_i}{\lambda_p}x_{Bp} & \text{for } i \neq p\\ \frac{x_{Bp}}{\lambda_p} & \text{for } i = p \end{cases}\\
&\text{or } x'_{Bi} = \begin{cases} x_{Bi} - \frac{\beta_{ir}\Delta a_{rk}}{1+\beta_{pr}\Delta a_{rk}}x_{Bp} & \text{for } i \neq p, i = 1, 2, \ldots, m\\ \frac{x_{Bp}}{1+\beta_{pr}\Delta a_{rk}} & \text{for } i = p \end{cases}
\end{aligned}$$

Thus, to maintain the feasibility of $x'_B$, we must have $x'_{Bi} \geq 0$ for $i \neq p$ and $x'_{Bp} \geq 0$.

$$\text{Now}, x'_{Bp} \geq 0 \Rightarrow \frac{x_{Bp}}{1+\beta_{pr}\Delta a_{rk}} \geq \quad \Rightarrow 1 + \Delta\beta_{pr}\Delta a_{rk} > 0 \text{ if } x_{Bp} \neq 0.$$

$$\begin{aligned}
&\text{Again}, x'_{Bi} \geq 0 (i \neq p) \Rightarrow x_{Bi} - \frac{\beta_{ir}\Delta a_{rk}}{1+\beta_{pr}\Delta a_{rk}}x_{Bp} \geq 0\\
&\Rightarrow x_{Bi} + \Delta a_{rk}(\beta_{pr}x_{Bi} - \beta_{ir}x_{Bp}) \geq 0\\
&\Rightarrow \Delta a_{rk}(\beta_{pr}x_{Bi} - \beta_{ir}x_{Bp}) \geq -x_{Bi}\\
&\Rightarrow \begin{cases} \Delta a_{rk} \geq -\dfrac{x_{Bi}}{\beta_{pr}x_{Bi} - \beta_{ir}x_{Bp}}, \text{for } \beta_{pr}x_{Bi} - \beta_{ir}x_{Bp} > 0\\[2ex] \Delta a_{rk} \leq -\dfrac{x_{Bi}}{\beta_{pr}x_{Bi} - \beta_{ir}x_{Bp}}, \text{for } \beta_{pr}x_{Bi} - \beta_{ir}x_{Bp} < 0 \end{cases}
\end{aligned}$$

Hence, in order to maintain the feasibility, combining these two inequalities, we have

$$\text{Max}\left\{\frac{-x_{Bi}}{Q_i}, Q_i > 0\right\} \le \Delta a_{rk} \le \text{Min}\left\{\frac{-x_{Bi}}{Q_i}, Q_i < 0\right\},$$

where $Q_i = \beta_{pr}x_{Bi} - \beta_{ir}x_{Bp}$.

The preceding inequalities give a range for discrete changes $\Delta a_{rk}$ in the coefficient matrix $A$ to maintain the feasibility of the solution.

**Condition for maintaining the optimality of the solution**

To maintain the optimality of the new solution, we must have

$z_j^* - c_j \ge 0$ for all $j = 1, 2, \ldots, n$.

$$\begin{aligned}
\text{Now,}\quad z_j^* - c_j &= \left\langle c_B^T, B^{*-1}a_j\right\rangle - c_j \\
&= \left\langle c_B^T, (EB^{-1})a_j\right\rangle - c_j \quad \left[\because B^{*-1} = EB^{-1}\right] \\
&= \left\langle c_B^T, (EB^{-1}a_j)\right\rangle - c_j \\
&= \left\langle c_B^T, (Ey_j)\right\rangle - c_j \quad \left[\because y_j = B^{-1}a_j\right] \\
&= \sum_{i=1}^{m} c_{Bi}y_{ij}^* - c_j,
\end{aligned}$$

$$\text{where } y_{ij}^* = \begin{cases} y_{ij} - \frac{\Delta a_{rk}\beta_{ir}}{1+\beta_{pr}\Delta a_{rk}}y_{pj}, & i \ne p \\ \frac{y_{pj}}{1+\beta_{pr}\Delta a_{rk}}, & i = p \end{cases}$$

$$\begin{aligned}
\therefore z_j^* - c_j &= \sum_{i=0}^{m} c_{Bi}\left(y_{ij} - \frac{\Delta a_{rk}\beta_{ir}y_{pj}}{1+\beta_{pr}\Delta a_{rk}}\right) - c_j \\
&= \sum_{i=1}^{m} c_{Bi}y_{ij} - c_j - \sum_{i=1}^{m} c_{Bi}\frac{\Delta a_{rk}\beta_{ir}y_{pj}}{1+\beta_{pr}\Delta a_{rk}} \\
&= z_j - c_j - \frac{y_{pj}\Delta a_{rk}\left\langle c_B^T, \beta_r\right\rangle}{1+\beta_{pr}\Delta a_{rk}}.
\end{aligned}$$

Thus, the condition of optimality requires $z_j^* - c_j \ge 0$; i.e. $z_j - c_j - \frac{y_{pj}\Delta a_{rk}\left\langle c_B^T, \beta_r\right\rangle}{1+\beta_{pr}\Delta a_{rk}} \ge 0$

or $z_j - c_j + \Delta a_{rk}\left\{(z_j - c_j)\beta_{pr} - y_{pj}\left\langle c_B^T, \beta_r\right\rangle\right\} \ge 0.$

Hence, $\Delta a_{rk} \ge -\frac{(z_j-c_j)}{(z_j-c_j)\beta_{pr}-y_{pj}\left\langle c_B^T,\beta_r\right\rangle}$, if $(z_j - c_j)\beta_{pr} - y_{pj}\left\langle c_B^T, \beta_r\right\rangle > 0$

and $\Delta a_{rk} \le -\frac{(z_j-c_j)}{(z_j-c_j)\beta_{pr}-y_{pj}\left\langle c_B^T,\beta_r\right\rangle}$, if $(z_j - c_j)\beta_{pr} - y_{pj}\left\langle c_B^T, \beta_r\right\rangle < 0.$

Hence, in order to maintain the optimality of the new solution, $\Delta a_{rk}$ must satisfy the following inequalities (which are the combination of the preceding two inequalities):

$$\text{Max}\left\{-\frac{(z_j - c_j)}{P_j}, P_j > 0\right\} \le \Delta a_{rk} \le \text{Min}\left\{\frac{-(z_j - c_j)}{P_j}, P_j < 0\right\},$$

where $P_j = \left(z_j - c_j\right)\beta_{pr} - y_{pj}\left\langle c_B^T, \beta_r\right\rangle$.

## 7.6 Addition of a Variable

Addition of a new variable in an LPP plays an important role in post-optimality analysis. Due to this change, structural variation in the simplex table takes place. Let a new variable $x_{n+1}$ be introduced in the following LPP whose optimal solution is known:

$$\begin{aligned}&\text{Maximize}\, z = \langle c^T, x\rangle\\&\text{subject to}\, Ax = b \;\text{ and }\; x \ge 0.\end{aligned}$$

Also, let $a_{n+1}$ be the coefficient vector associated with $x_{n+1}$ and let the profit coefficient for $x_{n+1}$ be $c_{n+1}$.

Since the requirement vector $b$ is not changed, the old optimal solution will be a feasible solution of the new LPP, but it may not be optimal.

Now, we have to compute

$$y_{n+1} = B^{-1}a_{n+1} \text{ and } z_{n+1} - c_{n+1} = \left\langle c_B^T, y_{n+1}\right\rangle - c_{n+1}.$$

If $z_{n+1} - c_{n+1} \ge 0$, then $x_{n+1} = 0$, and the current optimal solution remains optimal for the new problem. If $z_{n+1} - c_{n+1} < 0, y_{n+1}$ enters the basis, and the simplex method is applied to obtain the optimal solution of the resulting new problem.

*Example 7* Consider the LPP

$$\begin{aligned}\text{Maximize}\; z &= x_1 + 2x_2 + x_3\\\text{subject to}\; 2x_1 + x_2 - x_3 &\le 2\\2x_1 - x_2 + 5x_3 &\le 6\\4x_1 + x_2 + x_3 &\le 6\\\text{and}\, x_1,\, x_2,\, x_3 &\ge 0.\end{aligned}$$

The optimal simplex table is as follows:

| | | $c_j$ | 1 | 2 | 1 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 2 | $y_2$ | $x_2 = 4$ | 3 | 1 | 0 | 5/4 | 1/4 | 0 | |
| 1 | $y_3$ | $x_3 = 2$ | 1 | 0 | 1 | 1/4 | 1/4 | 0 | |
| 2 | $y_6$ | $x_6 = 0$ | 0 | 0 | 0 | −3/2 | −1/2 | 1 | |
| $z_B = 10$ | | | 6 | 0 | 0 | 1/4 | 3/4 | 0 | $\leftarrow \Delta_j$ |

If a new variable $x_7(\geq 0)$ is introduced with profit (i) 4, (ii) 6 and $a_7 = [2 \quad -1 \quad 4]^{\mathrm{T}}$, discuss the effect of the introduction of the new variable $x_7$.

**Solution** From the given optimal simplex table, we get the inverse of the basis matrix as

$$B^{-1} = [y_4 \quad y_5 \quad y_6] = \begin{bmatrix} 5/4 & 1/4 & 0 \\ 1/4 & 1/4 & 0 \\ -3/2 & -1/2 & 1 \end{bmatrix}.$$

Since $a_7 = [2 \quad -1 \quad 4]^{\mathrm{T}}$, the corresponding column vector to be introduced in the optimal simplex table is given by

$$y_7 = B^{-1}a_7 = \begin{bmatrix} 5/4 & 1/4 & 0 \\ 1/4 & 1/4 & 0 \\ -3/2 & -1/2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 9/4 \\ 1/4 \\ 3/2 \end{bmatrix}.$$

(i) When $c_7 = 4$,

$$z_7 - c_7 = 2 \times \frac{9}{4} + 1 \times \frac{1}{4} + 0 \times \frac{3}{2} - 4 = \frac{3}{4} > 0.$$

Hence, the optimality condition is satisfied for the modified problem also. Therefore, the optimal solution for the modified problem is given by

$$x_1 = 0, x_2 = 4, x_3 = 2, x_7 = 0 \,\text{with}\, z_{\text{Max}} = 10.$$

(ii) When $c_7 = 6$,

$$z_7 - c_7 = 2 \times \frac{9}{4} + 1 \times \frac{1}{4} + 0 \times \frac{3}{2} - 6 = -\frac{5}{4} < 0.$$

Hence, the optimality condition is not satisfied. We have to modify the optimal simplex table of the old problem with the added column vector $y_7 = [9/4 \quad 1/4 \quad 3/2]^{\mathrm{T}}$, $z_7 - c_7 = -\frac{5}{4}$ and $c_7 = 6$.

Now, to get the optimal solution, we shall apply the simplex method. The consecutive tables obtained are as follows:

| $c_j$ | | | 1 | 2 | 1 | 0 | 0 | 0 | 6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | Min ratio |
| 2 | $y_2$ | $x_2 = 4$ | 3 | 1 | 0 | 5/4 | 1/4 | 0 | 9/4 | 16/9 |
| 1 | $y_3$ | $x_3 = 2$ | 1 | 0 | 1 | 1/4 | 1/4 | 0 | 1/4 | 8 |
| 0 | $y_6$ | $x_6 = 0$ | 0 | 0 | 0 | −3/2 | −1/2 | 1 | 3/2 | 0 |
| $z_B = 10$ | | | 6 | 0 | 0 | 11/4 | 3/4 | 0 | −5/4 | $\leftarrow \Delta_j$ |

| $c_j$ | | | 1 | 2 | 1 | 0 | 0 | 0 | 6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | |
| 2 | $y_2$ | $x_2 = 4$ | 3 | 1 | 0 | 7/2 | 1 | −3/2 | 0 | |
| 1 | $y_3$ | $x_3 = 2$ | 1 | 0 | 1 | 1/2 | 1/3 | −1/6 | 0 | |
| 6 | $y_7$ | $x_7 = 0$ | 0 | 0 | 0 | −1 | −1/3 | 2/3 | 1 | |
| $z_B = 10$ | | | 6 | 0 | 0 | 3/2 | 1/3 | 5/6 | 0 | $\leftarrow \Delta_j$ |

From this table, it is seen that all $x_{B_i} \geq 0$ and all $\Delta_j \geq 0$, and so the optimality conditions are satisfied. Hence, the optimal solution is given by $x_1 = 0,\ x_2 = 4,\ x_3 = 2,\ x_7 = 0$ and $z_{\text{Max}} = 10$.

## 7.7 Deletion of a Variable

Like addition of a new variable, deletion of a variable (basic or non-basic) also changes the structure of the problem. According to the type of deleted variables, two cases may arise.

**Case 1**: If the variable to be deleted is non-basic, then the feasibility and optimality conditions are not affected. This means that the current optimal solution is the optimal solution of the new problem.

**Case 2**: If the deleted variable is basic, then the conditions of optimality may be affected, and a new solution must be obtained. To find this new optimal solution, we have to assign a profit $-M$ corresponding to the basic variable to be deleted and apply the simplex method, taking the modified optimal simplex table as the starting simplex table.

## 7.8 Addition of a New Constraint

Let us suppose that a new constraint, i.e. the $(m + 1)$th constraint

$$a_{m+1,1}x_1 + a_{m+1,2}x_2 + \cdots + a_{m+1,n}\, x_n \le b_{m+1}, \quad \text{i.e. } ux \le b_{m+1},$$

is added to the system of original constraints $Ax = b,\ x \ge 0$, where $u = (a_{m+1,1}, a_{m+1,2}, \ldots, a_{m+1,n})$ and $b_{m+1}$ is positive or 0 or negative.

Then two cases may arise.

(i) The current optimal solution $x_B$ of the original problem satisfies the new constraint. If so, the solution remains feasible as well as optimal.

(ii) If the current optimal solution $x_B$ does not satisfy the new constraint, then the new basic feasible solution will be $x_B^* = \begin{pmatrix} x_B \\ x_S \end{pmatrix}$, where $x_s$ is the slack variable introduced in the additional constraint. In that case the new basis matrix $B^*$ of order $m + 1$ is given by $B^* = \begin{pmatrix} B & O \\ \alpha & 1 \end{pmatrix}$, where $\alpha$ is a submatrix of $u$ [the components are the coefficients of basic variables].

$$\therefore B^{*-1} = \begin{pmatrix} B^{-1} & O \\ -\alpha B^{-1} & 1 \end{pmatrix}.$$

Now, $x_B^* = B^{*-1}b^* = \begin{pmatrix} B^{-1} & O \\ -\alpha B^{-1} & 1 \end{pmatrix}\begin{pmatrix} b \\ b_{m+1} \end{pmatrix} = \begin{pmatrix} x_B \\ b_{m+1} - \alpha x_B \end{pmatrix}.$

If $x_B^* \ge 0$, the current solution remains basic feasible for the new problem also; otherwise, we apply the dual simplex method to obtain the basic feasible solution.

In this case, note that the following three situations may arise depending on the nature of the solution to the original LPP.

(i) If the original LPP has an optimal solution, then the modified LPP may have an optimal solution or it will have no feasible solution.

(ii) If the original LPP has an unbounded solution, then the modified LPP may have an optimal solution or it will have no feasible solution or it will have an unbounded solution.

(iii) If the original LPP has no feasible solution, then the modified LPP will also have no feasible solution.

*Example 8* We are given the following LPP:

$$\text{Maximize}\, z = 3x_1 + 4x_2 + x_3 + 7x_4$$

subject to the constraints

$$\begin{aligned} 8x_1 + 3x_2 + 4x_3 + x_4 &\leq 7 \\ 2x_1 + 6x_2 + x_3 + 5x_4 &\leq 3 \\ x_1 + 4x_2 + 5x_3 + 2x_4 &\leq 8 \\ \text{and}\, x_1, x_2, x_3, x_4 &\geq 0. \end{aligned}$$

(i) Discuss the effect of discrete changes in $b_1$.
(ii) Let the variable $x_4$ be deleted from the LPP part of the problem. Obtain an optimum solution to the resulting LPP.
(iii) Let a linear constraint $2x_1 + 3x_2 + x_3 + 5x_4 \leq 4$ be the addition to the constraints of the problem. Check whether there is any change in the optimum solution of the original problem.
(iv) Also discuss the case where the upper limit of the preceding constraint is reduced to 2.

**Solution** For the given LPP, the optimum table is the following:

| $c_j$ | | | 3 | 4 | 1 | 7 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | |
| 3 | $y_1$ | $x_1 = 16/19$ | 1 | 9/38 | 1/2 | 0 | 5/38 | −1/38 | 0 | |
| 7 | $y_4$ | $x_4 = 5/19$ | 0 | 21/19 | 0 | 1 | −1/19 | 8/38 | 0 | |
| 0 | $y_7$ | $x_7 = 126/19$ | 0 | 59/38 | 9/2 | 0 | −1/38 | −15/38 | 1 | |
| $z_B = 83/19$ | | | 0 | 169/38 | 1/2 | 0 | 1/38 | 53/38 | 0 | $\leftarrow \Delta_j$ |

(i) From this optimum table, we have $x_B = [16/19,\ 5/19,\ 126/19]^T$, $b = [7,\ 3,\ 8]^T$

$$\text{and}\, B^{-1} = (y_5\ y_6\ y_7) = \begin{pmatrix} \frac{5}{38} & \frac{-1}{38} & 0 \\ \frac{-1}{19} & \frac{8}{38} & 0 \\ \frac{-1}{38} & \frac{-15}{38} & 1 \end{pmatrix}.$$

Let the requirement vector component $b_1$ be changed to $b_1 + \Delta b_1$. Then the range of $\Delta b_1$ for such a change, maintaining the optimality to the basic feasible solution, is given by

$$\text{Max}\left\{\frac{-x_{Bi}}{b_{i1}}, b_{i1} > 0\right\} \leq \Delta b_1 \leq \text{Min}\left\{\frac{-x_{Bi}}{b_{i1}}, b_{i1} < 0\right\}$$

or $\text{Max}\{(-16/19)/(5/38)\} \leq \Delta b_1 \leq \text{Min}\{(-5/19)/(-1/19), (-126/19)/(-1/38)\}$

or $-32/5 \leq \Delta b_1 \leq \text{Min}\{5, 252\}$ or, $-32/5 \leq \Delta b_1 \leq 5.$

(ii) Since the variable $x_4$ to be deleted is a basic variable, we assign a cost—$M$ to $x_4$ and take the current optimum simplex table as the initial simplex table for the new LPP.

| $c_j$ | | | 3 | 4 | 1 | $-M$ | 0 | 0 | 0 | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | |
| 3 | $y_1$ | 16/19 | 1 | 9/38 | 1/2 | 0 | 5/38 | −1/38 | 0 | 32/9 |
| $-M$ | $y_4$ | 5/19 | 0 | 21/19 | 0 | 1 | −1/19 | 4/19 | 0 | 5/21 |
| 0 | $y_7$ | 126/19 | 0 | 59/38 | 9/2 | 0 | −1/38 | −15/38 | 1 | 252/59 |
| $z_B = -5M/19 + 48/19$ | | | 0 | $-\frac{21M}{19} - \frac{125}{38}$ | $\frac{1}{2}$ | 0 | $\frac{M}{19} + \frac{15}{38}$ | $-\frac{4M}{19} - \frac{3}{38}$ | 0 | $\leftarrow \Delta_j$ |

Here $y_2$ is the incoming vector, $y_4$ is the outgoing vector and $y_{22}$ = 21/19 is the key element.

Now introducing $y_2$ and dropping $y_4$, we develop the next table.

| $c_j$ | | | 3 | 4 | 1 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_5$ | $y_6$ | $y_7$ | |
| 3 | $y_1$ | $x_1 = 209/266$ | 1 | 0 | 1/2 | 1/7 | −1/14 | 0 | |
| 4 | $y_2$ | $x_2 = 5/21$ | 0 | 1 | 0 | −1/21 | 4/21 | 0 | |
| 0 | $y_7$ | $x_7 = 13997/798$ | 0 | 0 | 9/2 | 1/21 | −187/266 | 1 | |
| $z_B = 139/42$ | | | 0 | 0 | 1/2 | 5/21 | 23/42 | 0 | $\leftarrow \Delta_j$ |

This shows that the new optimum solution to the LPP when the variable $x_4$ has been deleted is given by $x_1$ = 209/266, $x_2$ = 9/2, $x_3$ = 0 and $x_4$ = 0, Max $z$ = 139/ 42.

(iii) From the optimum simplex table of the given LPP, we have $x_1$ = 16/19, $x2$ = 0, $x_3$ = 0, $x_4$ = 5/19.

For this solution, $2x_1 + 3x_2 + x_3 + 5x_4 = 2.16/19 + 5.5/19 = (32 + 25)/19 = 3 < 4.$

Hence, the current optimal solution also satisfies the additional constraint. Thus, the additional constraint is redundant, and the optimum solution of the given LPP is also the optimum solution of the modified LPP (the original LPP including the additional constraint).

(iv) When the upper limit of the additional constraint is reduced to 2, the additional constraint will be $2x_1 + 3x_2 + x_3 + 5x_4 \leq 2$. The current optimum solution does not satisfy the additional constraint, so by introducing a slack variable $x_8(\geq 0)$ in the additional constraint, we have

$$2x_1 + 3x_2 + x_3 + 5x_4 + x_8 = 2.$$

Now appending this constraint in the optimum table, we have the following:

| $c_j$ | | | 3 | 4 | 1 | 7 | 0 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | $y_8$ |
| 3 | $y_1$ | 16/19 | 1 | 9/38 | 1/2 | 0 | 5/38 | −1/38 | 0 | 0 |
| 7 | $y_4$ | 5/19 | 0 | 21/19 | 0 | 1 | −1/19 | 4/19 | 0 | 0 |
| 0 | $y_7$ | 126/19 | 0 | 59/38 | 9/2 | 0 | −1/38 | −15/38 | 1 | 0 |
| 0 | $y_8$ | 2 | 2 | 3 | 1 | 5 | 0 | 0 | 0 | 1 |

The basis $B$ = ($y_1$, $y_2$, $y_7$, $y_8$) is obtained by performing the operation $R_4' = R_4 - 2R_1$ and then $R_4' = R_4 - 5R_2$, to obtain the following simplex table:

| $c_j$ | | | 3 | 4 | 1 | 7 | 0 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | $y_8$ | |
| 3 | $y_1$ | 16/19 | 1 | 9/38 | 1/2 | 0 | 5/38 | −1/38 | 0 | 0 | |
| 7 | $y_4$ | 5/19 | 0 | 21/19 | 0 | 1 | −1/19 | 4/19 | 0 | 0 | |
| 0 | $y_7$ | 126/19 | 0 | 59/38 | 9/2 | 0 | −1/38 | −15/38 | 1 | 0 | |
| 0 | $y_8$ | −1 | 0 | −3 | 0 | 0 | 0 | −1 | 0 | 1 | |
| $z_B = 83/19$ | | | 0 | 169/38 | 1/2 | 0 | 1/38 | 53/38 | 0 | 0 | $\leftarrow \Delta_j$ |

Since all $z_j - c_j \geq 0$ and $x_{B4} = -1$, an optimum or more than optimum but infeasible solution has been obtained.

Hence, we have to apply the dual simplex method to solve the problem. As $x_{B4} = -1$, which is negative, the corresponding basis vector $y_8$ leaves the basis.

$$\begin{aligned}
\text{Now} \quad & \text{Max}\left\{\frac{z_j - c_j}{y_{4j}}, y_{4j} < 0\right\}, \quad j = 1, 2, \ldots, 8 \\
&= \text{Max}\left\{\frac{169/38}{-3}, \frac{53/38}{-1}\right\} = \text{Max}\left\{-\frac{169}{114}, -\frac{53}{38}\right\} \\
&= -\frac{53}{38}, \text{which occurs for } j = 6.
\end{aligned}$$

Therefore, $y_6$ enters the basis, and $y_{46} = -1$ is the key element.
Now we construct the next simplex table.

| $c_j$ | | | 3 | 4 | 1 | 7 | 0 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | $y_8$ | |
| 3 | $y_1$ | 33/38 | 1 | 6/19 | 1/2 | 0 | 5/38 | 0 | 0 | −1/38 | |
| 7 | $y_4$ | 1/19 | 0 | 9/19 | 0 | 1 | −1/19 | 0 | 0 | 4/19 | |
| 0 | $y_7$ | 267/38 | 0 | 7/19 | 9/2 | 0 | −1/38 | 0 | 1 | −15/38 | |
| 0 | $y_6$ | 1 | 0 | 3 | 0 | 0 | 0 | 1 | 0 | −1 | |
| $z_B = 113/38$ | | | 0 | 5/19 | 1/2 | 0 | 1/38 | 0 | 0 | 53/38 | $\leftarrow \Delta_j$ |

Since all $z_j - c_j \geq 0$, an optimum solution has been attained.

From the preceding simplex table, we see that the additional constraint has decreased the optimum value of the objective function from 83/19 to 113/38.

*Example 9* We are given the following LPP:

$$\text{Maximize}\, Z = 3x_1 + 4x_2 + x_3 + 7x_4$$

subject to the constraints

$$\begin{aligned} 8x_1 + 3x_2 + 4x_3 + x_4 &\leq 7 \\ 2x_1 + 6x_2 + x_3 + 5x_4 &\leq 3 \\ x_1 + 4x_2 + 5x_3 + 2x_4 &\leq 8 \\ \text{and} \quad x_1, x_2, x_3, x_4 &\geq 0. \end{aligned}$$

The following table gives the optimum solution:

| $c_j$ | | | 3 | 4 | 1 | 7 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | |
| 3 | $y_1$ | 16/19 | 1 | 9/38 | 1/2 | 0 | 5/38 | −1/38 | 0 | |
| 7 | $y_4$ | 5/19 | 0 | 21/19 | 0 | 1 | −1/19 | 8/38 | 0 | |
| 0 | $y_7$ | 126/19 | 0 | 59/38 | 9/2 | 0 | −1/38 | −15/38 | 1 | |
| $z_B = 83/19$ | | | 0 | 169/38 | 1/2 | 0 | 1/38 | 53/38 | 0 | $\leftarrow \Delta_j$ |

Discuss the effect of discrete changes in the activity coefficients $a_{12}$ and $a_{24}$ of the coefficient matrix $A$ on the current optimum basic feasible solution to the given LPP.

**Solution** From the given optimum simplex table, we have

$$B^{-1} = \begin{array}{c} \begin{array}{ccc} \beta_1 & \beta_2 & \beta_3 \end{array} \\ \begin{pmatrix} \frac{5}{38} & \frac{-1}{38} & 0 \\ \frac{-1}{19} & \frac{8}{38} & 0 \\ \frac{-1}{38} & \frac{-15}{38} & 1 \end{pmatrix} \end{array},$$

where $B$ is the basis for the given problem.

Let the element $a_{12}$ of $A$ be changed to $a_{12}^* = a_{12} + \Delta a_{12}$.

In the optimum simplex table, the element $a_{12}$, i.e. the corresponding column vector $y_2$, is not in the basis.

We know that the range for discrete change $\Delta a_{rk}$ in the coefficient matrix $A$ for maintaining the optimality of the solution is given by

$$\Delta a_{rk} \geq \frac{-(z_k - c_k)}{\left\langle c_B^T, \beta_r \right\rangle} \quad \text{for} \quad \left\langle c_B^T, \beta_r \right\rangle > 0$$
$$\Delta a_{rk} \leq \frac{-(z_k - c_k)}{\left\langle c_B^T, \beta_r \right\rangle} \quad \text{for} \quad \left\langle c_B^T, \beta_r \right\rangle < 0$$
$$\Delta a_{pk} \text{ is unrestricted for } \left\langle c_s^T, \beta_r \right\rangle = 0.$$

Here $r = 1$, $k = 2$.

Now $\left\langle c_B^T, \beta_1 \right\rangle = 3(5/38) + 7(-1/19) + 0 = 1/38 > 0$.

Therefore, the range for discrete change $\Delta a_{12}$ in the coefficient matrix $A$ for maintaining the optimality of the solution is given by

$$\Delta a_{12} \geq \frac{-(z_2 - c_2)}{\left\langle c_B^T, \beta_1 \right\rangle} \quad [\because c_B \beta_1 > 0]$$
$$\text{or} \quad \Delta a_{12} \geq -(169/38)/(1/38)$$
$$\text{or} \quad \Delta a_{12} \geq -169.$$

Let the element $a_{24}$ of $A$ be changed to $a_{24}^* = a_{24} + \Delta a_{24}$.

In the optimum simplex table, the element $a_{24}$ corresponding to the column vector $y_4$ is in the basis.

Hence, the discrete change in $a_{24}$ belonging to vector $y_4$ may affect the feasibility as well as the optimality of the original optimum basic feasible solution.

We know that the range for discrete change in the element $a_{24}$ of the coefficient matrix in order to maintain the optimality of the new solution is given by

$$\text{Max}\left\{-\frac{(z_j - c_j)}{P_j}, P_j > 0\right\} \leq \Delta a_{rk} \leq \text{Min}\left\{-\frac{(z_j - c_j)}{P_j}, P_j < 0\right\},$$
$$\text{where} \quad P_j = (z_j - c_j)\beta_{pr} - y_{pj}\left\langle c_B^T, \beta_r\right\rangle.$$

Here $r = 2$, $k = 4$ and $p = 2$.

Now, $\left\langle c_B^T, \beta_2\right\rangle = 3(-1/38) + 7(8/38) + 0(-15/38) = 53/38.$

$$\text{Hence, } P_2 = (z_2 - c_2)\beta_{22} - y_{22}\left\langle c_B^T, \beta_2\right\rangle = (169/38)(8/38) - (21/19)(53/38) = -23/38$$
$$P_3 = (z_3 - c_3)b_{22} - y_{23}\left\langle c_B^T, \beta_2\right\rangle = (1/2)(8/38) - 0(53/38) = 2/19$$
$$P_5 = (z_5 - c_5)b_{22} - y_{25}\left\langle c_B^T, \beta_2\right\rangle = (1/38)(8/38) - (-1/19)(53/38) = 3/38$$
$$P_6 = (z_6 - c_6)b_{22} - y_{26}\left\langle c_B^T, \beta_2\right\rangle = (53/8)(8/38) - (8/38)(53/38) = 0.$$

Therefore, the range for discrete change in the element $a_{24}$ of the coefficient matrix in order to maintain the optimality of the new solution is given by

$$\text{or} \quad \text{Max}\left\{-\tfrac{(z_3 - c_3)}{P_3}, -\tfrac{(z_5 - c_5)}{P_5}\right\} \leq \Delta a_{24} \leq \text{Min}\left\{-\tfrac{(z_2 - c_2)}{P_2}\right\}$$
$$\text{or} \quad -1/3 \leq a_{24} \leq 169/23.$$

Again, we know that the range for discrete change in the element $a_{24}$ in order to maintain the optimality of the new solution is given by

$$\text{Max}\left\{-\frac{x_{B_i}}{Q_i}, Q_i > 0\right\} \leq \Delta a_{rk} \leq \text{Min}\left\{-\frac{x_{B_i}}{Q_i}, Q_i < 0\right\},$$

where $Q_i = \beta_{pr}x_{B_i} - \beta_{ir}x_{B_p}$.

Here $r = 2$, $k = 4$, $p = 2$.

Now $Q_1 = \beta_{22}x_{B_1} - \beta_{12}x_{B_2} = \frac{8}{38} \times \frac{16}{19} - \left(-\frac{1}{38}\right) \times \frac{5}{19} = \frac{7}{38}$

$$Q_3 = \beta_{22}x_{B_3} - \beta_{32}x_{B_2} = \frac{8}{38} \times \frac{126}{19} - \left(-\frac{15}{38}\right) \times \frac{5}{19} = \frac{3}{2}.$$

Therefore, the range for the discrete change in the element $a_{24}$ in order to maintain the optimality of the new solution is given by $\text{Max}\left\{-\frac{x_{B_1}}{Q_1}, -\frac{x_{B_3}}{Q_3}\right\} \leq \Delta a_{24} < \infty$

or $\text{Max}\left\{-\frac{16/19}{7/38}, -\frac{126/19}{3/2}\right\} \leq \Delta a_{24}$ or $-\frac{84}{19} \leq \Delta a_{24} < \infty.$

## 7.9 Exercises

1. Solve the following LPP:

$$\begin{aligned}
&\text{Maximize}\, z = x_1 + 3x_2 - 2x_3\\
&\text{subject to the constraints}\\
&3x_1 - x_2 + 2x_3 \le 7\\
&\quad - 2x_1 + 4x_2 \le 12\\
&\quad - 4x_1 + 3x_2 + 8x_3 \le 10.\\
&\quad\ \text{and}\, x_1, x_2, x_3 \ge 0.
\end{aligned}$$

Determine the separate range for discrete changes in $a_{13}$, $a_{23}$, $a_{21}$ consistent with the optimal solution of the given LPP.

2. The optimal solution to the problem

$$\begin{aligned}
&\text{Maximize}\, Z = x_1 + 2x_2 - x_3\\
&\text{subject to the constraints}\\
&3x_1 + x_2 - x_3 \le 10, -x_1 + 4x_2 + x_3 \ge 6, \quad x_2 + x_3 \le 4\\
&\quad \text{and}\, x_1, x_2, x_3 \ge 0
\end{aligned}$$

is given in the following table:

| $c_j \rightarrow$ | | | $-1$ | $2$ | $-1$ | $0$ | $0$ | $0$ | $-M$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $y_B$ | $c_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | |
| $y_4$ | 0 | 6 | 3 | 0 | $-2$ | 1 | $-1$ | 0 | 0 | |
| $y_2$ | 2 | 4 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | |
| $y_6$ | 0 | 10 | 1 | 0 | 3 | 0 | 4 | 1 | $-1$ | |
| $z_B = 8$ | | | 1 | 0 | 3 | 0 | 2 | 0 | $M$ | $\leftarrow \Delta_j$ |

Determine the ranges for discrete changes in the components $b2$ and $b3$ of the required vector so as to maintain the optimality of the current optimal solution.

3. Will the optimal solution to the following LPP remain optimal if the requirement vector $(3{,}5{,}6)^{\mathrm{T}}$ changes to $(3,\ 10,\ 6)^{\mathrm{T}}$?

$$\begin{aligned}\text{Minimize } z &= 8x_1 + 3x_2 + 6x_3 + 3x_4\\ \text{subject to } x_2 + 4x_3 + 5x_4 &\le 3\\ 3x_1 + 2x_3 &\le 5\\ x_1 + 2x_2 + x_4 &\le 6\\ x_1, x_2, x_3, x_4 &\ge 0\end{aligned}$$

4. Find the optimal solution of the LPP

$$\begin{aligned}\text{Maximize } z &= 4x_1 + 3x_2\\ \text{subject to } x_1 + x_2 &\le 5\\ 3x_1 + x_2 &\le 7\\ x_1 + 2x_2 &\le 10,\\ x_1, x_2 &\ge 0.\end{aligned}$$

Show how to find the optimal solution of the problem if:

(i) The first component of the original requirement vector is increased by one unit and the third component is decreased by one unit.
(ii) The second component of the original requirement vector is decreased by two units.

5. For the LPP

$$\begin{aligned}\text{Maximize } z &= x_1 + 2x_2 + x_3\\ \text{subject to } 2x_1 + x_2 - x_3 &\le 2\\ 2x_1 - x_2 + 5x_3 &\le 6\\ 4x_1 + x_2 + x_3 &\le 6\\ x_1, x_2, x_3 &\ge 0,\end{aligned}$$

the optimal table is as follows:

| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
|---|---|---|---|---|---|---|---|---|---|
| 2 | $y_2$ | 4 | 3 | 1 | 0 | 5/4 | 1/4 | 0 | |
| 1 | $y_3$ | 2 | 1 | 0 | 1 | 1/4 | 1/4 | 0 | |
| 0 | $y_6$ | 0 | 0 | 0 | 0 | −3/2 | −1/2 | 1 | |
| $z_B = 10$ | | | 6 | 0 | 0 | 1/4 | 3/4 | 0 | $\leftarrow \Delta_j$ |

Discuss the effect of deletion of the variable (i) $x_1$, (ii) $x_2$, (iii) $x_3$.

6. Let us consider the final table of an LPP.

| | | $c_j$ | 2 | 4 | 1 | 8 | 2 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | $y_8$ | |
| 2 | $y_1$ | 3 | 1 | 0 | 0 | −1 | 0 | 1/2 | 1/5 | −1 | |
| 4 | $y_2$ | 1 | 0 | 1 | 0 | 2 | 1 | −1 | 0 | ½ | |
| 1 | $y_3$ | 7 | 0 | 0 | 1 | 1 | −2 | 5 | −3/10 | 2 | |
| | | | 0 | 0 | 0 | −1 | 0 | 2 | 1/10 | 2 | $\leftarrow \Delta_j$ |

Here $y_6$, $y_7$ and $y_8$ are slack variables.
If the constraints

(i) $2x_1 + 3x_2 - x_3 + 2x_4 + 4x_5 \le 5$

(ii) $2x_1 + 3x_2 - x_3 + 2x_4 + 4x_5 \le 1$

are added, then find the solution of the changed LPP.

7. For the LPP

$$\begin{aligned} &\text{Maximize}\, z = 15x_1 + 45x_2 \\ &\text{subject to}\, 5x_1 + 2x_2 \le 162 \\ &\qquad\qquad x_1 + 16x_2 \le 240 \\ &\qquad\qquad x_2 \le 50 \\ &\qquad\qquad \text{and}\, x_1,\, x_2 \ge 0, \end{aligned}$$

find the optimal solution. Find the range of each cost coefficient (changed one at a time) for which the current solution will remain optimal.

8. Find the optimal solution of the LPP and the separate ranges of variations of $b_2$ and $b_3$ consistent with the optimality of the solution

$$\begin{aligned} &\text{Minimize}\, z = -x_1 + 2x_2 - x_3 \\ &\text{subject to}\, 3x_1 + x_2 - x_3 \le 10 \\ &\qquad\qquad -x_1 + 4x_2 + x_3 \ge 6 \\ &\qquad\qquad x_2 + x_3 \le 4 \\ &\qquad\qquad \text{and}\, x_1,\, x_2,\, x_3 \ge 0. \end{aligned}$$

Determine also the efficient discrete changes in the components of the cost vector which correspond to the basic variables.

9. Following is the optimal table for an LPP:

| | | $c_j$ | 2 | 1 | 1 | 2 | 0 |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ |
| 2 | $y_1$ | 3 | 1 | 0 | −1 | 3 | 2 |
| 1 | $y_2$ | 4 | 0 | 1 | 4 | −1 | −2 |
| | | | 0 | 0 | 1 | 3 | 2 |

(i) Find the limitations of the values of $c_3$, $c_4$, $c_5$ (taking one at a time) for which the current solution will remain optimal.
(ii) Find the optimal solution to the problem if $c_3$ is changed to 3.
(iii) Find the limitations of the values of $c_1$ for which the current solution remains optimal.
(iv) Find the optimal solution to this problem if $c_1$ is changed to 5.

10. Find the optimal solution of the LPP

$$\begin{aligned}
&\text{Maximize}\, z = 4x_1 + 3x_2 \\
&\text{subject to}\, x_1 + x_2 \le 5 \\
&\qquad 3x_1 + x_2 \le 7 \\
&\qquad x_1 + 2x_2 \le 10 \\
&\qquad \text{and}\, x_1,\ x_2 \ge 0.
\end{aligned}$$

Then find the optimal solution of the problem, if:

(i) The first component of the original requirement vector is increased by one unit and the third component is decreased by one unit.
(ii) The second component of the original requirement vector is decreased by two units.

# Chapter 8
# Integer Programming

## 8.1 Objectives

After studying this chapter, the readers are able to learn:

- The limitations of the simplex method in deriving integer solutions to LPPs
- Gomory's all-integer and mixed-integer programming techniques
- The branch and bound method to solve integer programming problems.

## 8.2 Introduction

Integer programming problems (IPPs) are a special class of linear programming problem (LPP) where all or some of the decision variables are restricted to the integers. If all the variables are restricted to take integer values, the corresponding IPP is called a pure IPP. On the other hand, if some of the variables are restricted to take integer values, then the problem is called a mixed-integer linear programming problem (mixed ILPP).

In the year 1956, R. E. Gomory first developed a method to solve pure IPPs. Later, he extended the method to solve mixed ILPPs. Another important approach, called the branch and bound method, was developed for solving both all-integer and mixed-integer programming problems.

The integer solution of a given ILPP can be obtained by finding the non-integer optimal solution through the simplex method or the graphical method and then rounding off the fractional values of the variables. But, in some cases, the derivation from the exact optimal integer solution (obtained as a result of rounding) may become large enough to give an infeasible solution. Hence, it is necessary to develop a systematic procedure to determine the optimal integer solution to such problems. The following example will make the concept more understandable.

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_8

Let us consider a simple problem: Maximize $z = 10x_1 + 4x_2$ subject to $3x_1 + 4x_2 \leq 8$, $x_1 \geq 0$, $x_2 \geq 0$, where $x_1$, $x_2$ are integers. First ignoring the integer-valued restrictions, we get the optimal solution $x_1 = 2\frac{2}{3}$, $x_2 = 0$ and Maximize $z = 26\frac{2}{3}$ using the graphical method. Then by rounding off the fractional value of $x_1 = 2\frac{2}{3}$, the optimum solution becomes $x_1 = 3$, $x_2 = 0$ with Maximize $z = 30$. But this solution does not satisfy the constraint $3x_1 + 4x_2 \leq 8$. Therefore, this solution is not feasible.

Now, let us take $x_1 = 2$, $x_3 = 0$, which is feasible and integer-valued. But this solution gives $z$ = 20, which is far away from the value $z = 26\frac{2}{3}$. So this is another disadvantage of getting an integer-valued solution by rounding down the solution. The result may be far away from the optimum solution.

There are two methods used to solve IPPs:

(i) The Gomory method or the Gomory cutting-plane method
(ii) The branch and bound method.

## 8.3 Gomory's All-Integer Cutting-Plane Method

A systematic procedure for solving all IPPs was first developed by R. E. Gomory in 1956, Later, he extended the method to deal with the more complicated case of mixed-integer programming problems.

In this method, first of all, we have to find the optimal solution of the given IPP by the simplex method, ignoring the restriction of integer values of the variables.

(i) If all the variables in the optimal solution obtained have integer values, the current solution will be the required optimal integer solution.
(ii) If not, the considered LPP requires modifications by introducing a secondary constraint (also called the Gomory constraint) that reduces some of the non-integer values of variables to integer ones but does not eliminate any feasible integers.
(iii) Then, an optimal solution to this modified LPP is obtained by using the dual simplex method. If all the variables in this solution are integers, then the optimal integer solution is obtained. Otherwise, another constraint is added to the LPP, and the entire procedure is repeated.

**All Integer Programming Algorithm**

Gomory's fractional cut method may be summarized in the following steps:

Step 1 Express the given LPP in its standard forms by converting the minimization problem to a maximization problem (if necessary) and introducing slack and/or surplus variables with the readjustment of the objective function.

Step 2 Determine the optimal solution by using the simplex method, ignoring the integer value restriction.

Step 3 Test the integrality of the optimum solution

(a) If the optimum solution contains all integer values, an optimal integer basic feasible solution (BFS) has been achieved. Stop the process.
(b) If not, go Step 4.

Step 4 Examine the values of basic variables which give fractional values. Express them as the sum of an integer and a proper fraction. Then, select the largest fractional value of the basic variables, i.e. find $\underset{i}{\text{Max}}\{f_{B_i}\}$. Let it be $f_{B_k}$ for $i = k$.

Step 5 Express each of the negative fractions, if any, in the $k$th row of the optimal simplex table as the sum of a negative integer and a non-negative proper fraction.

Step 6 Construct the Gomorian constraint

$$\sum_{j=1}^{n} f_{kj}x_j \geq f_{B_k}$$

or $-\sum_{j=1}^{n} f_{kj}x_j + G_1 = -f_{B_k}$, where $0 \leq f_{kj} < 1$ and $0 < f_{B_k} < 1$ and $G_1$ is the Gomorian slack variable.

Step 7 Add the Gomorian constraint constructed in Step 6 at the bottom of the optimum simplex table. Use the dual simplex method to find an improved optimal solution. Then go to Step 3.

## 8.4 Construction of Gomory's Constraint

Let the optimal non-integer solution to the maximization LPP be obtained. In our usual notation this solution is shown by the following simplex table:

| $c_B$ | $B$ | $X_B$ | $y_1(\beta_1)$ | $y_2(\beta_2)$ | $\cdots$ | $y_i(\beta_i)$ | | $y_m(\beta_m)$ | $y_{m+1}$ | $\cdots$ | $y_n$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_{B_1}$ | $\beta_1$ | $x_{B_1}$ | 1 | 0 | $\cdots$ | 0 | $\cdots$ | 0 | $y_{1,\,m+1}$ | $\cdots$ | $y_{1,\,n}$ | |
| $c_{B_2}$ | $\beta_2$ | $x_{B_2}$ | 0 | 1 | $\cdots$ | 0 | $\cdots$ | 0 | $y_{2,\,m+1}$ | $\cdots$ | $y_{2,\,n}$ | |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | | $\vdots$ | | $\vdots$ | $\vdots$ | | $\vdots$ | |
| $c_{B_i}$ | $B_i$ | $x_{B_i}$ | 0 | 0 | $\cdots$ | 1 | $\cdots$ | 0 | $y_{i,\,m+1}$ | $\cdots$ | $y_{i,\,n}$ | |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | | $\vdots$ | | $\vdots$ | $\vdots$ | | $\vdots$ | |
| $c_{B_m}$ | $\beta_m$ | $x_{B_m}$ | 0 | 0 | $\cdots$ | 0 | | 1 | $y_{m,\,m+1}$ | $\cdots$ | $y_{m,\,n}$ | |
| $z = <c_B^T,\, x_B>$ | | | 0 | 0 | $\cdots$ | 0 | $\cdots$ | 0 | $\Delta_{m+1}$ | | $\Delta_n$ | $\leftarrow \Delta_j$ |

In this table, the variables $x_{B_i} (i = 1, 2, \ldots, m)$ are taken as basic variables and the remaining $x_{m+1}, x_{m+2}, \ldots, x_n$ as the non-basic variables. (For simplicity, basic variables are taken consecutively.) Let the $i$th basic variable $x_{B_i}$ possess a non-integer value which is given by

$$0.x_1 + 0.x_2 + \cdots + x_i + \cdots + 0.x_m + y_{i,m+1}x_{m+1} + \cdots + y_{i,n}x_n = x_{B_i}$$

$$\text{or} \qquad x_i + \sum_{j=m+1}^{n} y_{ij}x_j = x_{B_i}. \tag{8.1}$$

Now, let us assume that $x_{B_i} = I_{B_i} + f_{B_i}$ and $y_{ij} = I_{ij} + f_{ij}$, where $I_{B_i}$ and $I_{ij}$ are the largest integral parts of $x_{B_i}$ and $y_{ij}$ respectively, such that $I_{B_i} < x_{B_i}$ and $I_{ij} \le y_{ij}$. Hence, $0 < f_{B_i} < 1$ and $0 \le f_{ij} < 1$; i.e. $f_{B_i}$ is a strictly positive proper fraction, while $f_{ij}$ is a non-negative proper fraction or zero. Now substituting $x_{B_i} = I_{B_i} + f_{B_i}$ and $y_{ij} = I_{ij} + f_{ij}$ in (8.1), we have

$$\begin{aligned} & x_i + \sum_{j=m+1}^{n} \left(I_{ij} + f_{ij}\right) x_j = I_{B_i} + f_{B_i} \\ \text{or} \qquad & x_i + \sum_{j=m+1}^{n} I_{ij}x_j + \sum_{j=m+1}^{n} f_{ij}x_j = I_{B_i} + f_{B_i} \end{aligned} \tag{8.2}$$

Now for all the variables $x_i$ (for $i = 1, 2, \ldots, m$) and $x_j$ (for $j = m + 1, m + 2, \ldots, n$) to be integer-valued, the sum of terms with fractional coefficients on the left-hand side must be greater than or equal to the fraction on the right-hand side,

$$\begin{aligned} \text{i.e.} \qquad & \sum_{j=m+1}^{n} f_{ij}x_j \ge f_{B_i} \\ \text{or} \qquad & \sum_{j=m+1}^{n} f_{ij}x_j - f_{B_i} \ge 0 \\ \text{or} \qquad & -\sum_{j=m+1}^{n} f_{ij}x_j + f_{B_i} \le 0 \end{aligned} \tag{8.3}$$

This constraint is known as the Gomorian constraint.

*Example 1* Find the optimal integer solution to the following LPP:

$$\begin{aligned} & \text{Maximize } z = x_1 + x_2 \\ & \text{subject to } 3x_1 + 2x_2 \le 5 \\ & x_2 \le 2,\ x_1,\ x_2 \ge 0 \text{ and are integers.} \end{aligned}$$

**Solution**:
The given problem is a maximization problem.

Now introducing two slack variables $x_3(\geq 0)$ and $x_4(\geq 0)$, the standard form of given LPP becomes

$$\begin{aligned}
&\text{Maximize } z = x_1 + x_2 + 0.x_3 + 0.x_4\\
&\text{subject to} \quad 3x_1 + 2x_2 + x_3 = 5\\
&\qquad\qquad x_2 + x_4 = 2, \quad x_1,\ x_2,\ x_3,\ x_4 \geq 0 \text{ and are integers.}
\end{aligned}$$

Now, ignoring the integer condition, we shall solve the problem by the simplex method.

Here the initial BFS is $x_3$ = 5, $x_4$ = 2 with the initial basis as $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Now we construct the initial simplex table.

| | | $c_j$ | 1 | 1 | 0 | 0 | Minimum ratio |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| 0 | $y_3$ | $x_3$ = 5 | 3 | 2 | 1 | 0 | 5/1 = 5 |
| 0 | $y_4$ | $x_4$ = 2 | 0 | 1 | 0 | 1 | – |
| | | $z_B = 0$ | −1 | −1 | 0 | 0 | $\leftarrow \Delta_j$ |
| | | | ↑ | | ↓ | | |

| | | $c_j$ | 1 | 1 | 0 | 0 | Minimum ratio |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| 1 | $y_1$ | $x_1$ = 5/3 | 1 | 2/3 | 1/3 | 0 | (5/3)/(2/3) = 5/2 |
| 0 | $y_4$ | $x_4$ = 2 | 0 | 1 | 0 | 1 | 2/1 = 2 |
| | | $z_B = \frac{5}{3}$ | 0 | −1/3 | 1/3 | 0 | $\leftarrow \Delta_j$ |
| | | | | ↑ | | ↓ | |

| | | $c_j$ | 1 | 1 | 0 | 0 |
|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ |
| 1 | $y_1$ | $x_1$ = 1/3 | 1 | 0 | 1/3 | −2/3 |
| 1 | $y_2$ | $x_2$ = 2 | 0 | 1 | 0 | 1 |
| | | $z_B = \frac{7}{3}$ | 0 | 0 | 1/3 | 1/3 |

Since all $\Delta_j \geq 0$, an optimal solution is obtained, given by $x_1 = \frac{1}{3}$, $x_2 = 2$. The optimal solution is not an integer solution because $x_1 = \frac{1}{3}$, so we have to add a fractional cut constraint, i.e. a Gomorian constraint, in the optimal simplex table.

Since $x_{B_1} = \frac{1}{3} = 0 + \frac{1}{3}$, we have $f_{B_1} = \frac{1}{3}$. Hence, $\text{Max}(f_{B_1}) = \frac{1}{3} = f_{B_1}$, which corresponds to the first row. So, the first row is the source row.

Now, from the first row of the optimal simplex table, we get

$$x_1 + \frac{1}{3}x_3 - \frac{2}{3}x_4 = \frac{1}{3}$$
$$x_1 + 0.x_3 + \frac{1}{3}x_3 - x_4 + \frac{1}{3}x_4 = \frac{1}{3}.$$

Hence, the Gomorian constraint will be

$$\frac{1}{3}x_3 + \frac{1}{3}x_4 \geq \frac{1}{3}$$
$$\text{or} \quad -\frac{1}{3}x_3 - \frac{1}{3}x_4 \leq -\frac{1}{3}$$
$$\text{or} \quad -\frac{1}{3}x_3 - \frac{1}{3}x_4 + G_1 = -\frac{1}{3},$$

where $G_1$ is the Gomorian slack variable.

Adding this new constraint at the bottom of the preceding optimal simplex table, we have used the dual simplex method to solve the problem.

| | | $c_j$ | 1 | 1 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $g_1$ | |
| 1 | $y_1$ | $x_1 = 1/3$ | 1 | 0 | 1/3 | −2/3 | 0 | |
| 1 | $y_2$ | $x_2 = 2$ | 0 | 1 | 0 | 1 | 0 | |
| 0 | $g_1$ | $G_1 = -1/3$ | 0 | 0 | −1/3 | −1/3 | 1 | |
| | | $z_B = \frac{7}{3}$ | 0 | 0 | 1/3 | 1/3 | 0 | $\leftarrow \Delta_j$ |

We have $x_{B_r} = \text{Min}[x_{B_i},\ x_{B_i} < 0] = \text{Min}\left[-\frac{1}{3}\right] = -\frac{1}{3} = x_{B_3}$. Hence, $r = 3$. So $y_{B_3}$, i.e. $g_1$, leaves the basis. Now,

$$\text{Max}\left\{\frac{\Delta_j}{y_{rj}}, y_{rj} < 0\right\} = \text{Max}\left\{\frac{\frac{1}{3}}{-\frac{1}{3}}, \frac{\frac{1}{3}}{-\frac{1}{3}}\right\}$$
$$= \text{Max}\{-1, -1\} = -1,$$

which corresponds to $y_3$ and $y_4$.

Arbitrarily, we choose $y_3$ as the entering vector, and $y_{33} = -\frac{1}{3}$ is the key element.

Now we construct the next simplex table.

| | | $c_j$ | 1 | 1 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $g_1$ | |
| 1 | $y_1$ | $x_1 = 0$ | 1 | 0 | 0 | −1 | 1 | |
| 1 | $y_2$ | $x_2 = 2$ | 0 | 1 | 0 | 1 | 0 | |
| 0 | $y_3$ | $x_3 = 1$ | 0 | 0 | 1 | 1 | −3 | |
| | | $z_B = 2$ | 0 | 0 | 0 | 0 | 1 | $\leftarrow \Delta_j$ |

Since all $\Delta_j \geq 0$ and all $x_{B_i} \geq 0$, we obtain an optimal feasible integer solution. The optimal integer solution is $x_1 = 0,\ x_2 = 2$ and $\text{Max}\, z = 0 + 2 = 2$.

*Example 2* Find the optimum integer solution to the following LPP:

$$\begin{aligned} &\text{Maximize } z = 2x_1 + 1.7x_2 \\ &\text{subject to} \quad 4x_1 + 3x_2 \leq 7 \\ &\qquad\qquad\quad x_1 + x_2 \leq 4 \\ &\qquad\qquad\quad x_1,\ x_2 \geq 0 \text{ and are integers.} \end{aligned}$$

**Solution**:

The given problem is a maximization problem.

Now introducing two slack variables $x_3(\geq 0)$ and $x_4(\geq 0)$, the standard form of the given LPP becomes

$$\begin{aligned} &\text{Maximize } z = 2x_1 + 1.7x_2 + 0.x_3 + 0.x_4 \\ &\text{subject to} \quad 4x_1 + 3x_2 + x_3 = 7 \\ &\qquad\qquad\quad x_1 + x_2 + x_4 = 4,\ \ x_1,\ x_2,\ x_3,\ x_4 \geq 0 \text{ and are integers.} \end{aligned}$$

Now, ignoring the integer condition, we shall solve the problem by the simplex method.

Here the initial BFS is $x_3 = 7$, $x_4 = 4$ with the initial basis as $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Now we construct the initial simplex table.

| | | $c_j$ | 2 | 1.7 | 0 | 0 | Minimum ratio |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| 0 | $y_3$ | $x_3 = 7$ | 4 | 3 | 1 | 0 | 7/4 = 1.75 |
| 0 | $y_4$ | $x_4 = 4$ | 1 | 1 | 0 | 1 | 4 |
| | | $z_B = 0$ | −2 | −1.7 | 0 | 0 | $\leftarrow \Delta_j$ |
| | | | $\uparrow$ | | $\downarrow$ | | |

| | | $c_j$ | 2 | 1.7 | 0 | 0 | Minimum ratio |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| 2 | $y_1$ | $x_1$ = 7/4 | 1 | 3/4 | 1/4 | 0 | 7/3 |
| 0 | $y_4$ | $x_4$ = 9/4 | 0 | 1/4 | −1/4 | 1 | 9 |
| | | $z_B = \frac{7}{2}$ | 0 | −1/5 | 1/2 | 0 | $\leftarrow \Delta_j$ |
| | | | ↓ | ↑ | | | |

| | | $c_j$ | 2 | 1.7 | 0 | 0 |
|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ |
| 1.7 | $y_2$ | $x_2$ = 7/3 | 4/3 | 1 | 1/3 | 0 |
| 0 | $y_4$ | $x_4$ = 5/3 | −1/3 | 0 | −1/3 | 1 |
| | | $z_B = \frac{11.9}{3}$ | 0.267 | 0 | 1.7/3 | 0 |

Since all $\Delta_j \geq 0$, an optimum solution is obtained, given by $x_2 = \frac{7}{3}$, $x_4 = \frac{5}{3}$. The optimal solution is not an integer solution because $x_2 = \frac{7}{3}$, $x_4 = \frac{5}{3}$, so we have to add a fractional cut constraint or Gomorian constraint in the optimal simplex table.

Since $x_{B_1} = \frac{7}{3} = 2 + \frac{1}{3}$ and $x_{B_2} = \frac{5}{3} = 1 + \frac{2}{3}$, we have $f_{B_1} = \frac{1}{3}$ and $f_{B_2} = \frac{2}{3}$. Hence, $\text{Max}\left(f_{B_1}, f_{B_2}\right) = \frac{2}{3} = f_{B_2}$, which corresponds to the second row. So, the second row is the source row.

Now, from the second row of the optimal simplex table, we get

$$-x_1 + \frac{2}{3}x_1 - x_3 + \frac{2}{3}x_3 + x_4 = 1 + \frac{2}{3}.$$

Hence, the Gomorian constraint will be

$$\begin{aligned} & \frac{2}{3}x_1 + \frac{2}{3}x_3 \geq \frac{2}{3} \\ \text{or} \quad & -\frac{2}{3}x_1 - \frac{2}{3}x_3 \leq -\frac{2}{3} \\ \text{or} \quad & -\frac{2}{3}x_1 - \frac{2}{3}x_3 + G_1 = -\frac{2}{3}, \end{aligned}$$

where $G_1$ is the Gomorian slack variable.

Adding this new constraint at the bottom of the preceding optimal simplex table, we have used the dual simplex method to solve the problem.

| | | $c_j$ | 2 | 1.7 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $g_1$ | |
| 1.7 | $y_2$ | $x_2 = 7/3$ | 4/3 | 1 | 1/3 | 0 | 0 | |
| 0 | $y_4$ | $x_4 = 5/3$ | −1/3 | 0 | −1/3 | 1 | 0 | |
| 0 | $g_1$ | $G_1 = -2/3$ | −2/3 | 0 | −2/3 | 0 | 1 | |
| | | $z_B = \frac{11.9}{3}$ | 0.267 | 0 | 1.7/3 | 0 | 0 | $\leftarrow \Delta_j$ |

We have $x_{B_r} = \text{Min}\,[x_{B_i}, x_{B_i} < 0] = \text{Min}\left[-\frac{2}{3}\right] = -\frac{2}{3} = x_{B_3}$. Hence $r = 3$. So $y_{B_3}$ i.e. $g_1$ leaves the basis. Now,

$$\text{Max}\left\{\frac{\Delta_j}{y_{rj}},\ y_{rj} < 0\right\} = \text{Max}\left\{\frac{0.267}{-\frac{2}{3}}, \frac{\frac{1.7}{3}}{-\frac{2}{3}}\right\}$$
$$= \text{Max}\{-0.801/2,\ -1.7/2\} = -0.801/2,$$

which corresponds to $y_1$.

Hence $y_1$ is the entering vector, and $y_{31} = -\frac{2}{3}$ is the key element.

Now we construct the next simplex table.

| | | $c_j$ | 2 | 1.7 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $g_1$ | |
| 1.7 | $y_2$ | $x_2 = 1$ | 0 | 1 | −1 | 0 | 2 | |
| 0 | $y_4$ | $x_4 = 2$ | 0 | 0 | 0 | 1 | −1/2 | |
| 2 | $y_1$ | $x_1 = 1$ | 1 | 0 | 1 | 0 | −3/2 | |
| | | $z_B = 3.7$ | 0 | 0 | 0.3 | 0 | 0.4 | $\leftarrow \Delta_j$ |

Since all $\Delta_j \geq 0$ and all $x_{B_i} \geq 0$, we obtain an optimal feasible integer solution. The optimal integer solution is $x_1 = 1,\ x_2 = 1$ and Max $z = 2 + 1.7 = 3.7$.

*Example 3* Find the optimum integer solution to the following LPP:

$$\text{Maximize } z = 7x_1 + 6x_2$$
$$\text{subject to}\quad \begin{aligned} 2x_1 + 3x_2 &\leq 12 \\ 6x_1 + 5x_2 &\leq 30 \end{aligned}$$
$$x_1,\ x_2 \geq 0 \text{ and are integers.}$$

**Solution**:

The given problem is a maximization problem.

Now introducing two slack variables $x_3(\geq 0)$ and $x_4(\geq 0)$, the standard form of the given LPP becomes

$$\text{Maximize } z = 7x_1 + 6x_2 + 0.x_3 + 0.x_4$$
$$\text{subject to } \quad 2x_1 + 3x_2 + x_3 = 12$$
$$6x_1 + 5x_2 + x_4 = 30, \quad x_1,\ x_2,\ x_3,\ x_4 \geq 0 \text{ and are integers.}$$

Now, ignoring the integer condition, we shall solve the problem by the simplex method.

Here the initial BFS is $x_3 = 12$, $x_4 = 30$ with the initial basis as $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Now we construct the initial simplex table.

| | | $c_j$ | 7 | 6 | 0 | 0 | Minimum ratio |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| 0 | $y_3$ | $x_3 = 12$ | 2 | 3 | 1 | 0 | 12/2 = 6 |
| 0 | $y_4$ | $x_4 = 30$ | 6 | 5 | 0 | 1 | 30/6 = 5 |
| | | $z_B = 0$ | −7 | −6 | 0 | 0 | $\leftarrow \Delta_j$ |
| | | | $\uparrow$ | | | $\downarrow$ | |

| | | $c_j$ | 7 | 6 | 0 | 0 | Minimum ratio |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | |
| 0 | $y_3$ | $x_3 = 2$ | 0 | 4/3 | 1 | −1/3 | 3/2 |
| 7 | $y_1$ | $x_1 = 5$ | 1 | 5/6 | 0 | 1/6 | 6 |
| | | $z_B = 35$ | 0 | −1/6 | 0 | 7/6 | $\leftarrow \Delta_j$ |
| | | | | $\uparrow$ | $\downarrow$ | | |

| | | $c_j$ | 7 | 6 | 0 | 0 |
|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ |
| 6 | $y_2$ | $x_2 = 3/2$ | 0 | 1 | 3/4 | −1/4 |
| 7 | $y_1$ | $x_1 = 15/4$ | 1 | 0 | −5/8 | 3/8 |
| | | $z_B = 141/4$ | 0 | 0 | 1/8 | 9/8 |

Since all $\Delta_j \geq 0$, an optimum solution is obtained, given by $x_2 = \frac{3}{2},\ x_1 = \frac{15}{4}$. The optimal solution is not an integer solution because $x_2 = \frac{3}{2},\ x_1 = \frac{15}{4}$, so we have to add a fractional cut constraint or Gomorian constraint in the optimal simplex table.

Since $x_{B_1} = \frac{3}{2} = 1 + \frac{1}{2}$ and $x_{B_2} = \frac{15}{4} = 3 + \frac{3}{4}$, we have $f_{B_1} = \frac{1}{2}$ and $f_{B_2} = \frac{3}{4}$. Hence $\text{Max}\left(f_{B_1}, f_{B_2}\right) = \frac{3}{4} = f_{B_2}$, which corresponds to the second row. So, the second row is the source row.

Now, from the second row of the optimal simplex table, we get

$$x_1 - x_3 + \frac{3}{8}x_3 + \frac{3}{8}x_4 = 3 + \frac{3}{4}.$$

Hence, the Gomorian constraint will be

$$\begin{aligned} &\frac{3}{8}x_3 + \frac{3}{8}x_4 \geq \frac{3}{4} \\ \text{or} \quad &-\frac{3}{8}x_3 - \frac{3}{8}x_4 \leq -\frac{3}{4} \\ \text{or} \quad &-\frac{3}{8}x_3 - \frac{3}{8}x_4 + G_1 = -\frac{3}{4}, \end{aligned}$$

where $G_1$ is the Gomorian slack variable.

Adding this new constraint at the bottom of the preceding optimal simplex table, we have used the dual simplex method to solve the problem.

| | | $c_j$ | 7 | 6 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $g_1$ | |
| 6 | $y_2$ | $x_2 = 3/2$ | 0 | 1 | 3/4 | −1/4 | 0 | |
| 7 | $y_1$ | $x_1 = 15/4$ | 1 | 0 | −5/8 | 3/8 | 0 | |
| 0 | $g_1$ | $G_1 = -3/4$ | 0 | 0 | −3/8 | −3/8 | 1 | |
| | | $z_B = 141/4$ | 0 | 0 | 1/8 | 9/8 | 0 | $\leftarrow \Delta_j$ |

We have $x_{B_r} = \text{Min}[x_{B_i},\ x_{B_i} < 0] = \text{Min}\left[-\frac{3}{4}\right] = -\frac{3}{4} = x_{B_3}$. Hence, $r = 3$. So $y_{B_3}$ i.e. $g_1$ leaves the basis. Now,

$$\begin{aligned} \text{Max}\left\{\frac{\Delta_j}{y_{rj}},\ y_{rj} < 0\right\} &= \text{Max}\left\{\frac{1/8}{-\frac{3}{8}}, \frac{9/8}{-\frac{3}{8}}\right\} \\ &= \text{Max}\{-1/3, -3\} = -1/3, \end{aligned}$$

which corresponds to $y_3$.

Hence $y_3$ is the entering vector, and $y_{33} = -\frac{3}{8}$ is the key element.

Now we construct the next simplex table.

| | | $c_j$ | 2 | 1.7 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $g_1$ | |
| 6 | $y_2$ | $x_2 = 0$ | 0 | 1 | 0 | −1 | 2 | |
| 7 | $y_1$ | $x_1 = 5$ | 1 | 0 | 0 | 1 | −5/3 | |
| 0 | $y_3$ | $x_3 = 4$ | 0 | 0 | 1 | 1 | −8/3 | |
| | | $z_B = 35$ | 0 | 0 | 0 | 1 | 1/3 | $\leftarrow \Delta_j$ |

Since all $\Delta_j \geq 0$ and all $x_{B_i} \geq 0$, we obtain an optimal feasible integer solution. The optimal integer solution is $x_1 = 5,\ x_2 = 0$ and $\text{Max}\, z = 5 \times 7 + 6 \times 0 = 35$.

## 8.5 Gomory's Mixed-Integer Cutting-Plane Method

In all-integer programming problems, the fractional cut or Gomory's constraint is formed under the assumption that all the variables including slack and surplus variables are integers. This means that the application of the fractional cut will not yield a feasible integer solution unless all variables assume integer values. The importance of this assumption is discussed in the following constraint:

$$x_1 + \frac{1}{3}x_2 \leq \frac{13}{2},\ x_1,\ x_2 \geq 0 \ \text{ and are integers.}$$

From the stand point of solving the associated integer linear programming program, the constraint is converted to an equation by introducing a non-negative slack variable $G_1$, i.e.

$$x_1 + \frac{1}{3}x_2 + G_1 = \frac{13}{2}.$$

From this equation, it is clear that it can have a feasible integer solution in $x_1$ and $x_2$ only if $G_1$ is non-integer. Thus, the application of the fractional cut will yield no feasible integer solution because all the variables $x_1$, $x_2$ and $G_1$ cannot be integers simultaneously.

To avoid this situation, a special cut, called Gomory's mixed-integer cut, is developed. For this cut, only a subset of variables is assumed to take integer values, and the remaining variables including slack and surplus variables are assumed as non-integer variables.

## 8.6 Construction of Additional Constraint for Mixed-Integer Programming Problems

Let us consider the following mixed-integer programming problem:

$$\text{Maximize } z = \sum_{j=1}^{x} c_j x_j$$

$$\text{subject to} \quad \sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = 1,\ 2,\ \ldots,\ m$$

$$\text{and } x_j \geq 0, \quad j = 1,\ 2,\ \ldots,\ n$$

where $x_j$ are integers for $j = 1,\ 2,\ \ldots,\ k(<n)$.

Let us suppose that the basic variable $x_r$ has the largest fractional value among the values of all other basic variables which are restricted to take integer values in the optimal simplex table (which is developed ignoring the integer restrictions of variables). Then, from the optimal simplex table, the $r$th constraint can be written as follows:

$$x_{B_r} = x_r + \sum_{j \in R} a_{rj} x_j \tag{8.4}$$

where $R$ is the set of indices corresponding to non-basic variables.

Let $x_{B_r} = [x_{B_r}] + f_r, \quad 0 < f_r < 1$

and $j^+ = \{j : a_{rj} \geq 0\}, \quad j^- = \{j : a_{rj} < 0\}$.

Then Eq. (8.4) can be rewritten as

$$\begin{aligned} [x_{B_r}] + f_r &= x_r + \sum_{j \in j^+} a_{rj} x_j + \sum_{j \in j^-} a_{rj} x_j \\ \text{or} \quad \sum_{j \in j^+} a_{rj} x_j + \sum_{j \in j^-} a_{rj} x_j &= [x_{B_r}] - x_r + f_r = I + f_r \end{aligned} \tag{8.5}$$

where $0 < f_r < 1$ and $I$ is the integer value.

Clearly, the left-hand side of (8.5) is either positive or negative according to whether $I + f_r$ is positive or negative respectively.

**Case I**: Let $I + f_r$ be positive.

In this case, $I + f_r$ must be greater than or equal to $f_r$,

$$\text{i.e.} \quad I + f_r \geq f_r.$$

Hence, from (8.5), we have

$$\sum_{j \in j^+} a_{rj} x_j + \sum_{j \in j^-} a_{rj} x_j \geq f_r \tag{8.6}$$

Since for $j \in j^-$, $a_{rj}$ are non-positive and $x_j \geq 0$, then

$$\begin{aligned} \sum_{j \in j^+} a_{rj} x_j &\geq \sum_{j \in j^+} a_{rj} x_j + \sum_{j \in j^-} a_{rj} x_j, \\ \text{i.e.} \quad \sum_{j \in j^+} a_{rj} x_j &\geq f_r \quad [\text{by } (8.6)] \end{aligned} \tag{8.7}$$

**Case II**: Let $I + f_r$ be negative.

In this case, the value of $I + f_r$ will be either $-1 + f_r$ or $-2 + f_r$ or $-3 + f_r$ or …

Therefore,

$$\begin{aligned}
&\sum_{j\in j^-} a_{rj}x_j \le \sum_{j\in j^+} a_{rj}x_j + \sum_{j\in j^-} a_{rj}x_j \le -1 + f_r \\
\text{or}\quad &\sum_{j\in j^-} a_{rj}x_j \le -1 + f_r \\
\text{or}\quad &\frac{1}{-1+f_r}\sum_{j\in j^-} a_{rj}x_j \ge 1 \quad [\text{Dividing both sides by } -1+f_r] \\
\text{or}\quad &\frac{f_r}{-1+f_r}\sum_{j\in j^-} a_{rj}x_j \ge f_r \quad [\text{Multiplying both sides by } f_r]
\end{aligned} \tag{8.8}$$

In both cases, the left-hand sides of inequalities (8.7) and (8.8) are positive and greater than or equal to $f_r$. Thus, any feasible solution of mixed-integer programming must satisfy the following inequality:

$$\sum_{j\in j^+} a_{rj}x_j + \frac{f_r}{-1+f_r}\sum_{j\in j^-} a_{rj}x_j \ge f_r \tag{8.9}$$

This inequality is not satisfied by the optimal solution of the LPP without the integer requirement, because by setting $x_j = 0\ \forall j$, the left-hand side becomes zero, whereas the right-hand side becomes positive.

Now, introducing a non-negative slack variable in (8.9), we get the following equation:

$$-\sum_{j\in j^+} a_{rj}x_j - \frac{f_r}{-1+f_r}\sum_{j\in j^-} a_{rj}x_j + G_1 = -f_r \tag{8.10}$$

This equation represents Gomory's constraint or Gomory's cut.

*Example 4* Solve the following mixed-integer programming problem:

$$\begin{aligned}
&\text{Maximize } z = 7x_1 + 9x_2 \\
&\text{subject to } \quad -x_1 + 3x_2 \le 6 \\
&\qquad\qquad\qquad 7x_1 + x_2 \le 35;\ x_1,\ x_2 \ge 0 \text{ and } x_1 \text{ is an integer.}
\end{aligned}$$

**Solution**:

Ignoring the integer restriction, we get the optimal simplex table as follows:

| | | $c_j$ | 7 | 9 | 0 | 0 |
|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ |
| 9 | $y_2$ | 7/2 | 0 | 1 | 7/22 | 1/22 |
| 7 | $y_1$ | 9/2 | 1 | 0 | −1/22 | 3/22 |
| $z_B = 63$ | | | 0 | 0 | 28/11 | 15/11 |

From the preceding table, it is observed that the optimal value of $x_1$ is not an integer. Hence from the second row, we have

$$\frac{9}{2} = x_1 - \frac{1}{22}x_3 + \frac{3}{22}x_4$$

$$\text{or} \quad \left(4 + \frac{1}{2}\right) = x_1 - \frac{1}{22}x_3 + \frac{3}{22}x_4.$$

Here $r = 2,\ f_{Br} = \frac{1}{2}$.

Hence, the corresponding fractional cut is

$$-\sum_{j \in j^+} a_{rj}x_j - \frac{f_{Br}}{-1 + f_{Br}} \sum_{j \in j^-} a_{rj}x_j + G_1 = -f_{Br},$$

where $j^+ = \{j : a_{rj} \geq 0\}$ and $j^- = \{j : a_{rj} < 0\}$ and $x_j$ are non-basic variables in the optimal simplex table

$$\text{or} \quad -\frac{3}{22}x_4 - \frac{\frac{1}{2}}{-1 + \frac{1}{2}}\left(-\frac{1}{22}\right)x_3 + G_1 = -\frac{1}{2}$$

$$\text{or} \quad -\frac{3}{22}x_4 - \frac{1}{22}x_3 + G_1 = -\frac{1}{2}.$$

Adding this new constraint at bottom of the preceding simplex table, we have used the dual simplex method to solve the problem.

| | | $c_j$ | 7 | 9 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $g_1$ |
| 9 | $y_2$ | 7/2 | 0 | 1 | 7/22 | 1/22 | 0 |
| 7 | $y_1$ | 9/2 | 1 | 0 | −1/22 | 3/22 | 0 |
| 0 | $g_1$ | −1/2 | 0 | 0 | −1/22 | −3/22 | 1 |
| | $z_B = 63$ | | 0 | 0 | 28/11 | 15/11 | 0 |

Since $x_{Br} = \text{Min}[x_{Bi}, x_{Bi} < 0] = \text{Min}\left[-\frac{1}{2}\right] = -x_{B3}$, then $B_3$ or $g_1$ leaves the basis.

Now $\text{Max}\left\{\frac{\Delta_j}{y_{3j}}, y_{3j} < 0\right\} = \text{Max}\left\{\frac{28/11}{-1/22}, \frac{15/11}{-3/22}\right\} = \text{Max}\{-56, -10\} = -10$, which corresponds to $y_4$.

Hence, $y_4$ enters the basis and $g_1$ leaves the basis.

| | | $c_j$ | 7 | 9 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $g_1$ |
| 9 | $y_2$ | 10/3 | 0 | 1 | 10/33 | 0 | 1/3 |
| 7 | $y_1$ | 4 | 1 | 0 | −1/11 | 0 | 1 |
| 0 | $y_4$ | 11/3 | 0 | 0 | 1/3 | 1 | −22/3 |
| | $z_B$ = 58 | | 0 | 0 | 23/11 | 0 | 10 |

The required optimum solution is $x_1 = 4,\ x_2 = \frac{10}{3}$ and Max $z = 58$ with $x_1$ an integer.

*Example 5* Solve the following mixed-integer programming problem:

$$\begin{aligned} &\text{Maximize } z = x_1 + x_2 \\ &\text{subject to} \quad 2x_1 + 5x_2 \le 16 \\ &\qquad\qquad 6x_1 + 5x_2 \le 30;\ x_2 \ge 0 \text{ and } x_1 \text{ is a non-negative integer.} \end{aligned}$$

**Solution**:

Ignoring the integer restriction, we get the optimal simplex table as follows:

| | | $c_j$ | 1 | 1 | 0 | 0 |
|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ |
| 1 | $y_2$ | 9/5 | 0 | 1 | 3/10 | −1/10 |
| 1 | $y_1$ | 7/2 | 1 | 0 | −1/4 | 1/4 |
| $z_B$ = 53/10 | | | 0 | 0 | 1/20 | 3/20 |

From the preceding table, it is observed that the optimal value of $x_1$ is not an integer. Hence from the second row, we have

$$\frac{7}{2} = x_1 + 0.x_2 - \frac{1}{4}x_3 + \frac{1}{4}x_4$$
$$\text{or} \quad \left(3 + \frac{1}{2}\right) = x_1 - \frac{1}{4}x_3 + \frac{1}{4}x_4.$$

Here $r$ = 2, $f_{Br} = \frac{1}{2}$.

Hence, the corresponding fractional cut is

$$-\sum_{j \in j^+} a_{rj}x_j - \frac{f_{Br}}{-1 + f_{Br}} \sum_{j \in j^-} a_{rj}x_j + G_1 = -f_{Br},$$

where $j^+ = \{j : a_{rj} \geq 0\}$ and $j^- = \{j : a_{rj} < 0\}$ and $x_j$ are non-basic variables in the optimal simplex table.

$$\text{or} \quad -\frac{1}{4}x_4 - \frac{\frac{1}{2}}{-1+\frac{1}{2}}\left(-\frac{1}{4}\right)x_3 + G_1 = -\frac{1}{2}$$

$$\text{or} \quad -\frac{1}{4}x_4 - \frac{1}{4}x_3 + G_1 = -\frac{1}{2}.$$

Adding this new constraint at bottom of the preceding simplex table, we have used the dual simplex method to solve the problem.

| | | $c_j$ | 1 | 1 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $g_1$ |
| 1 | $y_2$ | 9/5 | 0 | 1 | 3/10 | −1/10 | 0 |
| 1 | $y_1$ | 7/2 | 1 | 0 | −1/4 | 1/4 | 0 |
| 0 | $g_1$ | −1/2 | 0 | 0 | −1/4 | −1/4 | 1 |
| | $z_B$ = 53/10 | | 0 | 0 | 1/20 | 3/20 | 0 |

Since $x_{Br} = \text{Min}[x_{B_i},\ x_{B_i} < 0] = \text{Min}\left[-\frac{1}{2}\right] = -\frac{1}{2} = -x_{B_3}$ then $y_{B_3}$ i.e. $g_1$ leaves the basis.

Now, $\text{Max}\left\{\frac{\Delta_j}{y_{3j}}, y_{3j} < 0\right\} = \text{Max}\left\{\frac{1/20}{-1/4}, \frac{3/20}{-1/4}\right\} = \text{Max}\{-1/5, -3/5\} = -1/5$ which corresponds to $y_3$.

Hence, $y_3$ enters the basis and $g_1$ leaves the basis.

| | | $c_j$ | 1 | 1 | 0 | 0 | 0 |
|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $g_1$ |
| 1 | $y_2$ | 6/5 | 0 | 1 | 0 | −4/10 | 6/5 |
| 1 | $y_1$ | 4 | 1 | 0 | 0 | 1/2 | −1 |
| 0 | $y_3$ | 2 | 0 | 0 | 1 | 1 | −4 |
| | $z_B$ = 26/5 | | 0 | 0 | 0 | 1/10 | 1/5 |

The required optimum solution is $x_1 = 4,\ x_2 = \frac{6}{5}$ and Max $z = 26/5$ with $x_1$ as integer.

## 8.7 Branch and Bound Method

This method was first developed by A. H. Land and A. G. Doig, and it was further studied by J. D. C. Little and other researchers. This method can be used to solve all-integer, mixed-integer and zero-one LPPs.

The concept behind this method is to divide the entire feasible region of the LPP into smaller parts by forming (or branching) subproblems and then examine each successively until a feasible solution which gives the optimal solution is obtained. Here, we shall discuss its application to all-integer programming problems.

**Branch and Bound (B & B) Algorithm**

The iterative procedure of this method is summarized here for solving an integer programming problem (maximization case):

Let us consider the following all-integer programming problem (AIPP):

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

$$\text{subject to } \quad \sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = 1,\ 2,\ \ldots,\ m$$

$$x_j \geq 0\ (j = 1,\ 2,\ \ldots,\ n) \text{ are integers.}$$

Step 1 Ignoring the integer restrictions of the variables, obtain the optimal solution of the given LPP (we denote this LPP as LP-1).

Step 2 Test the integrality of the optimum solution obtained in Step 1. There may arise different possibilities.

(i) If the solution of LP-1 is infeasible or unbounded, then the solution to the given AIPP is also infeasible or unbounded respectively. Then stop the process.

(ii) If the optimum solution of LP-1 satisfies the integer restrictions, then the current solution is optimum to the given AIPP. Stop the process.

(iii) If one or more basic variables of LP-1 do not satisfy the integer restrictions, then go to Step 3. Let the optimal value of the objective function of LP-1 be $z_1$. This value provides an initial upper bound.

Step 3 Branching step

(i) The value of $z_1$ will be considered as the initial upper bound of the objective function value for the given AIPP. Let it be denoted by $z_U$. The lower bound of the objective function of the given integer LPP can be obtained by rounding off to integer all values of the variables.

(ii) Let the optimum value $x_k^*$ of the basic variable $x_k$ not be an integer. This selection can be done by considering the largest fractional value of the basic variables whose values are not integer, or it can be done arbitrarily.

(iii) Branch or partition the problem LP-1 into two new subproblems (called nodes) based on the integer value of $x_k$. This can be done by adding two mutually exclusive constraints

$$x_k \le \left[x_k^*\right] \text{ and } x_k \ge \left[x_k^*\right] + 1$$

to the LPP LP-1, where $\left[x_k^*\right]$ is the largest integer value not greater than $x_k^*$.
In this case, the two new LPPs (LP-2 and LP-3) are as follows:

$$\text{LP-2: Maximize } z = \sum_{j=1}^{n} c_j x_j$$

$$\text{subject to } \sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = 1,\ 2,\ \ldots,\ m$$

$$x_k > \left[x_k^*\right] + 1$$

$$\text{and } x_j \ge 0, \quad j = 1,\ 2,\ \ldots,\ n.$$

$$\text{LP-3: Maximize } z = \sum_{j=1}^{n} c_j x_j$$

$$\text{subject to } \sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = 1,\ 2,\ \ldots,\ m$$

$$x_k \ge \left[x_k^*\right] + 1$$

$$\text{and } x_j \ge 0, \quad j = 1,\ 2,\ \ldots,\ n.$$

Step 4 Bound step
Solve the two subproblems LP-2 and LP-3 obtained in Step 3. Let the optimal value of the objective function of LP-2 be $z_2$ and that of LP-3 be $z_3$. For the best integer solution (if it exists), the objective function value becomes the lower bound of the IPP objective function value (initially this is the rounding off value). Let the lower bound be denoted by $z_L$.

Step 5 Fathoming step
Examine the solutions of both LP-2 and LP-3, which contain the optimal solution. There may arise different cases:

(i) If the optimal solutions of the two subproblems are integral, then the required solution is the one that gives the larger value of $z$.
(ii) If the optimal solution of one subproblem is integral and the other subproblem has no feasible solution, the required solution is the same as that of the subproblem having an integer-valued solution.
(iii) If the optimal solution of one-subproblem is integral while that of the other is not integer-valued, then repeat Steps 3 and 4 for the non-integer-valued subproblem.

When the solution of any subproblem is either integer-valued or infeasible, then the subproblem or node is said to be fathomed.

If the solution of any subproblem is infeasible, then it is said to be fathomed by infeasibility. If $z_L \ge z_U$, it is said to be fathomed by bounding.

If $z_L$ is obtained as an integer feasible solution in any subproblem and $z_L < z_U$, it is said to be fathomed by integrality.

Step 6 Repeat Steps 3 through 5 until all integer-valued solutions are recorded.
Step 7 Choose the solution among the recorded integer-valued solutions that yields an optimum value of $z$.
Step 8 Stop.

Now, we shall illustrate the B & B algorithm by the following numerical example.

*Example 6* Solve the following IPP by the branch and bound method:

$$\begin{aligned}&\text{Maximize } z = 5x_1 + 4x_2\\&\text{subject to}\\&\qquad x_1 + x_2 \le 5\\&\qquad 10x_1 + 6x_2 \le 45\\&\qquad x_1, x_2 \ge 0 \text{ and they are integers.}\end{aligned}$$

**Solution**:

Let us consider the following IPP:

$$\begin{aligned}&\text{Maximize } z = 5x_1 + 4x_2\\&\text{subject to}\\&\qquad x_1 + x_2 \le 5\\&\qquad 10x_1 + 6x_2 \le 45\\&\qquad x_1, x_2 \ge 0 \text{ and are integers.}\end{aligned}$$

Ignoring the integer restrictions of the variables $x_1$ and $x_2$, the associated problem of the given IPP is denoted by LP-1 as follows:

$$\left.\begin{aligned}&\text{Maximize } z = 5x_1 + 4x_2\\&\text{subject to}\quad x_1 + x_2 \le 5,\ 10x_1 + 6x_2 \le 45,\ x_1,\ x_2 \ge 0\end{aligned}\right\}\text{(LP-1)}$$

Its optimum solution is $x_1 = 3.75,\ x_2 = 1.25$ and $z = 23.75$ (see Fig. 8.1).

The optimum solution of LP-1 does not satisfy the integer requirements. Hence, this solution is not optimal for the given IPP. Hence, the initial upper bound of the objective function to the given problem is given by $z_u = 23.75$. Again, the initial lower bound can also be obtained by rounding of the values of $x_1$ and $x_2$ to the nearest integer, i.e. $x_1 = 4,\ x_2 = 1$, which does not satisfy the second constraint. Now, we first select one of the integer variables whose value at LP-1 is not integer.

We select $x_1 (= 3.75)$ as it contains the largest fractional part. The region $3 < x_1 < 4$ of the LP-1 solution space does not contain any integer value. Hence, we can eliminate this region. So, we can write two new restrictions $x_1 \leq 3$ and $x_1 \geq 4$ which are mutually exclusive. As a result, we get two new LPPs, LP-2 and LP-3 as follows:

**LP-2**: Maximize $z = 5x_1 + 4x_2$

subject to $x_1 + x_2 \leq 5$

$10x_1 + 6x_2 \leq 45$

$x_1 \leq 3$ and $x_1, \ x_2 \geq 0$

[Here LP-2 solution space = LP-1 solution space + $(x \leq 3)$]

The optimal solution of LP-2 is given by $x_1 = 3$, $x_2 = 2$ and $z = 23$ (Fig. 8.2).

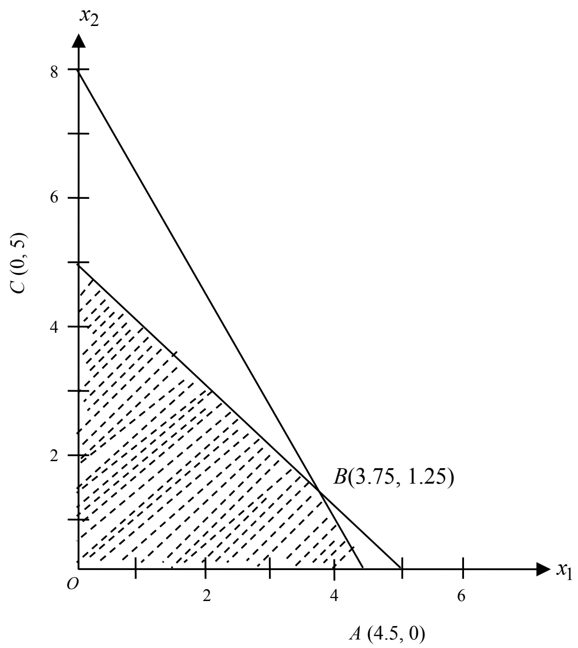
**LP-3**: Maximize $z = 5x_1 + 4x_2$

subject to $x_1 + x_2 \leq 5$

$10x_1 + 6x_2 \leq 45$

$x_1 \geq 4$ and $x_1, \ x_2 \geq 0$

**Fig. 8.1** Graphical representation of LP-1



[Here LP-3 solution space = LP-1 solution space + $(x_1 \geq 4)$]

The optimal solution of LP-3 is given by $x_1 = 4$, $x_2 = 0.83$ and $z = 23.33$.

Since this solution does not satisfy the integer restriction, we can write two new constraints, $x_2 \leq [0.83]$ and $x_2 \geq [0.83] + 1$, i.e. $x_2 \leq 0$ and $x_2 \geq 1$. These constraints are mutually exclusive. As a result, we get two new LPPs, LP-4 and LP-5 as follows:

**LP−4**: Maximize $z = 5x_1 + 4x_2$

subject to $x_1 + x_2 \leq 5$

$10x_1 + 6x_2 \leq 45$

$x_1 \geq 4,\ x_2 \leq 0.$

$$\begin{aligned}[\text{Here, LP-4 solution space} &= \text{LP-3 solution space} + (x_2 \leq 0)\\ &= \text{LP-1 solution space} + (x_1 > 4) + (x_2 \leq 0)]\end{aligned}$$

The optimal solution of LP-4 is $x_1 = 4.5$, $x_2 = 0$ and $z = 22.5$.

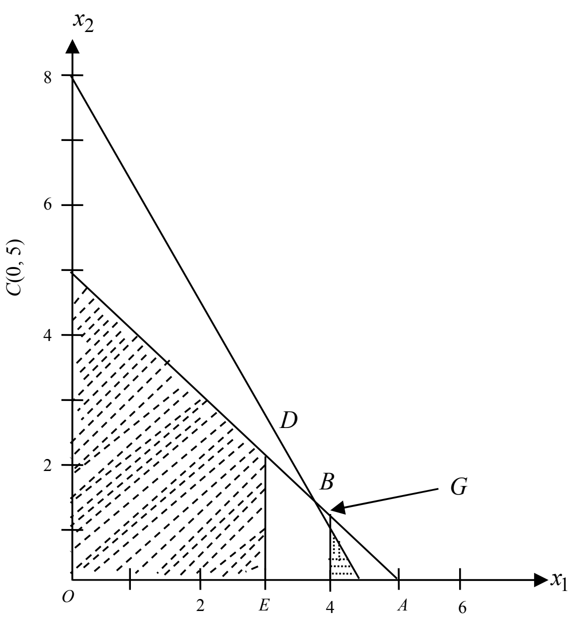
**LP**−5: Maximize $z = 5x_1 + 4x_2$

subject to $x_1 + x_2 \leq 5$

$10x_1 + 6x_2 \leq 45$

$x_1 \geq 4,\ x_2 \geq 1.$

**Fig. 8.2** Graphical representation of LP-2 & LP-3



[Here LP-5 solution space $=$ LP-3 solution space $+ (x_2 \geq 1)$
$=$ LP-1 solution space $+ (x_1 \geq 4) + (x_2 \geq 1)$]

LP-5 has no feasible solution. Since the solution of LP-4 does not satisfy the integer restriction, we can write two new mutually exclusive constraints $x_1$<=4 and $x_1$>=5. As a result, we get two new LPPs, LP-6 and LP-7 as follows:

**LP-6**: Maximize $z = 5x_1 + 4x_2$

subject to $x_1 + x_2 \leq 5$

$10x_1 + 6x_2 \leq 45$

$x_1 \leq 4,\ x_2 \leq 0$ and $x_1,\ x_2 \geq 0.$

[Here LP-6 solution space $=$ LP-4 solution space $+ (x_1 \leq 4)$
$=$ LP-1 solution space $+ (x_1 \geq 4) + (x_2 \leq 0) + (x_1 \leq 4)$]

The optimal solution of LP-6 is $x_1 = 4$, $x_2 = 0$ and $z = 20$.

**LP-7**: Maximize $z = 5x_1 + 4x_2$

subject to $x_1 + x_2 \leq 5$

$10x_1 + 6x_2 \leq 45$

$x_1 \geq 5,\ x_2 \leq 0$

[Here LP-7 solution space $=$ LP-4 solution space $+ (x \geq 5)$]
LP-7 has no feasible solution.

Hence, the optimal solution of the given AIPP is $x_1 = 3,\ x_2 = 2$ and Max $z = 23$.

## 8.8 Exercises

1. Find the optimum integer solution to the following all-integer programming problems:

   (i) Maximize $z = x_1 + 2x_2$

   subject to $2x_2 \leq 7,$

   $x_1 + x_2 \leq 7,$

   $2x_1 \leq 11$

   $x_1 \geq 0,\ x_2 \geq 0$ and they are integers.

(ii) Maximize $z = x_1 + 5x_2$

subject to $x_1 + 5x_2 \leq 20,$

$x_1 \leq 2$

$x_1 \geq 0,\ x_2 \geq 0$ and they are integers.

(iii) Maximize $z = x_1 + x_2$

subject to $3x_1 - 2x_2 \leq 5,$

$x_1 \leq 2$

$x_1 \geq 0, x_2 \geq 0$ and they are integers.

(iv) Maximize $z = 3x_1 + 4x_2$

subject to $3x_1 + 2x_2 \leq 8,$

$x_1 + 4x_2 \geq 10$

$x_1 \geq 0,\ x_2 \geq 0$ and they are integers.

$\left[\begin{array}{l}\text{Answer: } x_1 = 0, \\ {}_2 = 4,\ \text{Max } z = 16\end{array}\right]$

(v) Maximize $z = 2x_1 + 2x_2$

subject to $5x_1 + 3x_2 \leq 8$

$x_1 + 2x_2 \leq 4$

$x_1 \geq 0,\ x_2 \geq 0$ and they are integers.

(vi) Maximize $z = x_1 - x_2$

subject to $x_1 + 2x_2 \leq 4,$

$6x_1 + 2x_2 \leq 9$

$x_1 \geq 0,\ x_2 \geq 0$ and they are integers.

2. Find the optimum integer solution to the following mixed-integer programming problems.

(i) Maximize $z = 1.5x_1 + 3x_2 + 4x_3$

subject to $2.5x_1 + 2x_2 + 4x_3 \leq 12$

$2x_1 + 4x_2 - x_3 \leq 7$

$x_1,\ x_2,\ x_3 \geq 0$ and $x_3$ is an integer.

(ii) Maximize $z = 2x_1 + 6x_2$

subject to $x_1 + x_2 \leq 5$

$-x_1 + 2x_2 \leq 6$

$x_1, x_2 \geq 0$ and $x_2$ is an integer.

(iii) Maximize $z = x_1 + 6x_2$

subject to $3x_1 + 2x_2 \leq 5$

$x_2 \leq 2$

$x_1,\ x_2 \geq 0$ and $x_1$ is an integer.

3. Use the branch and bound method to solve the following all-integer programming problems.

(i) Maximize $z = x_1 + x_2$

subject to $4x_1 - x_2 \leq 10$

$2x_1 + 5x_2 \leq 10$

$x_1 \geq 0,\ x_2 \geq 0$ and they are integers.

(ii) Maximize $z = 7x_1 + 9x_2$

subject to $-x_1 + 3x_2 \leq 6$

$7x_1 + x_2 \leq 35$

$0 \leq x_1,\ x_2 \leq 7$ and they are integers.

(iii) Maximize $z = 21x_1 + 11x_2$

subject to $7x_1 + 4x_2 + x_3 = 13$

$x_2 \leq 5$

$x_1,\ x_2,\ x_3 \geq 0$ and they are integers.

(iv) Maximize $z = 4x_1 + 3x_2$

subject to $x_1 + 2x_2 \leq 4$

$2x_1 + x_2 \leq 6$

$x_1 \geq 0,\ x_2 \geq 0$ and they are integers.

# Chapter 9
# Basics of Unconstrained Optimization

## 9.1 Objectives

The objectives of this chapter are to:

- Discuss different techniques for solving unconstrained optimization problems of a single variable
- Study the necessary and sufficient conditions for optimizing a function of several variables.

## 9.2 Optimization of Functions of a Single Variable

Let $I \subset R$ be any interval, finite or infinite, and let $f$: $I \to R$ be a real-valued function.

**Definition**

(i) $x^* \in I$ is said to be a local minimum or minimizer of $f$ if $\exists$ $a\,\delta > 0$ such that $f(x^*) \le f(x)$, $\forall x \in (x^* - \delta, x^* + \delta)$.

(ii) $x^* \in I$ is said to be a global minimum or minimizer of $f$ if $f(x^*) \le f(x)$, $\forall x \in I$.

***Necessary condition for optimality***

The condition for optimality of $f(x)$ at the point $x = x^*$ is $f'(x^*) = 0$.

This is a necessary condition but not sufficient.

(i) For example, if $f(x) = x^3$, then at $x = 0$, $f'(0) = 0$.
But when $x > 0$, $f(x) > f(0)$ and $f(x) < f(0)$ for $x < 0$.

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_9

Here $f(x)$ has no extreme value at $x = 0$.

(ii) Even when $f'(x^*)$ does not exist, $f(x^*)$ may be optimum. For example, if $f(x) = |x|$, then obviously $f(0)$ is a minimum value of $f(x)$, but $f'(0)$ does not exist.

***Sufficient condition for optimality***

The sufficient condition for optimality of $f(x)$ is as follows: If at $x = x^*$ for the function $f(x)$:

(i) $f'(x^*) = f''(x^*) = \cdots = f^{n-1}(x^*) = 0$
(ii) $f^n(x^*)$ exists and $\neq 0$ then at $x = x^*$, $f(x)$ has
(a) no extreme value if $n$ is odd
(b) an extreme value if $n$ is even, i.e. a maximum if $f^n(x^*) < 0$, minimum if $f^n(x^*) > 0$

***Alternative approach***

**Theorem** If

(i) $f(x)$ is defined at $x = x^*$.
(ii) $f'(x)$ exists in a suitably small neighbourhood of $0 < |x - x^*| < \delta$ [$f'(x)$ may or may not exist].
(iii) $f'(x)$ has a fixed sign when $x < x^*$, also a unique sign when $x < x^*$.

Then with the increase of $x$ through $x^*$,

(a) $f(x)$ has no extreme value at $x = x^*$ if the sign of $f'(x)$ does not change.
(b) $f(x)$ has a maximum value at $x = x^*$ if the sign of $f'(x)$ changes from positive to negative.
(c) $f(x)$ has a minimum value at $x = x^*$ if the sign of $f'(x)$ changes from negative to positive.

**Note**: This method does not require higher order derivative information or the derivative information at the point where the optimality is tested.

*Example 1* Test for optimality at $x = 0$ for $f(x) = \sin x - x + \frac{x^3}{3!} - \frac{x^5}{5!}$.

**Solution** Here

$$\begin{aligned}
f(x) &= \sin x - x + \tfrac{x^3}{3!} - \tfrac{x^5}{5!} \\
f'(x) &= \cos x - 1 + \tfrac{x^2}{2!} - \tfrac{x^4}{4!}, && \therefore f'(0) = 0 \\
f''(x) &= -\sin x + x - \tfrac{x^3}{3!}, && \therefore f''(0) = 0 \\
f'''(x) &= -\cos x + 1 - \tfrac{x^2}{2!}, && \therefore f'''(0) = 0 \\
f^{iv}(x) &= \sin x - x, && \therefore f^{iv}(0) = 0 \\
f^{v}(x) &= \cos x - 1, && \therefore f^{v}(0) = 0 \\
f^{vi}(x) &= -\sin x, && \therefore f^{vi}(0) = 0 \\
f^{vii}(x) &= -\cos x, && \therefore f^{vii}(0) = -1 \neq 0.
\end{aligned}$$

Since the order of the first non-zero derivative is odd, then $f(x)$ is neither maximum nor minimum at $x = 0$.

*Example 2* If $f'(x) = (x-a)^{2m}(x-b)^{2n+1}$, where $m$ and $n$ are positive integers, show that $x = a$ gives neither a maximum nor a minimum value of $f(x)$. But $x = b$ gives a minimum.

**Solution** Let $h$ be any positive number, however small.

$$\begin{aligned}\text{Now, } f'(a-h) &= (a-h-a)^{2m}(a-h-b)^{2n+1} = (-h)^{2m}(a-b-h)^{2n+1}\\ &= h^{2m}(a-b-h)^{2n+1}\end{aligned}$$

$$\text{and } f'(a+h) = (a+h-a)^{2m}(a+h-b)^{2n+1} = h^{2m}(a-b+h)^{2n+1}.$$

Since $h$ is a very small quantity, $(a-b-h)^{2n+1}$ and $(a-b+h)^{2n+1}$ have the same sign.

$\therefore$ $f'(a-h)$ and $f'(a+h)$ have the same sign.
$\therefore$ $f(x)$ is neither maximum nor minimum at $x = a$.

Again,

$$\begin{aligned}f'(b-h) &= (b-h-a)^{2m}(b-h-b)^{2n+1}\\ &= (b-a-h)^{2m}(-h)^{2n+1}\\ &= -h^{2n+1}(b-a-h)^{2m}\\ f'(b+h) &= (b+h-a)^{2m}(b+h-b)^{2n+1}\\ &= h^{2n+1}(b-a+h)^{2m}.\end{aligned}$$

Since $h$ is a very small quantity, $(b-a+h)^{2m}$ and $(b-a-h)^{2m}$ have the same sign.

$$\therefore f'(b-h) < 0 \text{ and } f'(b+h) > 0,$$

i.e. $f'(x)$ changes its sign from negative to positive. Therefore, $f(x)$ is minimum at $x = b$.

*Example 3* If $f(x) = |x|$, show that $f(x)$ is minimum at $x = 0$.

**Solution** Since $f(x) = |x| \quad \therefore f(0) = 0$.

Let $h$ be any positive number, however small.

$$\text{Now, } f(0-h) = f(-h) = |-h| = h > 0 = f(0)$$
$$\therefore f(0-h) > f(0).$$
$$\text{Again, } f(0+h) = f(h) = |h| = h > 0 = f(0)$$
$$\therefore f(0+h) > f(0).$$

Hence, $f(x)$ has a minimum at $x = 0$.

*Example 4* Show that $(x-a)^{\frac{1}{3}}(2x-a)^{\frac{2}{3}}$ is maximum for $x = \frac{a}{2}$ and neither maximum nor minimum for $x = a$ $(a > 0)$.

**Solution** Let $f(x) = (x-a)^{\frac{1}{3}}(2x-a)^{\frac{2}{3}}$ $\quad \therefore f\left(\frac{a}{2}\right) = 0$ and $f(a) = 0$.

Again, let $h$ be any positive number, however small.

$$\begin{aligned} \therefore f(a-h) &= (a-h-a)^{\frac{1}{3}}(2a-2h-a)^{\frac{2}{3}} \\ &= (-h)^{\frac{1}{3}}(a-2h)^{\frac{2}{3}} = -h^{\frac{1}{3}}(a-2h)^{\frac{2}{3}} < 0 = f(a) \end{aligned}$$

$$\begin{aligned} f(a+h) &= (a+h-a)^{\frac{1}{3}}(2a+2h-a)^{\frac{2}{3}} \\ &= h^{\frac{1}{3}}(a+2h)^{\frac{2}{3}} > 0 = f(a) \quad \therefore f(a-h) < f(a) < f(a+h). \end{aligned}$$

Hence, $f(x)$ is neither maximum nor minimum for $x = a$ $(a > 0)$.

$$\begin{aligned} \text{Again, } f\left(\frac{a}{2}+h\right) &= \left(\frac{a}{2}+h-a\right)^{\frac{1}{3}}\left[2\left(\frac{a}{2}+h\right)-a\right]^{\frac{2}{3}} \\ &= \left\{-\left(\frac{a}{2}-h\right)\right\}^{\frac{1}{3}}(2h)^{\frac{2}{3}} \\ &= -\left(\frac{a}{2}-h\right)^{\frac{1}{3}}(2h)^{\frac{2}{3}} < 0 = f\left(\frac{a}{2}\right). \end{aligned}$$

$$\begin{aligned} \text{Now, } f\left(\frac{a}{2}-h\right) &= \left(\frac{a}{2}-h-a\right)^{\frac{1}{3}}\left[2\left(\frac{a}{2}-h\right)-a\right]^{\frac{2}{3}} \\ &= \left(-h-\frac{a}{2}\right)^{\frac{1}{3}}(a-2h-a)^{\frac{2}{3}} = (-1)^{\frac{1}{3}}\left(h+\frac{a}{2}\right)^{\frac{1}{3}}(-2h)^{\frac{2}{3}} \\ &= -\left(\frac{a}{2}+h\right)^{\frac{1}{3}}(2h)^{\frac{2}{3}} \\ &= \text{negative} \left[\because \left(\frac{a}{2}+h\right)^{\frac{1}{3}} > 0 \text{ and } (2h)^{\frac{2}{3}} > 0 \text{ as } h > 0\right. \\ \therefore f\left(\frac{a}{2}-h\right) &< 0 = f\left(\frac{a}{2}\right). \end{aligned}$$

$\therefore$ $f(x)$ is maximum at $x = \frac{a}{2}$.

## 9.3 Optimization of Functions of Several Variables

Here we consider the optimization problem

$$\begin{array}{ll} \text{Minimize} & f(x) \\ \text{subject to} & x \in \Omega, \end{array} \tag{9.1}$$

where the function $f: \mathbb{R}^n \to \mathbb{R}$ is a real-valued function, called the objective function. The vector $x$ is an $n$-tuple vector of independent variables $x_1, x_2, \ldots, x_n$. The set $\Omega$ is a subset of $\mathbb{R}^n$, called the feasible set.

There are also optimization problems that require maximization of the objective function, in which case we seek maximizers. Maximization problems, however, can be represented equivalently in the minimization form above because maximizing $f$ is equivalent to minimizing $-f$. Therefore, we can confine our attention to minimization problems without loss of generality.

**Definition** Let $f: \Omega \subseteq R^n \to R$ be a real-valued function. A point $x^* \in \Omega$ is said to be a:

(i) Local minimizer of $f$ if $\exists\, a\, \delta > 0$ such that $f(x^*) \le f(x), \quad \forall x \in \Omega \cap B(x^*, \delta)$, where $B(x^*, \delta)$ is an open ball whose centre is at $x^*$ and radius $\delta$
(ii) Global minimizer of $f$ if $f(x^*) \le f(x), \quad \forall x \in \Omega$.

If strict inequality holds in the preceding definition, $x^*$ is said to be a strict local minimizer/global minimizer.

***Necessary condition for optimality***

**Theorem** *The necessary condition for the continuous function f (x) to have an extreme point at $x = x^*$ is that the gradient $\nabla f(x^*) = 0$,*

$$\text{i.e. } \frac{\partial f(x^*)}{\partial x_1} = \frac{\partial f(x^*)}{\partial x_2} = \cdots = \frac{\partial f(x^*)}{\partial x_n} = 0 \tag{9.2}$$

*Proof* From the Taylor series expansion, we have

$$f(x+h) = f(x) + \langle h, \nabla f(x) \rangle + \frac{1}{2!} \langle H(x+\theta h)h, h \rangle \quad 0 < \theta < 1$$

Setting $x = x^*$ in this expression, we have

$$\begin{aligned} f(x^* + h) &= f(x^*) + \langle h, \nabla f(x^*) \rangle + \frac{1}{2!} \langle H(x^* + \theta h)h, h \rangle \\ \Rightarrow f(x^* + h) - f(x^*) &= \sum_{i=1}^{n} h_i \frac{\partial f(x^*)}{\partial x_i} + \frac{1}{2!} \langle H(x^* + \theta h)h, h \rangle. \end{aligned}$$

Since the term $\langle H(x^* + \theta h)h, h\rangle$ is of order $h_i^2$, so the term of order $h$ only will dominate the higher order terms for small $h$.

Thus, the sign of $f(x^* + h) - f(x^*)$ is decided by the sign of $h_i \frac{\partial f(x^*)}{\partial x_i}$.

Let us suppose that $\frac{\partial f(x^*)}{\partial x_i} > 0$; then the sign of $f(x^* + h) - f(x^*)$ will be positive for $h_i > 0$ ($i = 1, 2, \ldots, n$),

i.e. $f(x^* + h) > f(x^*)$ and negative for $h_i < 0$ ($i = 1, 2, \ldots, n$),

i.e. $f(x^* + h) < f(x^*)$.

This means that $x^*$ cannot be an extreme point. The same conclusion can be obtained if we assume that $\frac{\partial f(x^*)}{\partial x_i} < 0$.

These two conclusions contradict our original statement that $x^*$ is an extreme point.

Hence, we conclude that

$$\frac{\partial f(x^*)}{\partial x_i} = 0, \quad i = 1,\ 2,\ \ldots,\ n,$$
$$\text{i.e. } \nabla f(x^*) = 0.$$

***Sufficient condition for optimality***

**Theorem** *A sufficient condition for a stationary point $x^*$ to be an extreme point is that the Hessian matrix of $f(x)$ evaluated at $x^*$ is:*

(i) Positive definite when $x^*$ is a local minimizer.
(ii) Negative definite when $x^*$ is a local maximizer.

*Proof* From Taylor's theorem we can write

$$f(x^* + h) = f(x^*) + \sum_{i=1}^{n} h_i \frac{\partial f(x^*)}{\partial x_i} + \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j \frac{\partial^2 f(x^* + \theta h)}{\partial x_i \partial x_j}, \quad 0 < \theta < 1. \tag{9.3}$$

Since $x^*$ is an extreme point, from the necessary conditions for $f(x)$ to have an extreme point, we have $\frac{\partial f(x^*)}{\partial x_i} = 0, \quad i = 1,\ 2,\ \ldots,\ n.$

Then Eq. (9.3) reduces to

$$f(x^* + h) - f(x^*) = \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j \frac{\partial^2 f(x)}{\partial x_i \partial x_j}\bigg|_{x = x^* + \theta h}.$$

$\therefore$ The sign of $f(x^* + h) - f(x^*)$

will be the same as that of

$$\sum_i \sum_j h_i h_j \frac{\partial^2 f(x)}{\partial x_i \partial x_j}\bigg|_{x=x^* + \theta h}.$$

Since the second-order partial derivative $\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$ is continuous in the neighbourhood of $x^*$, then $\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\Big|_{x=x^* + \theta h}$ will have the same sign as $\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\Big|_{x=x^*}$ for sufficiently small components of $h$.

This implies that $f(x^* + h) - f(x^*)$ will be positive and $x^*$ will be a local minimum if $Q = \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \frac{\partial^2 f(x)}{\partial x_i \partial x_j}\Big|_{x=x^*}$ is positive.

This $Q$ is a quadratic form and can be written as

$$Q = \langle Hh, h \rangle |_{x=x^*},$$

where $H|_{x=x^*}$ is called a Hessian matrix of $f(x)$.

From linear algebra, we know that the quadratic form will be positive for all $h$ if and only if $H$ is positive definite at $x = x^*$. This means that a sufficient condition for the extreme point $x^*$ to be a local minimum is that the Hessian matrix of $f(x)$ evaluated at $x = x^*$ is positive definite.

This completes the proof for the minimization case.

By proceeding in a similar manner, it can be proved that the Hessian matrix will be negative definite if $x^*$ is a local maximum point.

*Remark* For the semi-definite case of the Hessian matrix of $f(x)$, no general conclusion can be drawn regarding the optimality of $f(x)$.

*Example 5* Examine whether $f(x_1, x_2) = \frac{1}{2}k_2 x_1^2 + \frac{1}{2}k_3 (x_2 - x_1)^2 + \frac{1}{2}k_1 x_2^2 - Px_2$ ($P$ being constant) is optimum or not.

**Solution**

It is given that $f(x_1,\ x_2) = \frac{1}{2}k_2 x_1^2 + \frac{1}{2}k_3 (x_2 - x_1)^2 + \frac{1}{2}k_1 x_2^2 - Px_2$.

Now for the extremum of $f$, we have

$$\frac{\partial f}{\partial x_1} = 0 \quad \text{i.e.,}\ k_2 x_1 - k_3 (x_2 - x_1) = 0 \tag{9.4}$$

$$\text{and } \frac{\partial f}{\partial x_2} = 0 \quad \text{i.e., } k_3(x_2 - x_1) + k_1 x_2 - P = 0. \tag{9.5}$$

From (9.4), we have

$$\begin{aligned} k_2 x_1 - k_3 x_2 + k_3 x_1 &= 0 \\ \text{or } (k_2 + k_3)x_1 - k_3 x_2 &= 0. \end{aligned} \tag{9.6}$$

From (9.5), we have

$$\begin{aligned} k_3(x_2 - x_1) + k_1 x_2 - P &= 0 \\ \text{or } k_3 x_2 - k_3 x_1 + k_1 x_2 &= P \\ \text{or } -k_3 x_1 + (k_1 + k_3)x_2 &= P. \end{aligned} \tag{9.7}$$

Now multiplying (9.6) and (9.7) by $k_3$ and ($k_2$ + $k_3$) respectively and then adding, we have

$$\begin{aligned} &\left\{-k_3^2 + (k_1 + k_3)(k_2 + k_3)(k_2 + k_3)\right\}x_2 = P(k_2 + k_3) \\ \text{or } &x_2 = \frac{P(k_2 + k_3)}{k_1 k_2 + k_1 k_3 + k_2 k_3}. \end{aligned}$$

Now multiplying (9.6) and (9.7) by ($k_1$ + $k_3$) and $k_3$ respectively and then adding, we have

$$\begin{aligned} \left(k_1 k_2 + k_1 k_3 + k_2 k_3 + k_3^2\right)x_1 &= P k_3 \\ \text{or } x_1 &= \frac{P k_3}{k_1 k_2 + k_1 k_3 + k_2 k_3}. \end{aligned}$$

Now the Hessian matrix of $f$ evaluated at $(x_1, x_2)$ is

$$H|_{(x_1,x_2)} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}_{(x_1,x_2)} = \begin{bmatrix} k_2 + k_3 & -k_3 \\ -k_3 & k_1 + k_3 \end{bmatrix}_{(x_1,x_2)} = \begin{bmatrix} k_2 + k_3 & -k_3 \\ -k_3 & k_1 + k_3 \end{bmatrix}.$$

The leading principal minors of the Hessian matrix are

$$\begin{aligned} & A_1 = |k_2 + k_3| = k_2 + k_3 > 0 \quad [\because k_2, k_3 > 0] \\ \text{and} \quad & A_2 = \begin{vmatrix} k_2 + k_3 & -k_3 \\ -k_3 & k_1 + k_3 \end{vmatrix} = (k_2 + k_3)(k_1 + k_3) - k_3^2 \\ & \qquad = k_1 k_2 + k_2 k_3 + k_3 k_1 + k_3^2 - k_3^2 \\ & \qquad = k_1 k_2 + k_2 k_3 + k_3 k_1 > 0 \quad [\because k_1, k_2, k > 0]. \end{aligned}$$

Therefore, the Hessian matrix of $f$ evaluated at $(x_1, x_2)$, i.e. $H|_{(x_1,x_2)}$, is positive definite.

Hence, $f$ is minimum at $(x_1, x_2)$, which are given by

$$\begin{aligned} x_1 &= \frac{Pk_3}{k_1k_2 + k_2k_3 + k_2k_1} \\ x_2 &= \frac{P(k_2 + k_3)}{k_1k_2 + k_2k_3 + k_3k_1}. \end{aligned}$$

*Example 6* Find the extreme point of $f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$.

**Solution**

$$\begin{aligned} \text{Here} \quad & f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6 \\ & \therefore \tfrac{\partial f}{\partial x_1} = 3x_1^2 + 4x_1 = x_1(3x_1 + 4) \\ \text{and} \quad & \tfrac{\partial f}{\partial x_2} = 3x_2^2 + 8x_2 = x_2(3x_2 + 8). \end{aligned}$$

The necessary conditions for the optimum of $f(x_1, x_2)$ are given by

$$\frac{\partial f}{\partial x_1} = 0 \quad \text{and} \quad \frac{\partial f}{\partial x_2} = 0,$$

$$\text{i.e. } x_1(3x_1 + 4) = 0 \tag{9.8}$$

$$\text{and } x_2(3x_2 + 8) = 0. \tag{9.9}$$

From (9.8), we have $x_1 = 0$ or $x_1 = -\frac{4}{3}$.

From (9.9), we have $x_2 = 0$ or $x_2 = -\frac{8}{3}$.

Hence, Eqs. (9.8) and (9.9) are satisfied at the points [or the solution of (9.8) and (9.9)]

$$(0, 0), \left(0, -\frac{8}{3}\right), \left(-\frac{4}{3}, 0\right) \text{ and } \left(-\frac{4}{3}, -\frac{8}{3}\right).$$

Since $f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$,

$$\frac{\partial^2 f}{\partial x_1^2} = 6x_1 + 4, \quad \frac{\partial^2 f}{\partial x_2^2} = 6x_2 + 8, \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0.$$

The Hessian matrix of $f(x_1, x_2)$ is given by

$$H = \begin{bmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{bmatrix}.$$

Now, the leading principal minors of the Hessian matrix are given by

$$H_1 = |6x_1 + 4| \text{ and } H_2 = \begin{vmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{vmatrix}.$$

Now, we construct the following table:

| Point $(x_1, x_2)$ | Value of $H_1$ | Value of $H_2$ | Nature of $H$ | Nature of $x$ | Value of $f(x)$ |
|---|---|---|---|---|---|
| (0, 0) | 4 | 32 | Positive definite | Local minimum | 6 |
| $\left(0, -\frac{8}{3}\right)$ | 4 | −32 | Indefinite | Saddle point | $\frac{418}{27}$ |
| $\left(-\frac{4}{3}, 0\right)$ | −4 | −32 | Indefinite | Saddle point | $\frac{194}{27}$ |
| $\left(-\frac{4}{3}, -\frac{8}{3}\right)$ | −4 | 32 | Negative | Local maximum | $\frac{50}{3}$ |

*Example 7* Examine the extreme values of

$$f(x, y, z) = 2xyz - 4zx - 2yz + x^2 + y^2 + z^2 - 2x - 4y + 4z.$$

**Solution**

$$\begin{aligned} \text{Here, } f_x &= 2yz - 4z + 2x - 2 \\ f_y &= 2xz - 2z + 2y - 4 \\ f_z &= 2xy - 4x - 2y + 2z + 4. \end{aligned}$$

For extreme values of $f(x, y, z)$, the necessary conditions are

$$f_x = 0, \quad f_y = 0, \quad f_z = 0,$$

$$\text{i.e.} \quad yz - 2z + x - 1 = 0 \tag{9.10}$$

$$xz - z + y - 2 = 0 \tag{9.11}$$

$$xy - 2x - y + z + 2 = 0. \tag{9.12}$$

From (9.10), we have

$$x = 1 + 2z - yz. \tag{9.13}$$

Substituting $x = 1 + 2z - yz$ in (9.11) and then simplifying, we have

$$(2 - y)z^2 + y - 2 = 0. \tag{9.14}$$

Again, substituting $x = 1 + 2z - yz$ in (9.11) and then simplifying, we have

$$z(y^2 - 4y + 3) = 0,$$
$$\text{i.e. } z = 0 \text{ or } y = 3, 1.$$

When $z = 0$, then from (9.13) and (9.14), we have $x = 1$, $y = 2$.

Again, for $y = 2$, $z = \pm 1$, $x = 0$, 2 and for $y = 1$, $z = \pm 1$, $x = 2$, 0.

Therefore, the solutions of (9.10) through (9.12) are

$$(1, 2, 0), (0, 3, 1), (2, 3, -1), (2, 1, 1), (0, 1, -1).$$

Now, $f_{xx} = 2, \quad f_{yy} = 2, \quad f_{zz} = 2$

$$f_{yx} = f_{xy} = 2z, \quad f_{yz} = f_{zy} = 2x - 2, \quad f_{zx} = f_{xz} = 2y - 4.$$

$\therefore$ The Hessian matrix of $f(x, y, z)$ is

$$H = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix} = \begin{bmatrix} 2 & 2z & 2y-4 \\ 2z & 2 & 2x-2 \\ 2y-4 & 2x-2 & 2 \end{bmatrix}.$$

Now at (1, 2, 0),

$$H = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

The leading principal minors of $H$ are

$$H_1 = |2| = 2, H_2 = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4 \text{ and } H_3 = |H| = 8.$$

As all the leading principal minors are positive, the Hessian matrix at (1, 2, 0) is positive definite.

Hence, the given function $f(x,y,z)$ is minimum at (1, 2, 0), and the minimum value of $f(x,y,z)$ is −5.

In a similar way, test the optimality at other points.

*Example 8* Show that $f(x,y,z) = (x+y+z)^3 - 3(x+y+z) - 24xyz + a^3$ has a minimum at (1, 1, 1) and a maximum at (−1, −1, −1).

**Solution**

$$\begin{aligned}
\text{Here} \quad & f(x,y,z) = (x+y+z)^3 - 3(x+y+z) - 24xyz + a^3 \\
\text{Hence,} \quad & f_x = 3(x+y+z)^2 - 3 - 24yz \\
& f_y = 3(x+y+z)^2 - 3 - 24zx \\
& f_z = 3(x+y+z)^2 - 3 - 24xy \\
\text{Again,} \quad & f_{xx} = 6(x+y+z) \\
& f_{yy} = 6(x+y+z) \\
& f_{zz} = 6(x+y+z) \\
& f_{xy} = f_{yx} = 6(x+y+z) - 24z \\
& f_{yz} = f_{zy} = 6(x+y+z) - 24x \\
& f_{xz} = f_{zx} = 6(x+y+z) - 24y.
\end{aligned}$$

At (1, 1, 1),

$$f_x = 3.3^2 - 3 - 24.1.1 = 0$$
$$f_y = 3.3^2 - 3 - 24.1.1 = 0$$
$$f_z = 3.3^2 - 3 - 24.1.1 = 0.$$

Again at (−1, −1, −1),

$$f_x = 3(-3)^2 - 3 - 24 = 0$$
$$f_y = 3(-3)^2 - 3 - 24 = 0$$
$$f_z = 3(-3)^2 - 3 - 24 = 0.$$

Hence, (1, 1, 1) and (−1, −1, −1) are the extreme points of $f(x, y, z)$.

The Hessian matrix for the function $f(x, y, z)$ is

$$\begin{aligned}
H &= \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix} \\
&= \begin{bmatrix} 6(x+y+z) & 6(x+y+z) - 24z & 6(x+y+z) - 24y \\ 6(x+y+z) - 24z & 6(x+y+z) & 6(x+y+z) - 24x \\ 6(x+y+z) - 24y & 6(x+y+z) - 24x & 6(x+y+z) \end{bmatrix}.
\end{aligned}$$

At (1, 1, 1),

$$H = \begin{bmatrix} 18 & -6 & -6 \\ -6 & 18 & -6 \\ -6 & -6 & 18 \end{bmatrix}.$$

The leading principal minors are

$$H_1 = |18| = 18, H_2 = \begin{vmatrix} 18 & -6 \\ -6 & 18 \end{vmatrix} = 288 \text{ and } H_3 = |H| = 18 \times 288 = 5184.$$

As all the leading principal minors are positive, the Hessian matrix $H$ is positive definite at (1, 1, 1), Therefore, $f(x, y, z)$ has a minimum value at $(1, 1, 1)$, and the minimum value of $f(x, y, z)$ is $a^{3-6}$.

At $(-1, -1, -1)$,

$$H = \begin{vmatrix} -18 & 6 & 6 \\ 6 & 18 & 6 \\ 6 & 6 & -18 \end{vmatrix}.$$

The leading principal minors are

$$H_1 = |-18| = -18, H_2 = \begin{vmatrix} 18 & -6 \\ -6 & 18 \end{vmatrix} = (18)^2 - 6^2 = 288$$

and

$$\begin{aligned} H_3 = |H| &= -18(324 - 36) - 6(-108 - 36) + 6(36 + 108) = -5184 + 2 \times 864 \\ &= -3456. \end{aligned}$$

Hence, $f(x, y, z)$ has a maximum value at $(-1, -1, -1)$ [since the Hessian matrix is negative definite].

## 9.4 Exercises

1. Show that $(x-a)^{\frac{1}{3}}(2x-a)^{\frac{2}{3}}$ is maximum for $x = \frac{a}{2}$ and neither maximum nor minimum for $x = a$ $(a > 0)$.

2. Find the extreme points of the function $f(x_1, x_2) = 2x_1^3 + 3x_2^3 + 5x_1^2 + 4x_2^2 + 6$.
3. Find the extreme points of the function $f(x_1, x_2) = 8x_1 + 4x_2 + x_1x_2 - x_1^2 - x_2^2$.
4. Examine the extreme values of

$$f(x, y, z) = 2xyz - 4zx - 2yz + x^2 + y^2 + z^2 - 2x - 4y + 4z.$$

# Chapter 10
# Constrained Optimization with Equality Constraints

## 10.1 Objective

The objective of this chapter is to introduce the most important methods, namely direct substitution, constrained variation and the Lagrange multiplier method, for solving non-linear constrained optimization problems with equality constraints.

## 10.2 Introduction

The general form of non-linear constrained optimization problems with equality constraints can be written as follows:

$$\begin{aligned}&\text{Optimize } z = f(x)\\&\text{subject to } g_i(x) = 0, \quad i = 1,\ 2,\ \ldots,\ m,\\&\text{where } x = (x_1,\ x_2,\ \ldots,\ x_n).\end{aligned}$$

subject to $g_i(x) = 0, \qquad i = 1, 2, \ldots, m$
$x \in \mathbb{R}^n$, $f\colon \mathbb{R}^n \to \mathbb{R}$, $g_i\colon \mathbb{R}^n \to \mathbb{R}$ and $m \le n$.

If $m > n$, the problem becomes overdefined, and in general there will be no solution.

For solving this problem, there are several methods, e.g.

(i) Direct substitution method
(ii) Constrained variation method
(iii) Lagrange multiplier method.

We shall discuss these methods in detail.

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_10

## 10.3 Direct Substitution Method

For an optimization problem with $n$ variables and $m$ equality constraints, it is possible (theoretically) to express any set of $m$ variables in terms of the remaining $(n-m)$ variables. When these expressions are substituted into the original objective function, then the reduced objective function involves only $n-m$ variables. This reduced objective function is not subjected to any constraint, so its optimum value can be found by using the unconstrained optimization technique discussed in Chap. 9.

Theoretically, the method of direct substitution is very simple. However, from a practical point of view, it is not convenient. In most of practical problems, the constraint equations are highly non-linear in nature. In those cases, it becomes impossible to solve them and express any $m$ variables in terms of the remaining $(n-m)$ variables from the given constraints.

*Example 1*

$$\begin{aligned} \text{Minimize} \quad & z = 9 - 8x_1 - 6x_2 - 4x_3 + 2x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 \\ \text{subject to} \quad & x_1 + x_2 + 2x_3 = 3. \end{aligned}$$

**Solution** From the given constraint, we have $x_2 = 3 - x_1 - 2x_3$, and substituting it in the objective function, we have

$$\begin{aligned} z &= 9 - 8x_1 - 6(3 - x_1 - 2x_3) - 4x_3 + 2x_1^2 + 2(3 - x_1 - 2x_3)^2 + x_3^2 + 2x_1(3 - x_1 - 2x_3) + 2x_1x_3 \\ &= 2x_1^2 + 9x_3^2 + 6x_1x_3 - 8x_1 - 16x_3 + 9. \end{aligned}$$

Now we have to optimize $z$ with respect to the decision variables $x_1$ and $x_3$.

The necessary conditions for optimality of $z$ are given by

$$\frac{\partial z}{\partial x_1} = 0 \text{ and } \frac{\partial z}{\partial x_3} = 0,$$

$$\text{i.e. } 2x_1 + 3x_3 = 4 \tag{10.1}$$

$$3x_1 + 9x_3 = 8. \tag{10.2}$$

Solving (10.1) and (10.2), we have $x_1 = \frac{4}{3}, x_3 = \frac{4}{9}$.

Now $x_2 = 3 - x_1 - 2x_3 = 3 - \frac{4}{3} - 2\frac{4}{9} = \frac{7}{9}$.

We have $\frac{\partial^2 z}{\partial x_1^2} = 4, \quad \frac{\partial^2 z}{\partial x_3^2} = 18, \quad \frac{\partial^2 z}{\partial x_1 \partial x_3} = 6.$

Hence, the Hessian matrix is given by

$$H = \begin{pmatrix} \frac{\partial^2 z}{\partial x_1^2} & \frac{\partial^2 z}{\partial x_1\,\partial x_3} \\ \frac{\partial^2 z}{\partial x_3\,\partial x_1} & \frac{\partial^2 z}{\partial x_3^2} \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 6 & 18 \end{pmatrix}.$$

The leading principal minors are $H_1 = |4| = 4$ and $H_2 = \begin{vmatrix} 4 & 6 \\ 6 & 18 \end{vmatrix} = 36.$

Since $H_1 > 0$ and $H_2 > 0$, $Z$ is minimum for $x_1 = \frac{4}{3}$, $x_2 = \frac{7}{9}$, $x_3 = \frac{4}{9}$ and the minimum value of $z$ is $\frac{1}{9}$.

*Example 2* Find the dimension of a box of largest volume that can be inscribed in a sphere of unit radius.

**Solution**

Let the origin $O$ of the Cartesian coordinate system $OX_1X_2X_3$ be the centre of the sphere and the sides of the box be $2x_1$, $2x_2$ and $2x_3$. Obviously, $x_1,\ x_2,\ x_3 > 0$.

The volume of the box is given by

$$f(x_1,\ x_2,\ x_3) = 8x_1x_2x_3.$$

Since the box is inscribed in a sphere of unit radius, the corners of the box lie on the surface of the sphere of unit radius; that is, $x_1$, $x_2$ and $x_3$ have to satisfy the constraint equation $x_1^2 + x_2^2 + x_3^2 = 1$, as the equation of a sphere with unit radius and centre at the origin is $x_1^2 + x_2^2 + x_3^2 = 1$.

Hence, our problem is

$$\text{Maximize} f(x_1,\ x_2,\ x_3) = 8x_1x_2x_3$$

subject to $x_1^2 + x_2^2 + x_3^2 = 1$ and $x_1,\ x_2,\ x_3 > 0$.

This problem has three variables and one equality constraint. Hence, any one variable can be eliminated using the constraint.

Now, from the constraint $x_1^2 + x_2^2 + x_3^2 = 1$, we have

$$x_3 = \left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}} \text{as}\, x_3 > 0$$

Substituting the expression of $x_3$ in $f(x_1,\ x_2,\ x_3)$, we get the reduced objective function as

$$f = 8x_1x_2\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}},$$

which can be maximized as an unconstrained function of two variables.

The necessary conditions for the optimality of $f$ are

$$\frac{\partial f}{\partial x_1} = 0 \text{ and } \frac{\partial f}{\partial x_2} = 0.$$

$$\text{Now, } \frac{\partial f}{\partial x_1} = 0 \text{ implies } 8x_2\left[\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}} - \frac{x_1^2}{\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}}}\right] = 0. \tag{10.3}$$

$$\text{Again, } \frac{\partial f}{\partial x_2} = 0 \text{ implies } 8x_1\left[\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}} - \frac{x_2^2}{\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}}}\right] = 0. \tag{10.4}$$

After simplifying (10.3) and (10.4), we have

$$\begin{aligned} &1 - 2x_1^2 - x_2^2 = 0 \\ &1 - x_1^2 - 2x_2^2 = 0 \quad \text{as } x_1, x_2 > 0. \end{aligned}$$

Solving these equations, we have

$$x_1 = x_2 = \frac{1}{\sqrt{3}}.$$

Hence, $x_3 = \left(1 - x_1^2 - x_2^2\right)^{\frac{1}{3}} = \dfrac{1}{\sqrt{3}}$.

$$\text{Now, } \frac{\partial^2 f}{\partial x_1^2} = -\frac{8x_1x_2}{\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}}} - \frac{8x_2}{1 - x_1^2 - x_2^2}\left[\frac{x_1^3}{\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}}} + 2x_1\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}}\right]$$

$$\frac{\partial^2 f}{\partial x_2^2} = -\frac{8x_1x_2}{\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}}} - \frac{8x_1}{1 - x_1^2 - x_2^2}\left[\frac{x_2^3}{\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}}} + 2x_2\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}}\right]$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = 8\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}} - \frac{8x_2^2}{\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}}} - \frac{8x_1^2}{\left(1 - x_1^2 - x_2^2\right)}\left[\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}} + \frac{x_2^2}{\left(1 - x_1^2 - x_2^2\right)^{\frac{1}{2}}}\right].$$

At $x_1 = \frac{1}{\sqrt{3}}$, $\quad x_2 = \frac{1}{\sqrt{3}}$, $\quad \frac{\partial^2 f}{\partial x_1^2} = -\frac{32}{\sqrt{3}}$, $\quad \frac{\partial^2 f}{\partial x_2^2} = -\frac{32}{\sqrt{3}}$, $\quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = -\frac{16}{\sqrt{3}}$.

Hence, at $x_1 = x_2 = \frac{1}{\sqrt{3}}$, the Hessian matrix $H$ of $f$ is given by

$$H = \begin{bmatrix} -\frac{32}{\sqrt{3}} & -\frac{16}{\sqrt{3}} \\ -\frac{16}{\sqrt{3}} & -\frac{32}{\sqrt{3}} \end{bmatrix}.$$

The leading principal minors are

$$H_1 = \left| -\frac{32}{\sqrt{3}} \right| = -\frac{32}{\sqrt{3}} < 0$$

$$\text{and } H_2 = \begin{vmatrix} -\frac{32}{\sqrt{3}} & -\frac{16}{\sqrt{3}} \\ -\frac{16}{\sqrt{3}} & -\frac{32}{\sqrt{3}} \end{vmatrix} = \frac{1024}{9} - \frac{256}{9} = \frac{768}{9} > 0.$$

Hence, $f$ is maximum at $x_1 = x_2 = \frac{1}{\sqrt{3}}$, and the maximum value of $f$ is $8\left(\frac{1}{\sqrt{3}}\right)^3 = \frac{8}{3\sqrt{3}}$.

Hence, the maximum volume of the box is $\frac{8}{3\sqrt{3}}$ with dimensions $\frac{2}{\sqrt{3}}, \frac{2}{\sqrt{3}}, \frac{2}{\sqrt{3}}$.

## 10.4 Method of Constrained Variation

The basic idea used in the method of constrained variation is to find a closed-form expression for the first-order differential of $f(x)$ (i.e. d$f$) at all points at which the constraints $g_j(x) = 0, \quad j = 1, 2, \ldots, m$, are satisfied. Then, by setting d$f$ = 0, the desired optimum points are obtained. Before discussing the general method, we shall indicate the important features through the following simple problem with $n$ = 2 and $m$ = 1.

$$\begin{aligned} &\text{Optimize} \quad f(x_1, x_2) \\ &\text{subject to} \quad g(x_1, x_2) = 0. \end{aligned}$$

**Theorem 1** *The necessary condition for the preceding problem to have an extremum at* $\left(x_1^*, x_2^*\right)$ *is*

$$\left(\frac{\partial f}{\partial x_1}\frac{\partial g}{\partial x_2} - \frac{\partial f}{\partial x_2}\frac{\partial g}{\partial x_1}\right)\Bigg|_{\left(x_1^*, x_2^*\right)} = 0.$$

*Proof* The necessary condition for $f$ to have an extremum at some point $\left(x_1^*, x_2^*\right)$ is that the total differential of $f(x_1, x_2)$ must be zero at $\left(x_1^*, x_2^*\right)$,

$$\text{i.e. } \mathrm{d}f = 0$$
$$\text{or } \frac{\partial f}{\partial x_1}\mathrm{d}x_1 + \frac{\partial f}{\partial x_2}\mathrm{d}x_2 = 0 \tag{10.5}$$

Since $g\left(x_1^*, x_2^*\right) = 0$ at the extreme point, any infinitesimal variations $\mathrm{d}x_1$ and $\mathrm{d}x_2$ taken about the point $\left(x_1^*, x_2^*\right)$ are called the admissible variations, provided that the new point lies on the constraint,

$$\text{i.e. } g\left(x_1^* + \mathrm{d}x_1, x_2^* + \mathrm{d}x_2\right) = 0. \tag{10.6}$$

The Taylor series expansion of the function $g\left(x_1^* + \mathrm{d}x_1,\ x_2^* + \mathrm{d}x_2\right)$ about the point $\left(x_1^*,\ x_2^*\right)$ gives

$$g\left(x_1^*, x_2^*\right) + \left.\frac{\partial g}{\partial x_1}\right|_{\left(x_1^*, x_2^*\right)}\mathrm{d}x_1 + \left.\frac{\partial g}{\partial x_2}\right|_{\left(x_1^*, x_2^*\right)}\mathrm{d}x_2 = 0. \tag{10.7}$$

Since $g\left(x_1^*, x_2^*\right) = 0$, Eq. (10.7) reduces to

$$\frac{\partial g}{\partial x_1}\mathrm{d}x_1 + \frac{\partial g}{\partial x_2}\mathrm{d}x_2 = 0 \quad \text{at } \left(x_1^*,\ x_2^*\right). \tag{10.8}$$

Thus, Eq. (10.8) has to be satisfied by all the admissible variations.

Now, assuming that $\frac{\partial g}{\partial x_2} \neq 0$, Eq. (10.8) can be rewritten as

$$\mathrm{d}x_2 = -\frac{\partial g/\partial x_1}{\partial g/\partial x_2}\mathrm{d}x_1 \quad \text{at } \left(x_1^*,\ x_2^*\right). \tag{10.9}$$

This relation indicates that once the variation in $x_1$ (i.e. $\mathrm{d}x_1$) is chosen arbitrarily, the variation in $x_2$ (i.e. $\mathrm{d}x_2$) is obtained automatically in order to have $\mathrm{d}x_1$ and $\mathrm{d}x_2$ as a set of admissible variations.

Now, by substituting the preceding expression for $\mathrm{d}x_2$ in (10.5), we have

$$\left.\left(\frac{\partial f}{\partial x_1} - \frac{\partial f}{\partial x_2}\frac{\partial g/\partial x_1}{\partial g/\partial x_2}\right)\right|_{\left(x_1^*, x_2^*\right)}\mathrm{d}x_1 = 0. \tag{10.10}$$

The expression on the left-hand side of Eq. (10.10) is called the constrained variation of $f$.

Since $\mathrm{d}x_1$ is chosen arbitrarily, the coefficient of $\mathrm{d}x_1$ in (10.10) must be zero,

$$\text{i.e.}\ \left.\left(\frac{\partial f}{\partial x_1} - \frac{\partial f}{\partial x_2}\frac{\partial g/\partial x_1}{\partial g/\partial x_2}\right)\right|_{(x_1^*, x_2^*)} = 0$$

$$\text{i.e.}\ \left.\left(\frac{\partial f}{\partial x_1}\frac{\partial g}{\partial x_2} - \frac{\partial f}{\partial x_2}\frac{\partial g}{\partial x_1}\right)\right|_{(x_1^*, x_2^*)} = 0, \tag{10.11}$$

which is the necessary condition for the given problem to have an extremum at $(x_1^*, x_2^*)$.

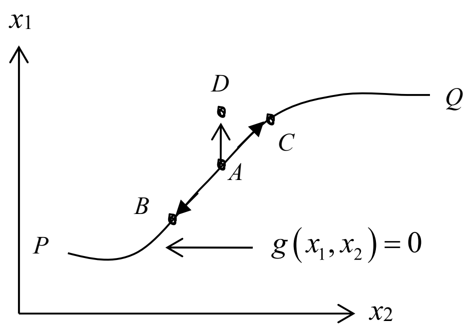
### 10.4.1 Geometrical Interpretation of Admissible Variation

We now illustrate the admissible variations geometrically in Fig. 10.1, where $PQ$ indicates the curve at each point of which $g(x_1,\ x_2) = 0$ is satisfied. If $A$ is taken as the point $(x_1^*,\ x_2^*)$, then the infinitesimal variations in $x_1$ and $x_2$ leading to the points $B$ and $C$ are called admissible variations. Thus, by the coordinates of $B$ and $C$, $g(x_1,\ x_2) = 0$ will be satisfied, as these points lie on the same curve. On the other hand, the infinitesimal variations in $x_1$ and $x_2$ representing $D$ are not admissible, since the point $D$ does not lie on the constraint curve $g(x_1,\ x_2) = 0$. Thus, any set of variations $(\mathrm{d}x_1,\ \mathrm{d}x_2)$ that does not satisfy Eq. (10.8) leads to points which do not satisfy the constraint $g(x_1,\ x_2) = 0$.

### 10.4.2 Necessary Conditions for Extremum of a General Problem for Constrained Variation Method

Let the function to be optimized be $f(x_1,\ x_2,\ \ldots,\ x_n)$ subject to the equality constraints

**Fig. 10.1** Variation about $A$



$$g_i(x_1, x_2, \ldots, x_n) = 0, \quad i = 1, 2, \ldots, m. \tag{10.12}$$

Let $x^* = \left(x_1^*,\ x_2^*,\ \ldots,\ x_n^*\right)$ be an extreme point and $(\mathrm{d}x_1,\ \mathrm{d}x_2,\ \ldots,\ \mathrm{d}x_n)$ be the infinitesimal admissible variations about the point $x^*$.

Then, $g_i\left(x_1^*,\ x_2^*,\ \ldots,\ x_n^*\right) = 0, \quad i = 1, 2, \ldots, m.$

Now, $g_i\left(x_1^* + \mathrm{d}x_1,\ \ldots,\ x_n^* + \mathrm{d}x_n\right) \simeq g_i\left(x_1^*,\ x_2^*,\ \ldots,\ x_n^*\right) + \sum_{j=1}^{n} \left.\frac{\partial g_i}{\partial x_j}\right|_{x^*} \mathrm{d}x_j.$

Since $(\mathrm{d}x_1,\ \mathrm{d}x_2,\ \ldots,\ \mathrm{d}x_n)$ are the admissible infinitesimal variations, then $g_i\left(x_1^* + \mathrm{d}x_1,\ \ldots,\ x_n^* + \mathrm{d}x_n\right) = 0.$

$$\begin{aligned} &\therefore g_i\left(x_1^*,\ x_2^*,\ \ldots,\ x_n^*\right) + \sum_{j=1}^{n} \left.\frac{\partial g_i}{\partial x_j}\right|_{x^*} \mathrm{d}x_j = 0 \\ &\text{or } \sum_{j=1}^{n} \left.\frac{\partial g_i}{\partial x_j}\right|_{x^*} \mathrm{d}x_j = 0, \quad \left[\because g_i\left(x_1^*,\ x_2^*,\ \ldots,\ x_n^*\right) = 0\right] \quad i = 1,\ 2,\ \ldots,\ m. \end{aligned} \tag{10.13}$$

There are $m$ equations in (10.13). These equations can be solved to express any $m$ variations, say, the first $m$ variations, in terms of the remaining variations as follows:

Writing $\left.\frac{\partial g_i}{\partial x_j}\right|_{x^*}$ simply as $\frac{\partial g_i}{\partial x_j}$, we get

$$\sum_{j=1}^{m} \frac{\partial g_i}{\partial x_j}\mathrm{d}x_j = -\sum_{k=m+1}^{n} \frac{\partial g_i}{\partial x_k}\mathrm{d}x_k = h_i \text{ (say)}, \quad i = 1,\ 2,\ \ldots,\ m.$$

By Cramer's rule, we have $\mathrm{d}x_1 = \Delta_1/\Delta$,

$$\text{where } \Delta_1 = \begin{vmatrix} h_1 & \frac{\partial g_1}{\partial x_2} & \frac{\partial g_1}{\partial x_3} & \cdots & \frac{\partial g_1}{\partial x_m} \\ h_2 & \frac{\partial g_2}{\partial x_2} & \frac{\partial g_2}{\partial x_3} & \cdots & \frac{\partial g_2}{\partial x_m} \\ \vdots & & & & \\ h_m & \frac{\partial g_m}{\partial x_2} & \frac{\partial g_m}{\partial x_3} & \cdots & \frac{\partial g_m}{\partial x_m} \end{vmatrix} \text{ and } \Delta = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_m} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_m} \\ \vdots & & & \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_m} \end{vmatrix}.$$

$$\text{Hence, } \mathrm{d}x_1 = \frac{\begin{vmatrix} -\sum\limits_{k=m+1}^{n} \frac{\partial g_1}{\partial x_k}\mathrm{d}x_k & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_m} \\ -\sum\limits_{k=m+1}^{n} \frac{\partial g_2}{\partial x_k}\mathrm{d}x_k & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_m} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ -\sum\limits_{k=m+1}^{n} \frac{\partial g_m}{\partial x_k}\mathrm{d}x_k & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_m} \end{vmatrix}}{J\left(\frac{g_1,\, g_2,\, \ldots,\, g_m}{x_1,\, x_2,\, \ldots,\, x_m}\right)} \left[\text{assuming } J\left(\frac{g_1,\ g_2,\ \ldots,\ g_m}{x_1,\ x_2,\ \ldots,\ x_m}\right) \neq 0\right]$$

$$= \frac{-\sum_{k=m+1}^{n} \mathrm{d}x_k J\left(\frac{g_1,\, g_2,\, \ldots,\, g_m}{x_k,\, x_2,\, \ldots,\, x_m}\right)}{J\left(\frac{g_1,\, g_2,\, \ldots,\, g_m}{x_1,\, x_2,\, \ldots,\, x_m}\right)}.$$

$$\therefore\ \mathrm{d}x_1 = \frac{-\sum_{k=m+1}^{n} \mathrm{d}x_k J\left(\frac{g_1,\, g_2,\, \ldots,\, g_m}{x_k,\, x_2,\, \ldots,\, x_m}\right)}{J\left(\frac{g_1,\, g_2,\, \ldots,\, g_m}{x_1,\, x_2,\, \ldots,\, x_m}\right)}.$$

$$\text{In general, } \mathrm{d}x_\alpha = \frac{-\sum_{k=m+1}^{n} \mathrm{d}x_k J\left(\frac{g_1,\, g_2,\, \ldots,\, g_{\alpha-1},\, g_\alpha,\, g_{\alpha+1},\, \ldots,\, g_m}{x_1,\, x_2,\, \ldots,\, x_{\alpha-1},\, x_k,\, x_{\alpha+1},\, \ldots,\, x_m}\right)}{J\left(\frac{g_1,\, g_2,\, \ldots,\, g_m}{x_1,\, x_2,\, \ldots,\, x_m}\right)}, \quad \alpha = 1,\ 2,\ \ldots,\ m.$$

At extreme point $x^*$, we have d$f = 0$

$$\text{or } \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n = 0 \quad \left[\text{Denoting } \left.\frac{\partial f}{\partial x_j}\right|_{x^*} \text{ by } \frac{\partial f}{\partial x_j}\right]$$

$$\text{or } \sum_{\alpha=1}^{m} \frac{\partial f}{\partial x_\alpha}dx_\alpha + \sum_{k=m+1}^{n} \frac{\partial f}{\partial x_k}dx_k = 0$$

Putting the expression d$x_\alpha$, $\alpha = 1, 2, \ldots, m$ in above, we have

$$\frac{1}{J\left(\frac{g_1,\, g_2,\, \ldots,\, g_n}{x_1,\, x_2,\, \ldots,\, x_n}\right)}\left[-\sum_{\alpha=1}^{m} \frac{\partial f}{\partial x_\alpha}\left\{\sum_{k=m+1}^{n} dx_k J\left(\frac{g_1,\ g_2,\ \ldots,\ g_{\alpha-1},\ g_\alpha,\ g_{\alpha+1},\ \ldots,\ g_m}{x_1,\ x_2,\ \ldots,\ x_{\alpha-1},\ x_k,\ x_{\alpha+1},\ \ldots,\ x_m}\right)\right\}\right.$$
$$\left. + \sum_{k=m+1}^{n} \frac{\partial f}{\partial x_k} J\left(\frac{g_1,\ g_2,\ \ldots,\ g_m}{x_1,\ x_2,\ \ldots,\ x_n}\right) dx_k\right] = 0$$

$$\text{or } \sum_{k=m+1}^{n}\left[-\sum_{\alpha=1}^{m} \frac{\partial f}{\partial x_\alpha} J\left(\frac{g_1,\ g_2,\ \ldots,\ g_{\alpha-1},\ g_\alpha,\ g_{\alpha+1},\ \ldots,\ g_m}{x_1,\ x_2,\ \ldots,\ x_{\alpha-1},\ x_k,\ x_{\alpha-1},\ \ldots,\ x_m}\right) + \frac{\partial f}{\partial x_k} J\left(\frac{g_1,\ g_2,\ \ldots,\ g_m}{x_1,\ x_2,\ \ldots,\ x_m}\right)\right] dx_k = 0.$$

Since d$x_{m+1}$, d$x_{m+2}$, $\ldots$, d$x_n$ are all independent variations, the coefficient of each variation must be zero. Thus, we have

$$-\sum_{\alpha=1}^{m}\frac{\partial f}{\partial x_\alpha}J\left(\frac{g_1,\ g_2,\ \ldots,\ g_{\alpha-1},\ g_\alpha,\ g_{\alpha+1},\ \ldots,\ g_m}{x_1,\ x_2,\ \ldots,\ x_{\alpha-1},\ x_k,\ x_{\alpha+1},\ \ldots,\ x_m}\right)+\frac{\partial f}{\partial x_k}J\left(\frac{g_1,\ g_2,\ \ldots,\ g_m}{x_1,\ x_2,\ \ldots,\ x_m}\right)=0$$

$$\text{for } k=m+1,\ m+2,\ \ldots,\ n$$

$$\text{or } -\frac{\partial f}{\partial x_1}J\left(\frac{g_1,\ g_2,\ \ldots,\ g_m}{x_k,\ x_2,\ \ldots,\ x_m}\right)-\frac{\partial f}{\partial x_2}J\left(\frac{g_1,\ g_2,\ \ldots,\ g_m}{x_1,\ x_k,\ \ldots,\ x_m}\right)\cdots$$
$$-\frac{\partial f}{\partial x_m}J\left(\frac{g_1,\ g_2,\ \ldots,\ g_{m-1},\ g_m}{x_1,\ x_2,\ \ldots,\ x_{m-1},\ x_k}\right)+\frac{\partial f}{\partial x_k}J\left(\frac{g_1,\ g_2,\ \ldots,\ g_m}{x_1,\ x_2,\ \ldots,\ x_m}\right)=0. \tag{10.14}$$

By transforming the column $\left[\frac{\partial g_1}{\partial x_k}\ \ \frac{\partial g_2}{\partial x_k}\ \ \frac{\partial g_3}{\partial x_k}\ \ldots\ \frac{\partial g_m}{\partial x_k}\right]^{\mathrm{T}}$ in each determinant of Eq. (10.14) to the first column, we get

$$\frac{\partial f}{\partial x_k}J\left(\frac{g_1,\ g_2,\ \ldots,\ g_m}{x_1,\ x_2,\ \ldots,\ x_m}\right)-\frac{\partial f}{\partial x_1}J\left(\frac{g_1,\ g_2,\ \cdots,\ g_m}{x_k,\ x_2,\ \ldots,\ x_m}\right)+\frac{\partial f}{\partial x_2}J\left(\frac{g_1,\ g_2,\ \ldots,\ g_m}{x_k,\ x_1,\ x_3,\ \ldots,\ x_m}\right)\cdots$$
$$+(-1)^m\frac{\partial f}{\partial x_m}J\left(\frac{g_1,\ g_2,\ \ldots,\ g_m}{x_k,\ x_1,\ \ldots,\ x_{m-1}}\right)=0,\quad k=m+1,\ m+2,\ \ldots,\ n$$

$$\text{or}\begin{vmatrix}\frac{\partial f}{\partial x_k} & \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_m}\\ \frac{\partial g_1}{\partial x_k} & \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_m}\\ \vdots & & & \\ \frac{\partial g_m}{\partial x_k} & \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_m}\end{vmatrix}=0 \tag{10.15}$$
$$\text{or } J\left(\frac{f,\ g_1,\ g_2,\ \ldots,\ g_m}{x_k,\ x_1,\ x_2,\ \ldots,\ x_m}\right)=0\quad \text{where } k=m+1,\ m+2,\ \ldots,\ n.$$

Thus, the $(n-m)$ equations represented by (10.15) give the necessary conditions for an extremum of $f(x_1,\ x_2,\ \ldots,\ x_n)$ under $m$ equality constraints $g_i(x_1,\ x_2,\ \ldots,\ x_n)=0,\ i=1,\ 2,\ \ldots,\ m.$

### 10.4.3 Sufficient Conditions for a General Problem in Constrained Variation Method

For the optimization problem with $n$ variables and $m$ equality constraints, $m$ variables are dependent and the other $n-m$ variables are independent. Let us assume that $x_1,\ x_2,\ \ldots,\ x_m$ are dependent variables and the others $x_{m+1},\ x_{m+2},\ \ldots,\ x_n$ are independent.

Then the Taylor series expansion of $f$, in terms of independent variables, about the extreme point $x^*$ gives

$$f\left(x^*+\mathrm{d}x\right) \simeq f\left(x^*\right)+\sum_{i=m+1}^{n}\left(\frac{\partial f}{\partial x_i}\right)_g \mathrm{d}x_i+\frac{1}{\underline{|2}}\sum_{i=m+1}^{n}\sum_{j=m+1}^{n}\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_g \mathrm{d}x_i \mathrm{d}x_j\,, \tag{10.16}$$

where $\left(\frac{\partial f}{\partial x_i}\right)_g$ denotes the partial derivative of $f$ with respect to $x_i$ (holding all the other variables $x_{m+1},\ x_{m+2},\ \ldots,\ x_{j-1},\ x_{j+1},\ \ldots,\ x_n$ as constant) when $x_1$, $x_2$, ..., $x_m$ are allowed to change so that the constraints $g_i(x^*+\mathrm{d}x)=0,\ i=1,\ 2,\ \ldots,\ m$ are satisfied. The second-order derivative $\left(\frac{\partial^2 f}{\partial x_j \partial x_k}\right)_g$ is used to denote a similar meaning.

Since all the variations $\mathrm{d}x_j$ appearing in (10.16) are independent, $\left(\frac{\partial f}{\partial x_j}\right)_g$ has to be zero for $j = m + 1, m + 2, \ldots, n$. Thus, the necessary conditions for the existence of a constrained optimum at $x^*$ can also be expressed as

$$\left(\frac{\partial f}{\partial x_j}\right)_g = 0, \quad j = m+1,\ m+2,\ \ldots,\ n. \tag{10.17}$$

Hence, from (10.16), we have

$$f\left(x^*+\mathrm{d}x\right)-f\left(x^*\right)=\frac{1}{\underline{|2}}\sum_{j=m+1}^{n}\sum_{k=m+1}^{n}\left(\frac{\partial^2 f}{\partial x_j \partial x_k}\right)_g \mathrm{d}x_j \mathrm{d}x_k\,.$$

Therefore, the sufficient condition for $x^*$ to be a local optimum is that the quadratic form $Q$ defined by

$$Q = \sum_{j=m+1}^{n}\sum_{k=m+1}^{n}\left(\frac{\partial^2 f}{\partial x_j \partial x_k}\right)_g \mathrm{d}x_j \mathrm{d}x_k$$

is positive (or negative) for all non-vanishing variations $\mathrm{d}x_j$. That is, the matrix

$$\begin{bmatrix} \left(\frac{\partial^2 f}{\partial x_{m+1}^2}\right)_g & \left(\frac{\partial^2 f}{\partial x_{m+1}\partial x_{m+2}}\right)_g & \cdots & \left(\frac{\partial^2 f}{\partial x_{m+1}\partial x_n}\right)_g \\ \vdots & & & \\ \left(\frac{\partial^2 f}{\partial x_n \partial x_{m+1}}\right)_g & \left(\frac{\partial^2 f}{\partial x_n \partial x_{m+2}}\right)_g & \cdots & \left(\frac{\partial^2 f}{\partial x_n^2}\right)_g \end{bmatrix}$$

is positive (or negative) definite.

*Remark* After a little manipulation, it can be shown that Eq. (10.17) and $J\left(\frac{f,\ g_1,\ g_2,\ \ldots,\ g_m}{x_k,\ x_1,\ x_2,\ \ldots,\ x_m}\right)=0$ are the same.

***Disadvantages of the method***

In this method, the computation of the constrained second-order partial derivative $\left(\frac{\partial^2 f}{\partial x_j \partial x_k}\right)_g$ is a formidable task for a problem with more than three constraints. Also, the necessary conditions involve the evaluation of $(n - m)$ numbers of determinants of order $m + 1$. This is also a complicated task.

*Example 3* Using the method of constrained variation,

$$\begin{aligned} \text{Minimize} \quad & f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + x_3^2 \\ \text{subject to} \quad & 2x_1 + 4x_2 + 3x_3 = 9 \\ & 4x_1 + 8x_2 + 5x_3 = 17. \end{aligned}$$

**Solution** The given problem is

$$f = x_1^2 + 2x_2^2 + x_3^2 \tag{10.18}$$

$$\text{subject to } g_1 = 2x_1 + 4x_2 + 3x_3 - 9 = 0 \tag{10.19}$$

$$g_2 = 4x_1 + 8x_2 + 5x_3 - 17 = 0. \tag{10.20}$$

$$\begin{aligned} \text{In this problem, } m &= \text{number of constraints} = 2 \\ n &= \text{number of variables} = 3. \end{aligned}$$

Hence $(n - m) = 1$ variable is independent and the other two are dependent. Let us choose $x_1, x_2$ as dependent variables.

Then, $J\left(\frac{g_1, g_2}{x_1, x_2}\right) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{vmatrix} = \begin{vmatrix} 2 & 4 \\ 4 & 8 \end{vmatrix} = 0.$

So $x_1$ and $x_2$ cannot be chosen as dependent variables.

Now, we consider $x_1$ and $x_3$ as dependent variables.

$$J\left(\frac{g_1, g_2}{x_1, x_3}\right) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_3} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_3} \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} = -2 \neq 0.$$

Thus, $x_1, x_3$ can be chosen as dependent variables.

The necessary condition for optimality of $f$ is given by

$$J\left(\frac{f, g_1, g_2}{x_2, x_1, x_3}\right) = 0,$$

$$\text{i.e.}\begin{vmatrix} \frac{\partial f}{\partial x_2} & \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_3} \\ \frac{\partial g_1}{\partial x_2} & \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_3} \\ \frac{\partial g_2}{\partial x_2} & \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_3} \end{vmatrix} = 0$$

$$\text{or,}\begin{vmatrix} 4x_2 & 2x_1 & 2x_3 \\ 4 & 2 & 3 \\ 8 & 4 & 5 \end{vmatrix} = 0 \quad \text{or,} \ -8x_2 + 8x_1 = 0 \quad \text{or, } x_1 = x_2. \tag{10.21}$$

Using (10.21) in (10.19) and (10.20), we have

$$\begin{aligned} 6x_2 + 3x_3 &= 9 \\ 12x_2 + 5x_3 &= 17. \end{aligned}$$

Solving, we get $x_2 = 1, \ x_3 = 1 \quad \therefore x_1 = x_2 = x_3 = 1.$

So the required optimum solution is $x_1 = x_2 = x_3 = 1$ with $f_{\min} = 6.$

*Example 4* Find at which point $f(x) = 2x_1^2 - x_2^2 + 4x_3^2 + x_4^2$ is optimum subject to

$$g_1(x) = x_1 + x_2 - x_3 + x_4 - 10 = 0 \tag{10.22}$$

$$g_2(x) = x_1 + x_2 + 2x_3 + 3x_4 - 12 = 0. \tag{10.23}$$

**Solution**

$$\begin{aligned} \text{Here } m &= \text{number of constraints} = 2 \\ n &= \text{number of variables} = 4. \end{aligned}$$

So $(n - m) = 2$ variables are independent and the other two are dependent. We first take $x_1$ and $x_2$ as dependent variables depending on $x_3, \ x_4$.

$$\text{Now, } J\left(\frac{g_1, \ g_2}{x_1, \ x_2}\right) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0.$$

So $x_1$, $x_2$ cannot be chosen as dependent variables.

$$\text{Again, } J\left(\frac{g_1, \ g_2}{x_1, \ x_3}\right) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_3} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_3} \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 1 & 2 \end{vmatrix} = 3 \neq 0.$$

So $x_1$, $x_3$ can be chosen as dependent variables.

Now the necessary conditions for optimality are given by

$$J\left(\frac{f,\ g_1,\ g_2}{x_2,\ x_1,\ x_3}\right) = 0 \text{ and } J\left(\frac{f,\ g_1,\ g_2}{x_4,\ x_1,\ x_3}\right) = 0$$

$$\text{i.e. } \begin{vmatrix} \frac{\partial f}{\partial x_2} & \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_3} \\ \frac{\partial g_1}{\partial x_2} & \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_3} \\ \frac{\partial g_2}{\partial x_2} & \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_3} \end{vmatrix} = 0 \text{ and } \begin{vmatrix} \frac{\partial f}{\partial x_4} & \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_3} \\ \frac{\partial g_1}{\partial x_4} & \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_3} \\ \frac{\partial g_2}{\partial x_4} & \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_3} \end{vmatrix} = 0$$

$$\text{i.e., } \begin{vmatrix} -2x_2 & 4x_1 & 8x_3 \\ 1 & 1 & -1 \\ 1 & 1 & 2 \end{vmatrix} = 0 \text{ and } \begin{vmatrix} 2x_4 & 4x_1 & 8x_3 \\ 1 & 1 & -1 \\ 3 & 1 & 2 \end{vmatrix} = 0.$$

$$\text{Hence, } -6x_2 - 12x_1 = 0 \text{ implies } x_2 = -2x_1 \tag{10.24}$$

$$\text{and } 6x_4 - 20x_1 - 16x_3 = 0 \text{ implies } 3x_4 - 10x_1 - 8x_3 = 0. \tag{10.25}$$

Now from (10.22) and (10.23) using (10.24), we have

$$-x_1 - x_3 + x_4 = 10 \tag{10.26}$$

$$\text{and } -x_1 + 2x_3 + 3x_4 = 12. \tag{10.27}$$

From (10.26) and (10.27), we have

$$3x_3 + 2x_4 = 2, \tag{10.28}$$

and from (10.25) and (10.26), we have

$$2x_3 - 7x_4 = -100. \tag{10.29}$$

Solving (10.28) and (10.29), we have $x_3 = -\frac{186}{25}$, $x_4 = \frac{304}{25}$.

$$\therefore x_1 = x_4 - x_3 - 10 = \frac{48}{5} \quad \therefore x_2 = -2x_1 = -\frac{96}{5}.$$

So at the point $x_1 = \frac{48}{5}$, $x_2 = -\frac{96}{5}$, $x_3 = -\frac{186}{25}$, $x_4 = \frac{304}{25}$, the given function $f(x)$ is optimum.

## 10.5 Lagrange Multiplier Method

In the area of optimization, the method of Lagrange multipliers (named after Joseph-Louis Lagrange) provides a strategy for finding the optimum of a function subject to the given equality constraints.

The Lagrange multiplier method is a very useful technique in multivariate calculus. One of the most common problems in calculus is finding the optimum of a function. However, it is difficult to find the closed-form solution for a problem to be optimized. Such difficulties often arise when one wishes to maximize or minimize a function subject to some equality constraints. The method of Lagrange multipliers is a very well-known method for solving non-linear constrained optimization problems. The method of Lagrange multipliers is a systematic process of generating the necessary conditions for a stationary point.

Let us consider the following optimization problem:

$$\begin{aligned}&\text{Optimize } z = f(x)\\ &\text{subject to } g_i(x) = 0, \quad i = 1,\ 2,\ \ldots,\ m\end{aligned}$$

where $x \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^m \to \mathbb{R}$ and $m \leq n$.

Now we introduce a new variable for each of the given constraints. These new variables $\lambda_1, \lambda_2, \ldots, \lambda_m$ are known as Lagrange multipliers. Using these multipliers, Lagrange defined a function now called the Lagrange function. It is denoted by $L(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m)$ and is defined by $L(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$ ($\lambda_i$ may be positive or negative).

Since the problem has $n$ variables and $m$ equality constraints, $m$ additional variables are to be added so that the Lagrange function will have $n + m$ variables.

### Stationary point

The points where the partial derivations of the Lagrange function $L$ are zero are called stationary points. If $(x_1, x_2, \ldots, x_n)$ is an optimum point for the original constrained optimization problem, then there exists a set of real numbers $(\lambda_1, \lambda_2, \ldots, \lambda_m)$ such that $(x_1, x_2, \ldots, x_n; \lambda_1, \lambda_2, \ldots, \lambda_m)$ is a stationary point for the Lagrange function. However, not all stationary points yield a solution of the original problem. Thus, the method of Lagrange multipliers yields necessary conditions for optimality in constrained optimization problems. Now we shall prove the theorem regarding the necessary conditions for optimality.

### 10.5.1 Necessary Conditions for Optimality

**Theorem 2** *The necessary condition for a function f(x) subject to the constraints* $g_i(x) = 0, \quad i = 1, 2, \ldots, m$ *to have a local optimum at a point* $x^*$ *is that the partial derivative of the Lagrange function defined by* $L = L(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$ *with respect to each of its arguments is zero at* $x^*$.

*Proof* Since $x^*$ is a local extreme point, then at $x^*$, $df = 0$

$$\text{or} \sum_{j=1}^{n} \left.\frac{\partial f}{\partial x_j}\right|_{x=x^*} dx_j = 0, \tag{10.30}$$

where $dx_j$ ($j = 1, 2, \ldots, n$) are the admissible variations about the point $x^*$.

Since $x^*$ is an extreme point and $dx_j$ ($j = 1, 2, \ldots, n$) are the admissible variations about the point $x^*$, then

$$g_i(x^*) = 0 \text{ and } g_i(x^* + dx) = 0. \tag{10.31}$$

From (10.31), we have

$$\begin{aligned} & g_i(x^* + dx) = 0 \\ & \text{or } g_i(x^*) + \sum_{j=1}^{n} \left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^*} dx_j = 0 \\ & \text{or} \sum_{j=1}^{n} \left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^*} dx_j = 0 \quad i = 1, 2, \ldots, m. \end{aligned} \tag{10.32}$$

Multiplying each equation in (10.32) by a constant $\lambda_i (i = 1, 2, \ldots, m)$, we have

$$\lambda_i \sum_{j=1}^{n} \left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^*} dx_j = 0 \quad (i = 1, 2, \ldots, m). \tag{10.33}$$

From (10.30) and (10.33), we have

$$\begin{aligned}&\sum_{j=1}^{n}\left.\frac{\partial f}{\partial x_j}\right|_{x=x^*}\mathrm{d}x_j+\sum_{i=1}^{m}\lambda_i\sum_{j=1}^{n}\left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^*}\mathrm{d}x_j=0\\ \text{or}\ &\sum_{j=1}^{n}\left.\frac{\partial f}{\partial x_j}\right|_{x=x^*}\mathrm{d}x_j+\sum_{j=1}^{n}\left(\sum_{i=1}^{m}\lambda_i\left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^*}\right)\mathrm{d}x_j=0\\ \text{or}\ &\sum_{j=1}^{n}\left[\left.\frac{\partial f}{\partial x_j}\right|_{x=x^*}+\sum_{i=1}^{m}\lambda_i\left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^*}\right]\mathrm{d}x_j=0.\end{aligned}\tag{10.34}$$

Since there are $m$ constraints, any $n-m$ variables can be chosen independently. Let $x_1,\ x_2,\ \ldots,\ x_m$ be a set of dependent variables and $x_{m+1},\ x_{m+2},\ \ldots,\ x_n$ be the set of independent variables.

We choose the values of $\lambda_i(i=1,\ 2,\ \ldots,\ m)$ in such a way that the coefficients of the first $m$ variations $\mathrm{d}x_i(i=1,\ 2,\ \ldots,\ m)$ become zero. Thus, $\lambda_i$ are defined by

$$\left.\frac{\partial f}{\partial x_j}\right|_{x=x^*}+\sum_{i=1}^{m}\lambda_i\left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^*}=0\quad \text{for}\, j=1,\ 2,\ \ldots,\ m.\tag{10.35}$$

Using (10.35), from (10.34) we have

$$\sum_{j=m+1}^{n}\left[\left.\frac{\partial f}{\partial x_j}\right|_{x=x^*}+\sum_{i=1}^{m}\lambda_i\left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^*}\right]\mathrm{d}x_j=0.$$

Since $\mathrm{d}x_{m+1},\ \mathrm{d}x_{m+2},\ \ldots,\ \mathrm{d}x_n$ are all independent, we have

$$\left.\frac{\partial f}{\partial x_j}\right|_{x=x^*}+\sum_{i=1}^{m}\lambda_i\left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^*}=0\quad \text{for}\, j=m+1,\ m+2,\ \ldots,\ n.\tag{10.36}$$

From (10.35) and (10.36), we have

$$\left.\frac{\partial f}{\partial x_j}\right|_{x=x^*}+\sum_{i=1}^{m}\lambda_i\left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^*}=0\quad \text{for}\, j=1,\ 2,\ \ldots,\ n.\tag{10.37}$$

$$\text{Also,}\ g_i(x^*)=0\quad \text{for}\, i=1,\ 2,\ \ldots,\ m.\tag{10.38}$$

Now $L = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$.

$$\therefore \left.\frac{\partial L}{\partial x_j}\right|_{x=x^*} = \left.\frac{\partial f}{\partial x_j}\right|_{x=x^*} + \sum_{i=1}^{m} \lambda_i \left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^*} = 0 \quad \text{for } j = 1,\ 2,\ \ldots,\ n \quad \text{[using (10.37)]}$$

and $\left.\dfrac{\partial L}{\partial \lambda_k}\right|_{x=x^*} = g_k(x^*) = 0 \quad \text{for } k = 1,\ 2,\ \ldots,\ m \quad$ [using (10.38)]

or $\left.\dfrac{\partial L}{\partial \lambda_k}\right|_{x=x^*} = 0 \quad \text{for } k = 1,\ 2,\ \ldots,\ m.$

These are the necessary conditions.

### *10.5.2 Sufficient Conditions for Optimality*

**Theorem 3** *A sufficient condition for f(x) to have a local minimum (maximum) subject to the constraints* $g_i(x) = 0, \quad i = 1,\ 2,\ \ldots,\ m$ *at* $x^*$ *is that the partial derivative of the Lagrange function* $L(x_1,\ x_2,\ \ldots,\ x_n;\quad \lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n)$ *at* $x^*$ *with respect to each of its arguments is zero, and the quadratic Q defined by*

$$Q = \sum_{j=1}^{n} \sum_{k=1}^{n} \left.\frac{\partial^2 L}{\partial x_j \partial x_k}\right|_{x=x^*} \mathrm{d}x_j \mathrm{d}x_k$$

must be positive (negative) for all choices of the admissible variations $\mathrm{d}x_j (j = 1,\ 2,\ \ldots,\ n)$.

*Proof* The Lagrange function is defined as

$$L(x,\ \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x).$$

Let d$x$ be the vector of admissible variations about the point $x^*$. Then the point $x^*$ + d$x$ must satisfy the given constraints.

$$\therefore g_i(x^* + \mathrm{d}x) = 0. \tag{10.39}$$

$$\text{Also, } \left.\frac{\partial L}{\partial \lambda_i}\right|_{x^*} = g_i(x^*) = 0. \tag{10.40}$$

$$\therefore L(x^* + \mathrm{d}x,\ \lambda) - L(x^*,\ \lambda) = f(x^* + \mathrm{d}x) + \sum_{i=1}^{m} \lambda_i g_i(x^* + \mathrm{d}x) - f(x^*) - \sum_{i=1}^{m} \lambda_i g_i(x^*)$$

$$\text{or } L(x^* + \mathrm{d}x,\ \lambda) - L(x^*,\ \lambda) = f(x^* + \mathrm{d}x) - f(x^*) \quad [\text{using (10.39) and (10.40)}]. \tag{10.41}$$

The Taylor series expansion of $L(x^* + \mathrm{d}x,\ \lambda)$ about the point $x^*$ (keeping all the components of $\lambda$ as fixed) gives

$$L(x^* + \mathrm{d}x,\ \lambda) = L(x^*,\ \lambda) + \sum_{j=1}^{n} \left.\frac{\partial L}{\partial x_j}\right|_{x=x^*} \mathrm{d}x_j + \sum_{j=1}^{n}\sum_{k=1}^{n} \left.\frac{\partial^2 L}{\partial x_j \partial x_k}\right|_{x=x^* + \theta \mathrm{d}x} \mathrm{d}x_j \mathrm{d}x_k, \quad 0<\theta<1$$

$$\text{or } L(x^* + \mathrm{d}x,\ \lambda) - L(x^*,\ \lambda) = 0 + \sum_{j=1}^{n}\sum_{k=1}^{n} \left.\frac{\partial^2 L}{\partial x_j \partial x_k}\right|_{x=x^* + \theta \mathrm{d}x} \mathrm{d}x_j \mathrm{d}x_k \quad \left[\because \left.\frac{\partial L}{\partial x_j}\right|_{x=x^*} = 0\right]. \tag{10.42}$$

If it is assumed that the second-order partial derivatives $\frac{\partial^2 L}{\partial x_i \partial x_j}$ are continuous in the neighbourhood of $x^*$, then the sign of $L(x^* + \mathrm{d}x,\ \lambda) - L(x^*,\ \lambda)$ will be the same as that of $\sum_{j=1}^{n}\sum_{k=1}^{n} \left.\frac{\partial^2 L}{\partial x_j \partial x_k}\right|_{x=x^*} \mathrm{d}x_j \mathrm{d}x_k$.

Therefore, from (10.41) and (10.42), it is clear that the signs of $f(x^* + \mathrm{d}x) - f(x^*)$ and $Q = \sum_{j=1}^{n}\sum_{k=1}^{n} \left.\frac{\partial^2 L}{\partial x_j \partial x_k}\right|_{x=x^*} \mathrm{d}x_j \mathrm{d}x_k$ be the same for all admissible variations.

If $Q > 0$ for all admissible variations, then $f(x^* + \mathrm{d}x) - f(x^*) > 0$ for all admissible variations.

That is, $f(x^* + \mathrm{d}x) > f(x^*)$.

$\therefore x^*$ is a local minimum.

If $Q < 0$ for all admissible variations, then

$$f(x^* + \mathrm{d}x) - f(x^*) < 0$$

$$\text{or } f(x^* + \mathrm{d}x) < f(x^*) \text{ for all admissible variations.}$$

$\therefore x^*$ is a local maximum.

### 10.5.3 Alternative Approach for Testing Optimality

To test the optimality of a constrained optimization problem by the above-described sufficient conditions is a formable task. Therefore, the following two methods are also used to test the optimality.

**Method-1**: In this method, a square matrix $H_{\mathrm{B}}$ of order $(m + n) \times (m + n)$, called a bordered Hessian matrix, is evaluated at the stationary points. This matrix is arranged as follows:

$$H_{\mathrm{B}} = \begin{bmatrix} O & U \\ U^{\mathrm{T}} & V \end{bmatrix},$$

where $O$ is an $m \times m$ null matrix, $U = \left[\frac{\partial g_i}{\partial x_j}\right]_{m\times n}, \quad V = \left[\frac{\partial^2 L}{\partial x_i \partial x_j}\right]_{n\times n}.$

Then the sufficient conditions for maximum and minimum stationary points are given as follows.

Let $(x^*, \lambda^*)$ be the stationary points for the Lagrangian function and $H_{\mathrm{B}}^*$ be the value of the corresponding bordered Hessian matrix computed at this stationary point. Then:

(i) $x^*$ is a local maximum if, starting with the leading principal minor of order $(2\,m + 1)$, the last $(n - m)$ leading principal minors of $H_{\mathrm{B}}^*$ have alternating sign starting with $(-1)^{m+n}$.

(ii) $x^*$ is a local minimum if, starting with the leading principal minor of order $(2m + 1)$, the last $(n - m)$ leading principal minors of $H_B^*$ have the sign of $(-1)^m$.

**Method-2**: In this method, optimality is tested by evaluating the roots of the polynomial in $z$, defined by the following Hancock determinantal equation:

$$\begin{vmatrix} L_{11} - z & L_{12} & L_{13} & \cdots & L_{1n} & g_{11} & g_{21} & \cdots & g_{m1} \\ L_{21} & L_{22} - z & L_{23} & \cdots & L_{2n} & g_{12} & g_{22} & \cdots & g_{m2} \\ \vdots & & & & & & & & \\ L_{n1} & L_{n2} & L_{n3} & \cdots & L_{nn} - z & g_{1n} & g_{2n} & \cdots & g_{mn} \\ g_{11} & g_{12} & g_{13} & \cdots & g_{1n} & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & & & \\ g_{m1} & g_{m2} & g_{n3} & \cdots & g_{mn} & 0 & 0 & \cdots & 0 \end{vmatrix} = 0$$

at $x = x^*$ and $\lambda = \lambda^*$,

where $L_{jk} = \frac{\partial^2 L}{\partial x_j \partial x_k}(x^*, \lambda^*), \quad g_{ij} = \frac{\partial g_i(x^*)}{\partial x_j}.$

This equation leads to an $(n - m)$th-order polynomial in $z$.

Then the sufficient conditions for minimum and maximum are, respectively:

(i) All positive roots of the polynomial
(ii) All negative roots of the polynomial.

If some of the roots of this polynomial are positive while the others are negative, the point $x^*$ is not an extreme point.

*Remark* It may be observed that the conditions in both the methods are only sufficient for identifying an extreme point but not necessary. That is, a stationary point may be an extreme point without satisfying the conditions described in Methods 1 and 2.

*Example 5* Solve the following problem by the Lagrange multiplier method:

$$\begin{array}{ll} \text{Minimize} & z = x_1^2 - 10x_1 + x_2^2 - 6x_2 + x_3^2 - 4x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 7. \end{array}$$

**Solution** The given problem is

$$\begin{array}{ll} \text{Minimize} & z = x_1^2 - 10x_1 + x_2^2 - 6x_2 + x_3^2 - 4x_3 \\ \text{subject to} & g(x) = x_1 + x_2 + x_3 - 7 = 0. \end{array}$$

Now the Lagrange function is

$$L(x_1, x_2, x_3, \lambda) = x_1^2 - 10x_1 + x_2^2 - 6x_2 + x_3^2 - 4x_3 + \lambda(x_1 + x_2 + x_3 - 7).$$

$\therefore$ The necessary conditions for the minimum of $z$ are given by

$$\frac{\partial L}{\partial x_1} = 0, \quad \frac{\partial L}{\partial x_2} = 0, \quad \frac{\partial L}{\partial x_3} = 0, \quad \frac{\partial L}{\partial \lambda} = 0.$$

$$\text{Now, } \frac{\partial L}{\partial x_1} = 0 \text{ gives } 2x_1 - 10 + \lambda = 0 \text{ or, } x_1 = \frac{10 - \lambda}{2} \tag{10.43}$$

$$\frac{\partial L}{\partial x_2} = 0 \text{ gives } 2x_2 - 6 + \lambda = 0 \quad \text{or,} \quad x_2 = \frac{6 - \lambda}{2} \tag{10.44}$$

$$\frac{\partial L}{\partial x_3} = 0 \text{ gives } 2x_3 - 4 + \lambda = 0 \quad \text{or,} \quad x_3 = \frac{4 - \lambda}{2} \tag{10.45}$$

$$\frac{\partial L}{\partial \lambda} = 0 \text{ gives } x_1 + x_2 + x_3 - 7 = 0 \text{ or, } x_1 + x_2 + x_3 = 7 \tag{10.46}$$

From (10.43), (10.44), (10.45) and (10.46), we have

$$\frac{20-3\lambda}{2}=7 \quad \text{or,} \quad \lambda=2.$$

$$\therefore x_1=4, \quad x_2=2, \quad x_3=1.$$

$$\text{Now } \frac{\partial^2 L}{\partial x_i^2}=2 \text{ for } \quad i=1,2,3$$

$$\frac{\partial^2 L}{\partial x_1 \partial x_j}=0 \quad \text{for} \quad j=2,3$$

$$\frac{\partial^2 L}{\partial x_2 \partial x_j}=0 \quad \text{for} \quad j=1,3$$

$$\frac{\partial^2 L}{\partial x_3 \partial x_j}=0 \quad \text{for} \quad j=1,2$$

$$\text{and } \frac{\partial g}{\partial x_i}=1 \quad \text{for} \quad i=1,2,3.$$

Hence, the bordered Hessian matrix at $x_1=4, x_2=2$ and $x_3=1$ is given by

$$H_B=\begin{bmatrix} O & U \\ U^T & V \end{bmatrix}, \text{ where O is a } 1\times 1 \text{ null matrix,}$$

$$U=\left[\frac{\partial g_i}{\partial x_j}\right]_{1\times 3} \quad \text{and} \quad V=\left[\frac{\partial^2 L}{\partial x_i \partial x_j}\right]_{3\times 3},$$

$$\text{i.e. } H_B=\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}.$$

Since $m = 1$ and $n = 3$, then $n - m = 3 - 1 = 2$ and $2m + 1 = 3$. Hence, only two leading principal minors of $H_B$ of order 3 and 4 (i.e. $\Delta_3$ and $\Delta_4$) are to be evaluated, and both the minors must have negative sign, as $(-1)^1$ = negative.

$$\text{Now, } \Delta_3=\begin{vmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{vmatrix}=-2-2=-4<0$$

$$\text{and } \Delta_4=\begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \end{vmatrix}=-\begin{vmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{vmatrix}+\begin{vmatrix} 1 & 2 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 2 \end{vmatrix}-\begin{vmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \\ 1 & 0 & 0 \end{vmatrix}$$
$$=-4+(-4)-4=-12<0$$

Hence, $z$ is minimum subject to $g(x) = 0$ at $x_1 = 4,\ x_2 = 2,\ x_3 = 1$ and $z_{\min} = 16 - 40 + 4 - 12 + 1 - 4 = -35$.

**Alternative test for optimality**
Again, the Hancock determinantal equation is

$$\begin{vmatrix} L_{11} - z & L_{12} & L_{13} & g_{11} \\ L_{21} & L_{22} - z & L_{23} & g_{12} \\ L_{31} & L_{32} & L_{23} - z & g_{13} \\ g_{11} & g_{12} & g_{13} & 0 \end{vmatrix} = 0,$$

where $L_{ij} = \left.\frac{\partial^2 L}{\partial x_i \, \partial x_j}\right|_{\substack{x = x^* \\ \lambda = \lambda^*}}$ and $g_{ij} = \left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^*}$

$$\text{or,} \quad \begin{vmatrix} 2 - z & 0 & 0 & 1 \\ 0 & 2 - z & 0 & 1 \\ 0 & 0 & 2 - z & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix} = 0$$

$$\text{or,} \quad (2 - z)\begin{vmatrix} 2 - z & 0 & 1 \\ 0 & 2 - z & 1 \\ 1 & 1 & 0 \end{vmatrix} - \begin{vmatrix} 0 & 2 - z & 0 \\ 0 & 0 & 2 - z \\ 1 & 1 & 1 \end{vmatrix} = 0$$

$$\text{or } (2 - z)\{(2 - z)(-1) - (2 - z)\} - (2 - z)^2 = 0$$

$$\text{or } (2 - z)(2 - z)(-2) - (2 - z)^2 = 0$$

$$\text{or } - 3(2 - z)^2 = 0$$

$$\text{or } z = 2, 2.$$

Hence, the extreme point (4, 2, 1) minimizes the objective function, and $z_{\min} = -35$.

*Example 6* Solve the following problem by the Lagrange multiplier method:

$$\text{Minimize } z = 4x_1^2 + 2x_2^2 + x_3^2 - 4x_1x_2$$
$$\text{subject to the constraints } x_1 + x_2 + x_3 = 15, 2x_1 - x_2 + 2x_3 = 20.$$

**Solution** The given problem is

$$\text{Minimize } z = f(x) = 4x_1^2 + 2x_2^2 + x_3^2 - 4x_1x_2$$
$$\text{subject to } g_1(x) = x_1 + x_2 + x_3 - 15 = 0$$
$$\text{and} \quad g_2(x) = 2x_1 - x_2 + 2x_3 - 20 = 0.$$

Now the Lagrange function is given by

$$\begin{aligned} L &= f(x) + \lambda_1 g_1(x) + \lambda_2 g_2(x) \\ &= 4x_1^2 + 2x_2^2 + x_3^2 - 4x_1x_2 + \lambda_1(x_1 + x_2 + x_3 - 15) + \lambda_2(2x_1 - x_2 + 2x_3 - 20). \end{aligned}$$

The necessary conditions for the minimum of $f(x)$ are given by

$$\frac{\partial L}{\partial x_1} = 0, \quad \frac{\partial L}{\partial x_2} = 0, \quad \frac{\partial L}{\partial x_3} = 0, \quad \frac{\partial L}{\partial \lambda_1} = 0, \quad \frac{\partial L}{\partial \lambda_2} = 0.$$

$$\text{Now, } \frac{\partial L}{\partial x_1} = 0 \text{ gives } 8x_1 - 4x_2 + \lambda_1 + 2\lambda_2 = 0 \tag{10.47}$$

$$\frac{\partial L}{\partial x_2} = 0 \text{ gives } 4x_2 - 4x_1 + \lambda_1 - \lambda_2 = 0 \tag{10.48}$$

$$\frac{\partial L}{\partial x_3} = 0 \text{ gives } 2x_3 + \lambda_1 + 2\lambda_2 = 0 \tag{10.49}$$

$$\frac{\partial L}{\partial \lambda_1} = 0 \text{ gives } x_1 + x_2 + x_3 - 15 = 0 \tag{10.50}$$

$$\frac{\partial L}{\partial \lambda_2} = 0 \text{ gives } 2x_1 - x_2 + 2x_3 - 20 = 0. \tag{10.51}$$

From (10.47) and (10.48), we have

$$4x_1 + 2\lambda_1 + \lambda_2 = 0$$

$$\text{or } x_1 = -\frac{2\lambda_1 + \lambda_2}{4}.$$

From (10.47), we have

$$4x_2 = -8\frac{2\lambda_1 + \lambda_2}{4} + \lambda_1 + 2\lambda_2 = -4\lambda_1 - 2\lambda_2 + \lambda_1 + 2\lambda_2 = -3\lambda_1$$

$$\text{or } x_2 = -\frac{3}{4}\lambda_1$$

From (10.49), we have

$$x_3 = -\frac{\lambda_1 + 2\lambda_2}{2}.$$

Now putting the values of $x_1, x_2, x_3$ in (10.50), we get

$$-\frac{2\lambda_1+\lambda_2}{4}-\frac{3}{4}\lambda_1-\frac{\lambda_1+2\lambda_2}{2}=15$$

$$\text{or } 7\lambda_1+5\lambda_2+60=0. \tag{10.52}$$

Putting the values of $x_1, x_2, x_3$ in (10.51), we get

$$\begin{aligned} &-2\cdot\frac{2\lambda_1+\lambda_2}{4}+\frac{3}{4}\lambda_1-2\frac{\lambda_1+2\lambda_2}{2}=20 \\ &\text{or } \lambda_1+2\lambda_2+16=0. \end{aligned} \tag{10.53}$$

Now solving (10.52) and (10.53) by the cross-multiplication method, we have

$$\begin{aligned} &\frac{\lambda_1}{80-120}=\frac{\lambda_2}{60-112}=\frac{1}{14-5} \\ \text{or}\quad &\lambda_1=-\frac{40}{9}\quad\text{and}\quad\lambda_2=-\frac{52}{9}. \\ &\therefore x_1=-\frac{2\lambda_1+\lambda_2}{4}=-\frac{1}{4}\left(-2\times\frac{40}{9}-\frac{52}{9}\right)=\frac{11}{3} \\ &x_2=-\frac{3}{4}\lambda_1=-\frac{3}{4}\left(-\frac{40}{9}\right)=\frac{10}{3} \\ \text{and}\quad &x_3=-\frac{\lambda_1+2\lambda_2}{2}=-\frac{1}{2}\left(-\frac{40}{9}-2\times\frac{52}{9}\right)=8. \\ \text{Now}\quad &\frac{\partial^2 L}{\partial x_1^2}=8\quad\frac{\partial^2 L}{\partial x_1\partial x_2}=-4\quad\frac{\partial^2 L}{\partial x_1\partial x_3}=0 \\ &\frac{\partial^2 L}{\partial x_2^2}=4\quad\frac{\partial^2 L}{\partial x_2\partial x_1}=-4\quad\frac{\partial^2 L}{\partial x_2\partial x_3}=0 \\ &\frac{\partial^2 L}{\partial x_3^2}=2\quad\frac{\partial^2 L}{\partial x_3\partial x_1}=0\quad\frac{\partial^2 L}{\partial x_3\partial x_2}=0. \\ \text{Again,}\quad &\frac{\partial g_1}{\partial x_1}=1\quad\frac{\partial g_1}{\partial x_2}=1\quad\frac{\partial g_1}{\partial x_3}=1\quad\frac{\partial g_2}{\partial x_1}=2\quad\frac{\partial g_2}{\partial x_2}=-1\quad\frac{\partial g_2}{\partial x_3}=2. \end{aligned}$$

Hence, the bordered Hessian matrix at $x_1=\frac{11}{3}$, $x_2=\frac{10}{3}$, $x_3=8$ is given by

$$H_B=\begin{bmatrix} O & U \\ U^T & V \end{bmatrix},$$

where O is an $m\times m$ null matrix,

$$U=\left[\frac{\partial g_i}{\partial x_j}\right]_{m\times n}\quad\text{and}\quad V=\left[\frac{\partial^2 L}{\partial x_i\partial x_j}\right]_{n\times n},$$

$$\text{i.e. } H_{\mathrm{B}} = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & -1 & 2 \\ 1 & 2 & 8 & -4 & 0 \\ 1 & -1 & -4 & 4 & 0 \\ 1 & 2 & 0 & 0 & 2 \end{bmatrix}.$$

Since $m = 2$ and $n = 3$, then $n - m = 1$ and $2m + 1 = 5$. Thus, only one leading principal minor of $H_{\mathrm{B}}$ of order 5 (i.e. $\Delta_5$) is to be evaluated, and the minor must have positive sign, as $(-1)^m =$ positive.

Here $\Delta_5 = |H_B| = 90 > 0$.

Hence, $f(x)$ is minimum subject to the constraints $g_1(x) = 0$ and $g_2(x) = 0$ at $x_1 = \frac{11}{3}$, $x_2 = \frac{10}{3}$ and $x_3 = 8$, and $f_{\min} = 4\left(\frac{11}{3}\right)^2 + 2\left(\frac{10}{3}\right)^2 + (8)^2 - 4 \times \frac{11}{3} \times \frac{10}{3} = \frac{820}{9}$.

*Example 7* Solve the following problem by using the Lagrange multiplier method:

$$\begin{aligned} &\text{Minimize } z = 2x_1^2 - 24x_1 + 2x_2^2 - 8x_2 + 2x_3^2 - 12x_3 + 200 \\ &\text{subject to the constraints } x_1 + x_2 + x_3 = 11 \text{ and } x_1, x_2, x_3 \geq 0. \end{aligned}$$

**Solution** The given problem can be written as

$$\begin{aligned} &\text{Minimize } z = 2x_1^2 - 24x_1 + 2x_2^2 - 8x_2 + 2x_3^2 - 12x_3 + 200 \\ &\text{subject to } g(x) = x_1 + x_2 + x_3 - 11 = 0. \end{aligned}$$

Now the Lagrange function is

$$\begin{aligned} L(x_1, x_2, x_3, \lambda) = {} & 2x_1^2 - 24x_1 + 2x_2^2 - 8x_2 + 2x_3^2 \\ & - 12x_3 + 200 + \lambda(x_1 + x_2 + x_3 - 11). \end{aligned}$$

The necessary conditions for the minimum of $Z$ are given by

$$\frac{\partial L}{\partial x_1} = 0, \quad \frac{\partial L}{\partial x_2} = 0, \quad \frac{\partial L}{\partial x_3} = 0, \quad \frac{\partial L}{\partial \lambda} = 0.$$

$$\text{Now, } \frac{\partial L}{\partial x_1} = 0 \text{ gives } 4x_1 - 24 + \lambda = 0 \Rightarrow x_1 = \frac{24 - \lambda}{4} \tag{10.54}$$

$$\frac{\partial L}{\partial x_2} = 0 \text{ gives } 4x_2 - 8 + \lambda = 0 \Rightarrow x_2 = \frac{8 - \lambda}{4} \tag{10.55}$$

$$\frac{\partial L}{\partial x_3} = 0 \text{ gives } 4x_3 - 12 + \lambda = 0 \Rightarrow x_3 = \frac{12 - \lambda}{4} \tag{10.56}$$

$$\frac{\partial L}{\partial \lambda} = 0 \text{ gives } x_1 + x_2 + x_3 - 11 = 0. \tag{10.57}$$

Substituting the values of $x_1, x_2$ and $x_3$ [from (10.54), (10.55), (10.56)] in (10.57), we have

$$\frac{44 - 3\lambda}{4} = 11 \Rightarrow \lambda = 0$$

$$\therefore\ x_1 = 6,\ \ x_2 = 2,\ \ x_3 = 3$$

$$\text{Now, } \frac{\partial^2 L}{\partial x_1^2} = 4, \quad \frac{\partial^2 L}{\partial x_1 \partial x_j} = 0, \quad j = 2, 3$$

$$\frac{\partial^2 L}{\partial x_2^2} = 4, \quad \frac{\partial^2 L}{\partial x_2 \partial x_j} = 0, \quad j = 1, 3$$

$$\frac{\partial^2 L}{\partial x_3^2} = 4, \quad \frac{\partial^2 L}{\partial x_3 \partial x_j} = 0, \quad j = 1, 2.$$

$$\text{Again, } \frac{\partial g}{\partial x_1} = 1, \quad \frac{\partial g}{\partial x_2} = 1, \quad \frac{\partial g}{\partial x_3} = 1.$$

Now we shall test the optimality by the bordered Hessian matrix method.

The bordered Hessian matrix at $x_1 = 6, x_2 = 2, x_3 = 3$ is given by

$$H_{\mathrm{B}} = \begin{bmatrix} O & U \\ U^{\mathrm{T}} & V \end{bmatrix},$$

where 0 is an $m \times m$ matrix, $U = \left[\frac{\partial g_i}{\partial x_j}\right]_{m\times n}$ and $V = \left[\frac{\partial^2 L}{\partial x_i \partial x_j}\right]_{n\times n}$,

$$\text{i.e. } H_{\mathrm{B}} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 4 & 0 & 0 \\ 1 & 0 & 4 & 0 \\ 1 & 0 & 0 & 4 \end{bmatrix}.$$

Since $m = 1$, $n = 3$, $n - m = 3 - 1 = 2$ and $2m + 1 = 3$, only two leading principal minors of $H_{\mathrm{B}}$ of order 3 and 4 (i.e. $\Delta_3$ and $\Delta_4$) are to be evaluated, and both the minors must have negative sign as $(-1)^1 = -1$ = negative.

$$\text{Now, } H_3 = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 4 & 0 \\ 1 & 0 & 4 \end{vmatrix} = -4 - 4 = -8 < 0$$

$$\text{and } H_4 = |H_B| = \begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 4 & 0 & 0 \\ 1 & 0 & 4 & 0 \\ 1 & 0 & 0 & 4 \end{vmatrix} = -1\begin{vmatrix} 1 & 0 & 0 \\ 1 & 4 & 0 \\ 1 & 0 & 4 \end{vmatrix} + \begin{vmatrix} 1 & 4 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 4 \end{vmatrix} - \begin{vmatrix} 1 & 4 & 0 \\ 1 & 0 & 4 \\ 1 & 0 & 0 \end{vmatrix}$$
$$= -16 + (-4)(4-0) = (-4)(0-4) = -16 - 16 - 16 = -48 < 0.$$

Hence, $z$ is minimum subject to the given constraint at $x_1 = 6, x_2 = 2, x_3 = 3$ and $Z_{\min} = 102$.

**Alternative test for optimality**

For the given problem, the Hancock determinantal equation at $x_1 = 6, x_2 = 2, x_3 = 3$ is given by

$$\begin{vmatrix} L_{11} - z & L_{12} & L_{13} & g_{11} \\ L_{21} & L_{22} - z & L_{23} & g_{12} \\ L_{31} & L_{32} & L_{33} - z & g_{13} \\ g_{11} & g_{12} & g_{13} & 0 \end{vmatrix} = 0,$$

$$\text{where } L_{ij} = \left.\frac{\partial^2 L}{\partial x_i \partial x_j}\right|_{\substack{x = x^* \\ \lambda = \lambda^*}}, \quad g_{ij} = \left.\frac{\partial g_i}{\partial x_j}\right|_{x=x^*}$$

$$\text{or} \begin{vmatrix} 4-z & 0 & 0 & 1 \\ 0 & 4-z & 0 & 1 \\ 0 & 0 & 4-z & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix} = 0$$

$$\text{or } (4-z)\begin{vmatrix} 4-z & 0 & 1 \\ 0 & 4-z & 1 \\ 1 & 1 & 0 \end{vmatrix} - \begin{vmatrix} 0 & 4-z & 0 \\ 0 & 0 & 4-z \\ 1 & 1 & 1 \end{vmatrix} = 0$$

or $(4-z)[(z-4)-(4-z)] - (4-z)(4-z) = 0$

or $(4-z)(z-4-4+z) - (4-z)^2 = 0$

or $(4-z)^2 = 0$

or $z = 4, 4$.

Since all the values of $z$ are positive, $Z$ is minimum subject to the given constraint at $x_1 = 6, x_2 = 2, x_3 = 3$, and $Z_{\min} = 102$.

*Example 8* Find the dimension of a cylindrical tin (with top and bottom) made up of sheet metal to maximize its volume such that the total surface area is equal to $24\pi$.

**Solution** Let $x_1$ be the radius of the top or bottom and $x_2$ be the height of the tin.

Then the volume of the tin $= \pi x_1^2 x_2$ and the total surface area $= 2\pi x_1^2 + 2\pi x_1 x_2$. Therefore, the problem is

$$\begin{aligned} \text{maximize} \quad & f(x_1, x_2) = \pi x_1^2 x_2 \\ \text{subject to} \quad & 2\pi x_1^2 + 2\pi x_1 x_2 = 24\pi \\ \text{or} \quad & \pi x_1^2 + \pi x_1 x_2 - 12\pi = 0 \\ \text{or} \quad & g_1(x_1, x_2) = \pi x_1^2 + \pi x_1 x_2 - 12\pi = 0. \end{aligned}$$

Now, the Lagrange function is

$$L(x_1, x_2, \lambda) = \pi x_1^2 x_2 + \lambda\left(\pi x_1^2 + \pi x_1 x_2 - 12\pi\right).$$

The necessary conditions for the maximum of $f$ are given by

$$\frac{\partial L}{\partial x_1} = 0, \quad \frac{\partial L}{\partial x_2} = 0, \quad \frac{\partial L}{\partial \lambda} = 0.$$

$$\text{Now, } \frac{\partial L}{\partial x_1} = 0 \text{ implies } 2\pi x_1 x_2 + \lambda\pi(2x_1 + x_2) = 0. \tag{10.58}$$

$$\begin{aligned} \text{Again, } \frac{\partial L}{\partial x_2} = 0 \text{ gives } & \pi x_1^2 + \lambda\pi x_1 = 0 \\ \text{or, } & \pi x_1^2 + \lambda\pi x_1 = 0 \\ \text{or, } & x_1 + \lambda = 0 \quad [\because x_1 \neq 0] \\ \text{or, } & x_1 = -\lambda. \end{aligned} \tag{10.59}$$

$$\text{Again, } \frac{\partial L}{\partial \lambda} = 0 \text{ gives } x_1^2 + x_1 x_2 - 12 = 0. \tag{10.60}$$

Now setting $x_1 = -\lambda$ in (10.58), we have

$$\begin{aligned} & 2\pi(-\lambda)x_2 + \lambda\pi(-2\lambda + x_2) = 0 \\ \text{or } & -\pi\lambda x_2 = 2\lambda^2\pi \\ \text{or } & x_2 = -2\lambda \quad [\because \lambda \neq 0] \\ \text{or } & x_2 = 2x_1 \quad [\text{From } (10.59), x_1 = -\lambda \text{ or, } \lambda = -x_1]. \end{aligned}$$

Setting $x_2 = 2x_1$ in (10.60), we have

$$\begin{aligned} & x_1^2 + x_1 \cdot 2x_1 - 12 = 0 \text{ or, } x_1^2 = 4 \\ \text{or } & x_1 = 2 \text{ [As } x_1 \text{ is the radius of the top or bottom of the tin].} \\ & \therefore x_2 = 2x_1 = 2.2 = 4 \quad \therefore x^* = (2, 4). \end{aligned}$$

From (10.59), we have $x_1 = -\lambda$ or, $\lambda = -2$.

Now the Hancock determinantal equation is

$$\begin{vmatrix} L_{11} - z & L_{12} & g_{11} \\ L_{21} & L_{22} - z & g_{12} \\ g_{11} & g_{12} & 0 \end{vmatrix} = 0. \tag{10.61}$$

We have $L_{11} = \left.\dfrac{\partial^2 L}{\partial x_1^2}\right|_{\substack{x=x^* \\ \lambda=-2}} = [2\pi x_2 + 2\pi\lambda]_{x'} = 2\pi \cdot 4 + 2\pi \cdot (-2) = 8\pi - 4\pi = 4\pi$

$$L_{12} = \left.\frac{\partial^2 L}{\partial x_1 \partial x_2}\right|_{\substack{x=x^* \\ \lambda=-2}} = [2\pi x_1 + 2\pi]_{x'} = 2\pi \cdot 2 + (-2)\pi = 2\pi.$$

Similarly, $L_{21} = 2\pi$.

$$\text{Again, } L_{22} = \left.\tfrac{\partial^2 L}{\partial x_2^2}\right|_{\substack{x=x^* \\ \lambda=-2}} = 0$$

Therefore, (10.61) becomes

$$\begin{aligned} & \begin{vmatrix} 4\pi - z & 2\pi & 8\pi \\ 2\pi & 0 - z & 2\pi \\ 8\pi & 2\pi & 0 \end{vmatrix} = 0 \\ \text{or, } & (4\pi - z)(0 - 4\pi^2) - 2\pi(0 - 16\pi^2) + 8\pi(4\pi^2 + 8\pi z) = 0 \\ \text{or, } & 68\pi^2 z + 48\pi^3 = 0 \\ \text{or, } & 68\pi^2 z = -48\pi^3 \quad \text{or, } \quad z = -\frac{48}{68}\pi = -\frac{12}{17}\pi. \end{aligned}$$

Since the value of $z$ is negative, $f$ is maximum for $x^* = (2, 4)$.

## 10.6 Interpretation of Lagrange Multipliers

To find the physical meaning of the Lagrange multipliers, let us consider the following optimization problem involving only one constraint:

$$\text{Minimize } f(x)$$
$$\text{subject to } h(x) = b,$$

$$\text{i.e. Minimize } f(x) \tag{10.62}$$

$$\text{subject to } g(x) = b - h(x) = 0, \tag{10.63}$$

where $b$ is a constant.

The necessary conditions for the minimum of $f(x)$ are given by

$$\frac{\partial L}{\partial x_j} = \frac{\partial f}{\partial x_j} + \lambda \frac{\partial g}{\partial x_j} = 0, \quad j = 1, 2, \ldots, n \tag{10.64}$$

$$\text{and } \frac{\partial L}{\partial \lambda} = g = 0,$$

$$\text{where } L(x, \lambda) = f(x) + \lambda g(x). \tag{10.65}$$

Let the solution of (10.64) and (10.65) be given by $x^*$, $\lambda^*$ and $f^* = f(x^*)$.

Now, we want to find the effect of a small relaxation or tightening of the constraint on the optimum value of the objective function (i.e. we want to find the effect of a small change in $b$ on $f^*$).

From (10.63), $g(x) = 0$, where $g(x) = b - h(x)$.

$$\therefore \, \mathrm{d}g = 0, \text{ which implies } \mathrm{d}b - \mathrm{d}h = 0$$

$$\text{or, } \mathrm{d}b = \mathrm{d}h, \text{ i.e. } \mathrm{d}b = \sum_{j=1}^{n} \frac{\partial h}{\partial x_j} \mathrm{d}x_j. \tag{10.66}$$

Equation (10.64) can be rewritten as

$$\frac{\partial f}{\partial x_j} - \lambda \frac{\partial h}{\partial x_j} = 0 \quad \text{or, } \frac{\partial h}{\partial x_j} = \frac{1}{\lambda} \frac{\partial f}{\partial x_j}, \quad j = 1, 2, \ldots, n. \tag{10.67}$$

From (10.66) and (10.67), we have

$$\mathrm{d}b = \sum_{j=1}^{n} \frac{1}{\lambda} \frac{\partial f}{\partial x_j} \mathrm{d}x_j = \frac{\mathrm{d}f}{\lambda} \quad \left[ \text{Since } \mathrm{d}f = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} \mathrm{d}x_j \right] \tag{10.68}$$

$$\text{or, } \lambda = \frac{\mathrm{d}f}{\mathrm{d}b} \quad \text{or,} \quad \lambda^* = \frac{\mathrm{d}f^*}{\mathrm{d}b} \quad \text{or,} \quad \mathrm{d}f^* = \lambda^* \mathrm{d}b. \tag{10.69}$$

Thus, $\lambda^*$ denotes the sensitivity (or rate of change) of $f$ with respect to $b$ or the marginal or incremental change in $f^*$ with respect to $b$ at $x^*$. In other words, $\lambda^*$ indicates how the constraint can be relaxed or tightly binding at the optimum point.

Depending on the value of $\lambda^*$, three cases may arise: Case I ($\lambda^* > 0$), Case II ($\lambda^* < 0$), Case III ($\lambda^* = 0$).

**Case-I**: $\lambda^* > 0$

In this case, for a unit decrease in $b$, the decrease in $f^*$ will be exactly equal to $\lambda^*$ as $\mathrm{d}f = \lambda^*(-1) = -\lambda^*$. Hence, $\lambda^*$ may be interpreted as the marginal gain (or reduction) in $f^*$ due to the tightening of the constraint. On the other hand, if $b$ is increased by one unit, $f$ will be increased to a new optimum level by the amount $\mathrm{d}f = \lambda^*(+1) = \lambda^*$. In this case, $\lambda^*$ may be interpreted as the marginal cost (increase) in $f^*$ due to relaxation of the constraint.

**Case-II**: $\lambda^* < 0$

In this case, the optimum value of $f$ will be decreased with the increase of $b$. Here, the marginal gain (reduction) in $f^*$ due to a relaxation of the constraint by one unit is determined by the value of $\lambda^*$ as $\mathrm{d}f^* = \lambda^*(+1) < 0$. If $b$ is decreased by one unit, the marginal cost (increase) in $f^*$ by the tightening of the constraint is $\mathrm{d}f^* = \lambda^*(-1) > 0$, since, in this case, the minimum value of the objective function increases.

**Case-III**: $\lambda^* = 0$

In this case, any incremental change in $b$ has absolutely no effect on the optimum value of $f$, and hence the constraint will not be binding. This means that the optimization of $f$ subject to $g = 0$ leads to the same optimum point $x^*$ as with the unconstrained optimization of $f$.

In economics and operations research, Lagrange multipliers are known as shadow prices of the constraints, since they indicate the changes in optimal value of the objective function per unit change in the right-hand side of the equality constraints.

**Procedure for finding the effect of changes of $b$ on $f^*$**

This can be done in two ways:

(i) For finding the effect of changes of $b$ (right-hand side of the constraints) on the optimal value $f^*$, one can solve the problem with the new value of $b$.

(ii) Another procedure is to use the relation $\mathrm{d}f^* = \lambda^* \, \mathrm{d}b$, where $\lambda^*$ is the optimal value.

*Example 9* Find the maximum of the function $f(x) = 2x_1 + x_2 + 10$ subject to $g(x) = x_1 + 2x_2^2 = 3$ using the Lagrange multiplier method. Also, find the effect of changing the right-hand side of the constraint on the optimum value of $f$.

**Solution** Here the Lagrange function is given by

$$\begin{aligned} L &= f(x) + \lambda g(x) \\ &= 2x_1 + x_2 + 10 + \lambda\left(3 - x_1 - 2x_2^2\right). \end{aligned} \tag{10.70}$$

The necessary conditions for the maximum of $f(x)$ are given by

$$\frac{\partial L}{\partial x_1} = 0, \quad \frac{\partial L}{\partial x_2} = 0, \quad \frac{\partial L}{\partial \lambda} = 0. \tag{10.71}$$

$$\begin{aligned} &\text{Now, } \frac{\partial L}{\partial x_1} = 0 \text{ gives } 2 - \lambda = 0 \quad \text{or} \quad \lambda = 2 \\ &\frac{\partial L}{\partial x_2} = 0 \text{ gives } 1 - 4\lambda x_2 = 0 \quad \text{or} \quad x_2 = \frac{1}{4\lambda} = \frac{1}{8} = 0.13 \\ &\frac{\partial L}{\partial \lambda} = 0 \text{ gives } x_1 + 2x_2^2 - 3 = 0 \quad \text{or} \quad x_1 = 3 - 2x_2^2 = 3 - \frac{1}{32} = \frac{95}{32} = 2.97. \end{aligned}$$

Hence, the solution of (10.71) is given by

$$x_1^* = 2.97, \quad x_2^* = 0.13, \quad \lambda^* = 2.$$

To test the sufficient condition, solve the Hancock determinantal equation for finding the value of $z$. The value of $z$ will be negative.

Hence, $f$ will be maximum for $x_1^* = 2.97$, $x_2^* = 0.13$, and the maximum value of $f$ will be $f^* = f(x^*) = 16.07$.

**Second part**: When the original constraint is tightened by one unit, i.e. $db = -1$,

$$df^* = \lambda^* db = 2(-1) = -2.$$

Due to this change, the new value of $f^*$ is $f_{\text{old}}^* + df^* = 16.07 - 2 = 14.07$.

On the other hand, if we relax the original constraint by one unit, i.e. $db = +1$,

$$df^* = \lambda^* db = 2(1) = 2.$$

Due to this change, the new value of $f^*$ is $f_{\text{old}}^* + df^* = 16.07 + 2 = 18.07$

*Example 10* Find the maximum of the function $f(x) = 4x_1 + x_2 + 15$ subject to $g(x) = 5 - x_1 - 4x_2^2 = 0$, using the Lagrange multiplier method. Also, find the effect of changing the right-hand side of the constraint on the optimum value of $f$.

**Solution** The Lagrange function is given by

$$\begin{aligned} L &= f(x) + \lambda\, g(x) \\ &= 4x_1 + x_2 + 15 + \lambda\left(5 - x_1 - 4x_2^2\right). \end{aligned} \tag{10.72}$$

The necessary conditions for the maximum of $f(x)$ are given by

$$\frac{\partial L}{\partial x_1} = 0, \quad \frac{\partial L}{\partial x_2} = 0, \quad \frac{\partial L}{\partial \lambda} = 0. \tag{10.73}$$

$$\begin{aligned} \text{Now, } \frac{\partial L}{\partial x_1} &= 0 \text{ gives } 4 - \lambda = 0 \Rightarrow \lambda = 4 \\ \frac{\partial L}{\partial x_2} &= 0 \text{ gives } 1 - 8\lambda x_2 = 0 \Rightarrow x_2 = \frac{1}{8\lambda} = \frac{1}{32} = 0.03 \\ \frac{\partial L}{\partial \lambda} &= 0 \text{ gives } 5 - x_1 - 4x_2^2 = 0 \Rightarrow x_1 = -4x_2^2 + 5 = 5 - \frac{4}{(32)^2} = 4.99. \end{aligned}$$

Hence, the solution of (10.73) is given by

$$x_1^* = 4.99, \quad x_2^* = 0.03, \lambda = 4.$$

To test the sufficient condition, solve the Hancock determinantal equation for finding the value of $z$.

$$\begin{aligned} \text{Now, } \frac{\partial^2 L}{\partial x_1^2} &= 0, \quad \frac{\partial^2 L}{\partial x_1 \partial x_2} = 0 \\ \frac{\partial^2 L}{\partial x_2^2} &= -8\lambda = -32, \quad \frac{\partial^2 L}{\partial x_2 \partial x_1} = 0 \\ \text{and } \frac{\partial g}{\partial x_1} &= -1, \quad \frac{\partial g}{\partial x_2} = -8x_2 = -0.24. \end{aligned}$$

Now the Hancock determinantal equation at $x_1^* = 4.99, x_2^* = 0.03$ and $\lambda^* = 4$ is given by

$$\begin{vmatrix} 0 - z & 0 & -1 \\ 0 & -32 - z & -0.24 \\ -1 & -0.24 & 0 \end{vmatrix} = 0$$

$$\text{or, } z = -\frac{32}{1.0576} = -30.26,$$

since the value of $z$ is negative.

Hence, $f$ will be maximum for $x_1^* = 4.99, x_2^* = 0.03$, and the maximum value of $f$ will be $f^* = f(x^*) = 34.99$

**Second part**: When the original constraint is tightened by one unit, i.e. $\mathrm{d}b = -1$,

$$\mathrm{d}f^* = \lambda^* \mathrm{d}b = 4(-1) = -4.$$

Due to this change, the new value of $f^*$ is

$$f_{\text{old}}^* + \mathrm{d}f^* = 34.99 - 4 = 30.99.$$

On the other hand, if we relax the original constraint by one unit, i.e. $\mathrm{d}b = +1$,

$$\mathrm{d}f^* = \lambda^* \mathrm{d}b = 4(1) = 4.$$

Due to this change, the new value of $f^*$ is

$$f_{\text{old}}^* + \mathrm{d}f^* = 34.99 + 4 = 38.99.$$

## 10.7 Applications of Lagrange Multiplier Method

**Financial mathematics**: Constrained optimization plays an important role in financial mathematics; e.g. the choice problem for a consumer is represented as one of maximizing a utility function subject to a budget constraint. The Lagrange multiplier has an economic interpretation as the shadow price associated with the constraint, since the Lagrange multipliers indicate the changes in the optimal value of the objective function per unit change on the right-hand side of the equality constraint. Other examples include profit maximization for a firm, along with various applications in the branch of financial mathematics dealing with performance, structure, behaviour and decision-making of the whole economy.

**Control theory**: The Lagrange multiplier method is used in optimal control theory to find the best possible control for taking a dynamical system from one state to another, especially in the presence of constraints for the state or input controls.

## 10.8 Exercises

1. Solve the following problems by the Lagrange multiplier method:

   (i) Optimize $z = x_1^2 - 10x_1 + x_2^2 - 6x_2 + x_3^2 - 4x_3$
   subject to $x_1 + x_2 + x_3 = 7$.

   (ii) Optimize $z = 3x_1^2 + x_2^2 + x_3^2$
   subject to $x_1 + x_2 + x_3 = 2$.

2. Solve the problem by the Lagrange multiplier method:

   Minimize $z = 4x_1^2 + 2x_2^2 + x_3^2 - 4x_1x_2$
   subject to the constraints
   $$x_1 + x_2 + x_3 = 15$$
   $$2x_1 - x_2 + 2x_3 = 20.$$

3. Solve the following problem by the Lagrange multiplier method:

   Minimize $z = x_1^2 + x_2^2 + x_3^2$
   subject to the constraints
   $$x_1 + x_2 + 3x_3 = 2$$
   $$5x_1 + 2x_2 + x_3 = 5$$

4. Solve the problem by the Lagrange multiplier method:

   Maximize $z = 6x_1 + 8x_2 + x_1^2 - x_2^2$
   subject to the constraints
   $4x_1 + 3x_2 = 16$
   $3x_1 + 5x_2 = 15.$

5. Solve the following optimization problems:

   (i) Minimize $z = x_1x_2x_3$
   subject to the constraints
   subject to $x_1 + x_2 + x_3 = 1$
   and $x_1, x_2, x_3 > 0$.

(ii) Minimize $z = x_1^2 + x_2^2$
subject to the constraint $x_1 + x_2 = 2$ .

(iii) Maximize $z = x_1 + x_2$
subject to the constraint $x_1^2 + x_2^2 = 1$.

(iv) Minimize $z = \frac{1}{2}\left(x_1^2 + x_2^2 + x_3^2\right)$
subject to the constraint $x_1 + x_2 + x_3 = 1$.

(v) Minimize $z = 2x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 - 8x_1 - 6x_2 - 4x_3 + 9$
subject to the constraint $x_1 + x_2 + 2x_3 = 3$ .

# Chapter 11
# Constrained Optimization with Inequality Constraints

## 11.1 Objective

The objective of this chapter is to derive the Kuhn-Tucker necessary and sufficient conditions to solve multivariate constrained optimization problems with inequality constraints.

## 11.2 Introduction

Let us consider a multivariate minimization problem as follows:

$$\begin{aligned}&\text{Minimize } f(x)\\&\text{subject to } g_i(x) \le 0, \quad i = 1, 2, \ldots, m..\\&\text{where } x \in \mathbb{R}^n,\ f, g_i\colon \mathbb{R}^n \to \mathbb{R}\end{aligned}$$

Now adding the non-negative slack variables, $y_i^2$ in the given inequality constraints, we get the equality constraints as

$$g_i(x) + y_i^2 = 0, \quad i = 1, 2, \ldots, m.$$

Let $G_i(x, y) = g_i(x) + y_i^2 = 0, i = 1, 2, \ldots, m$, where $y = (y_1, y_2, \ldots, y_m)^{\mathrm{T}}$ is the vector of slack variables.

Therefore, the reduced problem is

$$\begin{aligned}&\text{Minimize } f(x)\\&\text{subject to } G_i(x, y) = g_i(x) + y_i^2 = 0, \quad i = 1, 2, \ldots, m.\end{aligned}$$

This problem can be solved by the method of Lagrange multiplier.

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*, Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_11

In this case, the Lagrange function is

$$L(x, y, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i G_i(x, y),$$

where $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_m)^{\mathrm{T}}$ is the vector of Lagrange multiplier.

Now, the necessary conditions for a stationary point to be a local minimum are given by

$$\frac{\partial L}{\partial x_j} = 0, \quad j = 1, 2, \ldots, n$$

$$\frac{\partial L}{\partial \lambda_i} = 0, \quad i = 1, 2, \ldots, m$$

$$\text{and} \quad \frac{\partial L}{\partial y_i} = 0, \quad i = 1, 2, \ldots, m,$$

$$\text{i.e.} \quad \frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0, \quad j = 1, 2, \ldots, n \tag{11.1}$$

$$g_i(x) + y_i^2 = 0, \quad i = 1, 2, \ldots, m \tag{11.2}$$

and

$$2\lambda_i y_i = 0, \quad i = 1, 2, \ldots, m. \tag{11.3}$$

From (11.3), we have $\lambda_i y_i = 0$.

Hence, either $\lambda_i = 0$ or $y_i = 0, \quad i = 1, 2, \ldots, m$.

For any $i$, if $\lambda_i = 0$, it means that the corresponding constraint is inactive and hence can be ignored. On the other hand, if $y_i = 0$, it means that the constraint is active at the optimum point.

Now, we divide the indices of the constraints into two subsets $J_1$ and $J_2$. Let the set $J_1$ indicate the indices of those constraints which are active at the optimum point, and let $J_2$ include the indices of all the inactive constraints,

$$\text{i.e.} \quad J_1 = \{i\colon y_i = 0, \quad i = 1, 2, \ldots, m\}$$

$$J_2 = \{i\colon \lambda_i = 0, \quad i = 1, 2, \ldots, m\},$$

i.e. $i \in J_1$ when $y_i = 0$ and $i \in J_2$ when $\lambda_i = 0, i = 1, 2, \ldots, m$.

Hence, (11.1) becomes

$$\frac{\partial f}{\partial x_j} + \sum_{i \in J_1} \lambda_i \frac{\partial g_i}{\partial x_j} = 0, \quad j = 1, 2, \ldots, n. \tag{11.4}$$

Similarly, (11.2) becomes

$$g_i(x) = 0 \quad \text{for} \quad i \in J_1 \tag{11.5}$$

and

$$g_i(x) + y_i^2 = 0 \quad \text{for} \quad i \in J_2. \tag{11.6}$$

If the number of active constraints is $p$, then the number of inactive constraints is $m - p$. Therefore, in (11.4), (11.5), and (11.6), the number of equations are $n + p + m - p = m + n$ and the unknowns are $x_1$, $x_2$, …, $x_n$, $\lambda_i(i \in J_1)$ and $y_i(i \in J_2)$. Therefore, the number of unknowns are $n + p + m - p = m + n$. Hence, we can solve the system (11.4), (11.5), and (11.6) easily.

Let

$$\nabla f = \left(\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_2}\cdots\frac{\partial f}{\partial x_n}\right)^{\mathrm{T}} \quad \text{and} \quad \nabla g_i = \left(\frac{\partial g_i}{\partial x_1}\frac{\partial g_i}{\partial x_2}\cdots\frac{\partial g_i}{\partial x_n}\right)^{\mathrm{T}}$$

be the gradients of the objective function and the $i$th constraint respectively.

Then Eq. (11.4) may be written as

$$-\nabla f = \sum_{i \in J_1} \lambda_i \nabla g_i. \tag{11.7}$$

Thus, the negative of the gradient of the objective function can be expressed as a linear combination of the gradients of the active constraints at the optimum point.

Now, it can be proved that:

(a) For a minimization problem with $g_i(x) \leq 0$, we have $\lambda_i \geq 0$.
(b) For a maximization problem with $g_i(x) \leq 0$, we have $\lambda_i \leq 0$.
(c) For a minimization problem with $g_i(x) \geq 0$, we have $\lambda_i \leq 0$.
(d) For a maximization problem with $g_i(x) \geq 0$, we have $\lambda_i \geq 0$.

## 11.3 Feasible Direction

The direction of a vector $S$ is called a feasible direction from a point $x = (x_1, x_2, \ldots, x_n)$ if at least a small step can be taken along $S$ that does not immediately leave the feasible region. Thus, for problems with sufficiently smooth constraint surfaces, any vector $S$ satisfying the relation

$$< S^{\mathrm{T}}, \nabla g_i > < 0$$

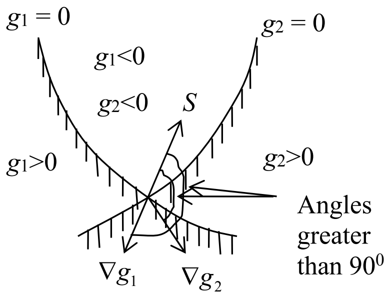



**Fig. 11.1** Feasible direction $S$ when both constraint functions are convex

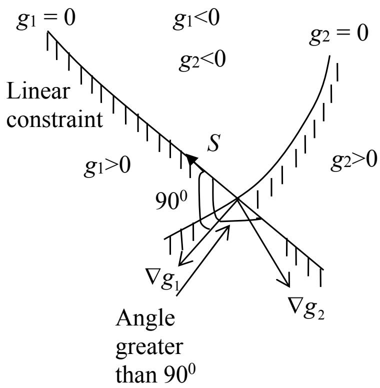



**Fig. 11.2** Feasible direction $S$ when one constraint function is linear

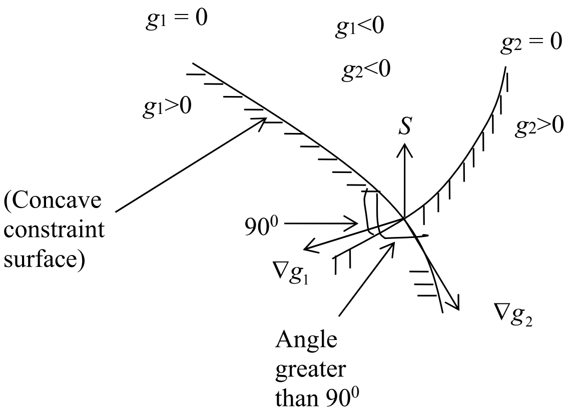



**Fig. 11.3** Feasible direction $S$ when one constraint function is concave and other one is convex

can be called a feasible direction (see Fig. 11.1). On the other hand, if the constraint is either linear or concave (see Figs. 11.2 and 11.3), any vector satisfying the relation

$$<S^T, \nabla g_i><0$$

is called a feasible direction.

Let us suppose that only two constraints are active at the optimum point, i.e. $p = 2$. Then Eq. (11.7) reduces to

$$-\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2. \tag{11.8}$$

Let $S$ be a feasible direction at the optimum point.

Now by premultiplying both sides of (11.8) by $S^{\mathrm{T}}$, we have

$$-<S^{\mathrm{T}}, \nabla f> \; = \lambda_1 <S^{\mathrm{T}}, \nabla g_1> \; + \lambda_2 <S^{\mathrm{T}}, \nabla g_2> .$$

Since $S$ is a feasible direction, then we have

$$<S^{\mathrm{T}}, \nabla g_1><0 \quad \text{and} \quad <S^{\mathrm{T}}, \nabla g_2><0. \tag{11.9}$$

Thus, if $\lambda_1 > 0$ and $\lambda_2 > 0$, the quantity $<S^{\mathrm{T}}, \nabla f>$ is always positive. As $\nabla f$ indicates the gradient direction along which the values of the function $f$ increase at the maximum rate, $<S^{\mathrm{T}}, \nabla f>$ represents the component of the increment of $f$ along the direction $S$. If $<S^{\mathrm{T}}, \nabla f> \; > 0$, the function value increases as we move along the direction $S$.

Hence, if $\lambda_1, \lambda_2$ are positive, we will not be able to find any direction in the feasible region along which the function value can further be decreased. Since the point at which Eq. (11.9) is valid is assumed to be optimum, $\lambda_1$ and $\lambda_2$ have to be positive. This reason can be extended to cases where there are more than two active constraints.

## 11.4 Kuhn-Tucker Conditions

In this section, we shall derive the necessary and sufficient conditions for a local optimum of the general non-linear programming problem.

### *11.4.1 Kuhn-Tucker Necessary Conditions*

Let us consider a minimization problem as follows:

$$\begin{array}{l} \text{Minimize } f(x) \\ \text{subject to } g_i(x) \le 0, \quad i = 1, 2, \ldots, m. \end{array} \quad \text{where} x \in \mathbb{R}^n,\ f, g_i\colon \mathbb{R}^n \to \mathbb{R}$$

Now adding the non-negative slack variables $y_i^2$ in the given inequality constraints, we get the equality constraints as

$$g_i(x) + y_i^2 = 0, \quad i = 1, 2, \ldots, m.$$

Therefore, the reduced problem is

$$\begin{aligned} &\text{Minimize}\, f(x) \\ &\text{subject to}\, g_i(x) + y_i^2 = 0, \quad i = 1, 2, \ldots, m. \end{aligned}$$

This problem can be solved by the Lagrange multiplier method.

In this case, the Lagrange function is

$$L(x, y, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i \left[ g_i(x) + y_i^2 \right],$$

where $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)^{\mathrm{T}}$ is the vector of Lagrange multipliers.

Now the necessary conditions for a stationary point to be a local minimum are given by

$$\begin{aligned} \frac{\partial L}{\partial x_j} &= 0, \quad j = 1, 2, \ldots, n \\ \frac{\partial L}{\partial \lambda_i} &= 0, \quad i = 1, 2, \ldots, m \\ \text{and} \quad \frac{\partial L}{\partial y_i} &= 0, \quad i = 1, 2, \ldots, m, \end{aligned}$$

$$\text{i.e.} \quad \frac{\partial L}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0, \quad j = 1, 2, \ldots, n \tag{11.10}$$

$$g_i(x) + y_i^2 = 0, \quad i = 1, 2, \ldots, m \tag{11.11}$$

and

$$2\lambda_i y_i = 0, \quad i = 1, 2, \ldots, m. \tag{11.12}$$

If a constraint is satisfied with an equality sign $g_i(x) = 0$ at the optimum point, then it is called the active constraint. Otherwise, it will be called inactive.

Equation (11.12) implies that either $\lambda_i = 0$ or $y_i = 0$. Now, if $\lambda_i = 0$ and $y_i^2 = 0$, then the $i$th constraint is inactive and can be discarded. If $y_i = 0, \lambda_i > 0$, then Eq. (11.11) implies $g_i(x) = 0$ at the optimum point. This means that the constraint is active. Therefore, Eqs. (11.11) and (11.12) imply

$$\lambda_i g_i(x) = 0. \tag{11.13}$$

Again, $y_i^2 \geq 0$. Then from (11.11), we have

$$g_i(x) \leq 0. \tag{11.14}$$

Thus, (11.13) implies that if $g_i(x) < 0$, then $\lambda_i = 0$ and when $g_i(x) = 0,\ \lambda_i > 0$.

Hence, the necessary conditions to be satisfied at a local minimum point are as follows:

$$\begin{aligned}
&\frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0, \quad j = 1, 2, \ldots, m \\
&\lambda_i g_i(x) = 0 \\
&g_i(x) \leq 0 \\
&\lambda_i \geq 0, \quad i = 1, 2, \ldots, m.
\end{aligned}$$

This set of necessary conditions is known as the Kuhn-Tucker necessary conditions.

**Note**: The first two conditions will be the same for both minimization as well as maximization problems. However, the last two conditions will be different, and these are dependent on the type of problem. The constraint conditions are given in the table.

| Type of problem | Type of constraint | Form of last two necessary conditions |
|---|---|---|
| Minimization | $g_i(x) \leq 0$ | $g_i(x) \leq 0, \lambda_i \geq 0$ |
| Minimization | $g_i(x) \geq 0$ | $g_i(x) \geq 0, \lambda_i \leq 0$ |
| Maximization | $g_i(x) \leq 0$ | $g_i(x) \leq 0, \lambda_i \leq 0$ |
| Maximization | $g_i(x) \geq 0$ | $g_i(x) \geq 0, \lambda_i \geq 0$ |

### *11.4.2 Kuhn-Tucker Sufficient Conditions*

**Theorem** *The Kuhn-Tucker necessary conditions for the optimization problem*

$$\begin{gathered}
\text{Minimize}\, f(x) \\
\text{subject to the constraints } g_i(x) \leq 0, \quad i = 1, 2, \ldots, m
\end{gathered}$$

*are also sufficient conditions if $f(x)$ is convex and all $g_i(x)$ are convex functions of $x$.*

*Proof* The Kuhn-Tucker necessary conditions for the given problem

$$\begin{aligned}&\text{Minimize } f(x)\\ &\text{subject to } g_i(x) \le 0, \quad i = 1, 2, \ldots, m\\ &\text{are } \frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0, \quad j = 1, 2, \ldots, n\\ &\lambda_i g_i(x) = 0, \quad i = 1, 2, \ldots, m\\ &g_i(x) \le 0, \quad i = 1, 2, \ldots, m\\ &\lambda_i \ge 0, \quad i = 1, 2, \ldots, m.\end{aligned}$$

Now, we have to prove that these conditions are sufficient if $f(x)$ is convex and all $g_i(x)$ are convex functions of $x$.

The Lagrangian function of the given problem is

$$L(x, y, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i \left[g_i(x) + y_i^2\right].$$

If $\lambda_i \ge 0$, then $\lambda_i g_i(x)$ is convex.

Further, since $\lambda_i y_i = 0$, we have $g_i(x) + y_i^2 = 0$.

Thus, it follows that $L(x, y, \lambda)$ is a convex function.

Again, the necessary condition for $f(x)$ to be minimum at $x$ is that $L(x, y, \lambda)$ has a stationary point.

However, if $L(x, y, \lambda)$ is convex, its partial derivatives must vanish at only one point. According to the convexity analysis, this point must be a global minimum point. Hence, the Kuhn-Tucker necessary conditions are also sufficient conditions for a global minimum of $f(x)$ at $x$ if $f(x)$ and all $g_i(x)$ are convex.

*Example 1* Determine $x_1$, $x_2$, $x_3$ so as to maximize

$$z = -x_1^2 - x_2^2 - x_3^2 + 4x_1 + 6x_2$$

subject to the constraints

$$x_1 + x_2 \le 2, 2x_1 + 3x_2 \le 12 \text{ and } x_1, x_2, x_3 \ge 0.$$

**Solution**

Let

$$f = -x_1^2 - x_2^2 - x_3^2 + 4x_1 + 6x_2$$

subject to $g_1 = x_1 + x_2 - 2 \le 0$, $g_2 = 2x_1 + 3x_2 - 12 \le 0$.

Hence, the Kuhn-Tucker necessary conditions are

$$\frac{\partial f}{\partial x_j} + \sum_{i=1}^{2} \lambda_i \frac{\partial g_i}{\partial x_j} = 0, \quad j = 1, 2, 3 \tag{11.15}$$

$$\lambda_i g_i = 0, \quad i = 1, 2 \tag{11.16}$$

$$g_i \le 0, \quad i = 1, 2 \tag{11.17}$$

$$\lambda_1 \le 0, \quad \lambda_2 \le 0. \tag{11.18}$$

From (11.15), we have

$$-2x_1 + 4 + \lambda_1 + 2\lambda_2 = 0 \tag{11.19a}$$

$$-2x_2 + 6 + \lambda_1 + 3\lambda_2 = 0 \tag{11.19b}$$

$$-2x_3 = 0. \tag{11.19c}$$

From (11.16), we have $\lambda_i g_i = 0, i = 1, 2,$

$$\text{i.e. } \lambda_1(x_1 + x_2 - 2) = 0 \tag{11.20a}$$

$$\lambda_2(2x_1 + 3x_2 - 12) = 0. \tag{11.20b}$$

From (11.17), we have $g_i \le 0, i = 1, 2,$

$$\text{i.e. } x_1 + x_2 - 2 \le 0 \tag{11.21a}$$

and

$$2x_1 + 3x_2 - 12 \le 0. \tag{11.21b}$$

According to (11.18), four different cases may arise:

**Case 1**: $\lambda_1 = 0, \quad \lambda_2 = 0$

**Case 2**: $\lambda_1 = 0, \quad \lambda_2 \ne 0$

**Case 3**: $\lambda_1 \ne 0, \quad \lambda_2 = 0$

**Case 4**: $\lambda_1 \ne 0, \quad \lambda_2 \ne 0.$

Now we shall examine all the cases.

**Case 1**: $\lambda_1 = 0, \quad \lambda_2 = 0$

In this case, from (11.19a), (11.19b) and (11.19c), we have $-2x_1 + 4 = 0$, $-2x_2 + 6 = 0$ and $x_3 = 0$, i.e. $x_1 = 2, x_2 = 3, x_3 = 0$.

However, this solution violates both constraints (11.21a) and (11.21b). So, this solution is rejected.

**Case 2**: $\lambda_1 = 0, \lambda_2 \ne 0$

In this case, (11.20a) and (11.20b) give $2x_1 + 3x_2 - 12 = 0$, and (11.19a) and (11.19b) give $-2x_1 + 4 + 2\lambda_2 = 0$, $-2x_2 + 6 + 3\lambda_2 = 0$, i.e. $x_1 = \lambda_2 + 2$ and $x_2 = \frac{3}{2}\lambda_2 + 3$.

Substituting these results in $2x_1 + 3x_2 = 12$, we have

$$2(\lambda_2 + 2) + 3\left(\frac{3}{2}\lambda_2 + 3\right) = 12$$

$$\text{or} \quad 2\lambda_2 + 4 + \frac{9}{2}\lambda_2 + 9 = 12 \;\text{ or }\; \frac{13}{2}\lambda_2 = -1 \;\text{ or, }\; \lambda_2 = -\frac{2}{13} < 0.$$

$$\therefore\; x_1 = -\frac{2}{13} + 2 = \frac{24}{13},\;\; x_2 = \frac{3}{2}\lambda_2 + 3 = \frac{3}{2}\left(-\frac{2}{13}\right) + 3 = -\frac{3}{13} + 3 = \frac{36}{13}.$$

Again, from (11.19c), $x_3 = 0$

$$\therefore\; x_1 = \frac{24}{13},\;\; x_2 = \frac{36}{13},\;\; x_3 = 0.$$

This solution violates the constraint $x_1 + x_2 - 2 \le 0$.
So this solution is rejected.

**Case 3**: $\lambda_1 \neq 0, \lambda_2 = 0$

In this case, (11.20a) gives $x_1 + x_2 = 2$, and (11.19a) and (11.19b) give $x_1 = 2 + \frac{\lambda_1}{2}, x_2 = 3 + \frac{\lambda_1}{2}$.

$$\therefore\; 2 + \frac{\lambda_1}{2} + 3 + \frac{\lambda_1}{2} = 2 \;\text{ or, }\; \lambda_1 = -3 < 0$$

$$\therefore\; x_1 = 2 - \frac{3}{2} = \frac{1}{2},\;\; x_2 = 3 - \frac{3}{2} = \frac{3}{2}.$$

Again from (11.19c), we have $x_3 = 0$.
Hence, the solution is $x_1 = \frac{1}{2},\;\; x_2 = \frac{3}{2},\;\; x_3 = 0$.
This solution satisfies the constraint $2x_1 + 3x_2 \le 12$.

**Case 4**: $\lambda_1 \neq 0, \lambda_2 \neq 0$

In this case, (11.20a) and (11.20b) give $x_1 + x_2 = 2$, $2x_1 + 3x_2 = 12$.
Solving these equations, we have $x_1 = -6$, $x_2 = 8$.
Now, from (11.19a) and (11.19b) we have $\lambda_1 + 2\lambda_2 = -16, \lambda_1 + 3\lambda_2 = 10$.
Solving these equations, we have $\lambda_2 = 26, \lambda_1 = -68$.
Since $\lambda_2 = 26$ violates the condition $\lambda_2 \le 0$, this solution is also rejected.
Hence, the optimum (maximum) solution of the given problem is $x_1 = \frac{1}{2}$, $x_2 = \frac{3}{2}, x_3 = 0$ with $\lambda_1 = 3, \lambda_2 = 0$ and Max $z = \frac{17}{2}$.

*Example 2*

Optimize $z = 2x_1 + 3x_2 - \left(x_1^2 + x_2^2 + x_3^2\right)$
subject to the constraints
$x_1 + x_2 \leq 1$
$2x_1 + 3x_2 \leq 6.$

**Solution**
Here we have

$$f(x) = 2x_1 + 3x_2 - x_1^2 - x_2^2 - x_3^2$$
$$g_1(x) = x_1 + x_2 - 1 \leq 0$$
$$g_2(x) = 2x_1 + 3x_2 - 6 \leq 0.$$

First of all, we have to determine whether the given problem is of minimization type or maximization type. For this purpose, we have to construct the bordered Hessian matrix as follows:

$$H_B = \begin{bmatrix} \mathrm{O} & U \\ U^{\mathrm{T}} & V \end{bmatrix},$$

where O is an $m \times m$ null matrix,

$$U = \left[\frac{\partial g_i}{\partial x_j}\right]_{m\times n} \quad \text{and} \quad V = \left[\frac{\partial^2 L}{\partial x_i\, \partial x_j}\right]_{n\times n}.$$

Here $m = 2$, $n = 3$.
Hence,

$$|H_B| = \begin{vmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 3 & 0 \\ 1 & 2 & -2 & 0 & 0 \\ 1 & 3 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{vmatrix}$$
$$= -2.$$

Since $n - m = 1$, $2m + 1 = 5$, $(-1)^{m+n}$= negative and $|H_B| < 0$, the solution point should maximize the objective function. Hence the Kuhn-Tucker conditions for the given problem are given by

$$\frac{\partial f}{\partial x_j} + \sum_{i=1}^{2} \lambda_i \frac{\partial g_i}{\partial x_j} = 0, \quad j = 1, 2, 3 \tag{11.22}$$

$$\lambda_i g_i = 0, \quad i = 1, 2 \tag{11.23}$$

$$g_i \leq 0, \quad i = 1, 2 \tag{11.24}$$

$$\lambda_1 \leq 0, \quad \lambda_2 \leq 0. \tag{11.25}$$

From (11.22), we have

$$2 - 2x_1 + \lambda_1 + 2\lambda_2 = 0 \tag{11.26a}$$

$$3 - 2x_2 + \lambda_1 + 3\lambda_2 = 0 \tag{11.26b}$$

$$-2x_3 = 0. \tag{11.26c}$$

From (11.23), we have

$$\left.\begin{array}{l} \lambda_1(x_1 + x_2 - 1) = 0 \\ \lambda_2(2x_1 + 3x_2 - 6) = 0 \end{array}\right\}. \tag{11.27a, b}$$

From (11.24), we have

$$x_1 + x_2 - 1 \leq 0 \tag{11.28a}$$

$$2x_1 + 3x_2 - 6 \leq 0. \tag{11.28b}$$

According to (11.25), four different cases may arise:

**Case 1**: $\lambda_1 = 0, \lambda_2 = 0$

**Case 2**: $\lambda_1 = 0, \lambda_2 \neq 0$

**Case 3**: $\lambda_1 \neq 0, \lambda_2 = 0$

**Case 4**: $\lambda_1 \neq 0, \lambda_2 \neq 0$.

Now we shall examine the different cases.

**Case 1**: $\lambda_1 = 0, \lambda_2 = 0$

In this case, from (11.26a)–(11.26c), we have $x_1 = 1, x_2 = 3/2, x_3 = 0$. This solution does not satisfy the inequalities (11.28a) and (11.28b). So this solution is rejected.

**Case 2**: $\lambda_1 = 0, \lambda_2 \neq 0$

In this case, solving (11.26a)–(11.26c), (11.27a) and (11.27b), we have

$$x_1 = \frac{12}{13}, \quad x_2 = \frac{18}{13}, \quad x_3 = 0 \;\text{ and }\; \lambda_2 = \frac{1}{13}.$$

This solution does not satisfy inequality (11.28a) and is therefore rejected.

**Case 3**: $\lambda_1 \neq 0, \lambda_2 = 0$

In this case, the solution of Eqs. (11.26a)–(11.26c) and (11.27a) yields $x_1 = \frac{1}{4}, x_2 = \frac{3}{4}, x_3 = 0$ and $\lambda_1 = \frac{3}{2}$. This solution satisfies all the conditions and is therefore accepted.

**Case 4**: $\lambda_1 \neq 0, \lambda_2 \neq 0$

In this case, the solution of Eqs. (11.26a)–(11.26c), (11.27a) and (11.27b) gives $x_1 = -3, x_2 = 4, x_3 = 0$, $\lambda_1 = -34, \lambda_2 = 13$. This solution violates the conditions (11.25) and is therefore discarded.

Since only one solution satisfies all the conditions, this one is thus the optimal solution, and it is given by $x_1 = \frac{1}{4}, x_2 = \frac{3}{4}, x_3 = 0$ and $\text{Max}\, z = \frac{17}{8}$.

*Example 3* Solve the following problem:

$$\begin{aligned}
&\text{Minimize}\, f = x_1^2 + x_2^2 + x_3^2 + 40x_1 + 20x_2 - 3000 \\
&\text{subject to}\, g_1 = x_1 - 50 \geq 0 \\
&\qquad\qquad g_2 = x_1 + x_2 - 100 \geq 0 \\
&\qquad\qquad g_3 = x_1 + x_2 + x_3 - 150 \geq 0.
\end{aligned}$$

**Solution**

The Kuhn-Tucker necessary conditions are

$$\frac{\partial f}{\partial x_j} + \sum_{i=1}^{3} \lambda_i \frac{\partial g_i}{\partial x_j} = 0, \quad j = 1, 2, 3 \tag{11.29}$$

$$\lambda_i g_i = 0, \quad i = 1, 2, 3 \tag{11.30}$$

$$g_i \geq 0, \quad i = 1, 2, 3 \tag{11.31}$$

$$\lambda_i \leq 0, \quad i = 1, 2, 3. \tag{11.32}$$

From (11.29), we have

$$\frac{\partial f}{\partial x_j} + \sum_{i=1}^{3} \lambda_i \frac{\partial g_i}{\partial x_j} = 0, \quad j = 1, 2, 3,$$

$$\text{i.e.} \quad \frac{\partial f}{\partial x_j} + \lambda_1 \frac{\partial g_1}{\partial x_j} + \lambda_2 \frac{\partial g_2}{\partial x_j} + \lambda_3 \frac{\partial g_3}{\partial x_j} = 0, \; j = 1, 2, 3,$$

$$\text{i.e.} \; 2x_1 + 40 + \lambda_1 + \lambda_2 + \lambda_3 = 0 \tag{11.33a}$$

$$2x_2 + 20 + \lambda_2 + \lambda_3 = 0 \tag{11.33b}$$

$$2x_3 + \lambda_3 = 0. \tag{11.33c}$$

From (11.30), we have $\lambda_i g_i = 0, i = 1, 2, 3,$

$$\text{i.e.} \; \lambda_1(x_1 - 50) = 0 \tag{11.34a}$$

$$\lambda_2(x_1 + x_2 - 100) = 0 \tag{11.34b}$$

$$\lambda_3(x_1 + x_2 + x_3 - 150) = 0. \tag{11.34c}$$

From (11.31), we have $g_i \geq 0, i = 1, 2, 3,$

$$\begin{aligned} \text{i.e.} \; g_1 &\geq 0 \\ g_2 &\geq 0 \\ g_3 &\geq 0, \end{aligned}$$

$$\text{i.e.} \; x_1 - 50 \geq 0 \tag{11.35a}$$

$$x_1 + x_2 - 100 \geq 0 \tag{11.35b}$$

$$x_1 + x_2 + x_3 - 150 \geq 0. \tag{11.35c}$$

From (11.32), we have $\lambda_i \leq 0, i = 1, 2, 3,$

$$\text{i.e.} \; \lambda_1 \leq 0 \tag{11.36a}$$

$$\lambda_2 \leq 0 \tag{11.36b}$$

$$\lambda_3 \leq 0. \tag{11.36c}$$

Now from (11.34a), we have $\lambda_1(x_1 - 50) = 0$.
Hence, either $\lambda_1 = 0$ or $x_1 - 50 = 0$,
i.e. either $\lambda_1 = 0$ or $x_1 = 50$.
**Case 1**: When $\lambda_1 = 0$
Now from Eqs. (11.33a)–(11.33c), we have

$$2x_1 = -40 - \lambda_1 - \lambda_2 - \lambda_3 \quad \text{or} \;\; x_1 = -20 - \frac{\lambda_2}{2} - \frac{\lambda_3}{2} \quad [\because \; \lambda_1 = 0]$$

$$\left.\begin{array}{l} 2x_2 = -20 - \lambda_2 - \lambda_3 \;\; \text{or,} \;\; x_2 = -10 - \frac{\lambda_2}{2} - \frac{\lambda_3}{2} \\ 2x_3 = -\lambda_3 \;\; \text{or,} \;\; x_3 = -\frac{\lambda_3}{2} \end{array}\right\}. \tag{11.37}$$

Substituting these values of $x_1$, $x_2$, $x_3$ in Eqs. (11.34b) and (11.34c), we have

$$\lambda_2\left(-20 - \frac{\lambda_2}{2} - \frac{\lambda_3}{2} - 10 - \frac{\lambda_2}{2} - \frac{\lambda_3}{2} - 100\right) = 0$$

and

$$\begin{array}{l} \lambda_3\left(-20 - \frac{\lambda_2}{2} - \frac{\lambda_3}{2} - 10 - \frac{\lambda_2}{2} - \frac{\lambda_3}{2} - \frac{\lambda_3}{2} - 150\right) = 0, \\ \text{i.e.} \quad \lambda_2(-130 - \lambda_2 - \lambda_3) = 0 \;\; \text{and} \;\; \lambda_3\left(-180 - \lambda_2 - \frac{3}{2}\lambda_3\right) = 0. \end{array} \tag{11.38}$$

From the first equation of (11.38), we have

$$\text{either}\, \lambda_2 = 0 \;\; \text{or} \;\; -130 - \lambda_2 - \lambda_3 = 0.$$

From the second equation of (11.38), we have

$$\text{either} \;\; \lambda_3 = 0 \;\; \text{or} \;\; -180 - \lambda_2 - \frac{3}{2}\lambda_3 = 0.$$

Hence, the four possible solutions of (11.38) are given by

$$\lambda_2 = 0, \;\; -180 - \lambda_2 - \frac{3}{2}\lambda_3 = 0 \tag{A}$$

$$\lambda_3 = 0, \;\; -130 - \lambda_2 - \lambda_3 = 0 \tag{B}$$

$$\lambda_2 = 0, \quad \lambda_3 = 0 \tag{C}$$

$$-130 - \lambda_2 - \lambda_3 = 0, \quad -180 - \lambda_2 - \frac{3}{2}\lambda_3 = 0. \tag{D}$$

From (A), $\lambda_2 = 0, -180 - \lambda_2 - \frac{3}{2}\lambda_3 = 0$, i.e. $\lambda_2 = 0$ and $-\frac{3}{2}\lambda_3 = 180$ or $\lambda_2 = 0$ and $\lambda_3 = -120$.

Now substituting $\lambda_2 = 0$ and $\lambda_3 = -120$ in (11.37), we have

$$\begin{aligned}
x_1 &= -20 - \frac{0}{2} - \frac{-120}{2} \quad \text{i.e.,} \quad x_1 = 40 \\
x_2 &= -10 - \frac{0}{2} - \frac{-120}{2} \quad \text{i.e.,} \quad x_2 = 50 \\
x_3 &= -\frac{-120}{2} = 60.
\end{aligned}$$

Hence, (A) gives the solution $x_1 = 40$, $x_2 = 50$, $x_3 = 60$.

This solution satisfies (11.35c), but it does not satisfy (11.35a) and (11.35b). So, we reject this solution.

Now from (B), we have $\lambda_3 = 0, -130 - \lambda_2 - \lambda_3 = 0,$

$$\text{i.e.} \quad \lambda_3 = 0 \quad \text{and} \quad \lambda_2 = -130.$$

Hence, from (11.37), we have

$$\begin{aligned}
x_1 &= -20 - \frac{-130}{2} - \frac{0}{2} = -20 + 65 = 45, \\
x_2 &= -10 - \frac{-130}{2} - \frac{0}{2} = 55, \quad x_3 = 0. \\
\therefore\ x_1 &= 45, \quad x_2 = 55, \quad x_3 = 0.
\end{aligned}$$

This solution satisfies (11.36), but it does not satisfy (11.35a) and (11.35c). Hence, we reject this solution.

Now from (C), we have $\lambda_2 = 0, \lambda_3 = 0.$

From (11.37), we have

$$x_1 = -20, \quad x_2 = -10, \quad x_3 = 0.$$

This solution satisfies (11.36), but it does not satisfy (11.35a)–(11.35c). Hence, we reject this solution.

Now from (D), we have

$$\begin{aligned}
&-130 - \lambda_2 - \lambda_3 = 0, \quad -180 - \lambda_2 - \frac{3}{2}\lambda_3 = 0, \\
&\text{i.e.} \quad \lambda_2 + \lambda_3 = -130 \quad \text{and} \quad \lambda_2 + \frac{3}{2}\lambda_3 = -180.
\end{aligned}$$

Subtracting the first equation from the second, we have

$$\frac{1}{2}\lambda_3 = -50 \quad \text{or} \quad \lambda_3 = -100.$$

From the first, $\lambda_2 - 100 = -130 \quad \text{or,} \quad \lambda_2 = -30.$

Hence, from (11.37), we have

$$x_1 = -20 - \frac{-30}{2} - \frac{-100}{2} = -20 + 15 + 50 = 45$$
$$x_2 = -10 - \frac{-30}{2} - \frac{-100}{2} = -10 + 15 + 50 = 55$$
$$x_3 = -\frac{-100}{2} = 50.$$

This solution satisfies (11.35b) and (11.35c) but does not satisfy (11.35a). Hence, we reject this solution.

**Case 2**: When $x_1 = 50$

Now from Eqs. (11.33a)–(11.33c), we have

$$\begin{aligned} \lambda_3 &= -2x_3 \\ \lambda_2 &= -20 - 2x_2 - \lambda_3 = -20 - 2x_2 + 2x_3 \\ \text{and } \lambda_1 &= -40 - 2x_1 - \lambda_2 - \lambda_3 = -40 - 2.50 - (-20 - 2x_2 + 2x_3) + 2x_3 \\ &= -40 - 100 + 20 + 2x_2 - 2x_3 + 2x_3 \\ &= -120 + 2x_2. \end{aligned}$$

$$\left.\begin{aligned} \therefore\ \lambda_1 &= -120 + 2x_2 \\ \lambda_2 &= -20 - 2x_2 + 2x_3 \\ \lambda_3 &= -2x_3 \end{aligned}\right\}. \tag{11.39}$$

Now substituting these values of $\lambda_1, \lambda_2, \lambda_3$ in (11.34b) and (11.34c), we have

$$(-20 - 2x_2 + 2x_3)(x_1 + x_2 - 100) = 0 \tag{11.40a}$$

and

$$-2x_3(x_1 + x_2 + x_3 - 150) = 0. \tag{11.40b}$$

From (11.40a), we have

$$\text{either } -20 - 2x_2 + 2x_3 = 0 \text{ or, } x_1 + x_2 - 100 = 0.$$

From (11.40b), we have either $x_3 = 0$ or $x_1 + x_2 + x_3 - 150 = 0$.

Therefore, the four possible solutions of (11.40a) and (11.40b) are given by

$$-20 - 2x_2 + 2x_3 = 0, \quad x_1 + x_2 + x_3 - 150 = 0 \tag{11.41a}$$

$$-20 - 2x_2 + 2x_3 = 0, \quad x_3 = 0 \tag{11.41b}$$

$$x_1 + x_2 - 100 = 0, \quad x_3 = 0 \tag{11.41c}$$

$$x_1 + x_2 - 100 = 0, \quad x_1 + x_2 + x_3 - 150 = 0. \tag{11.41d}$$

From (11.41a), we have

$$\begin{aligned} & -20 - 2x_2 + 2x_3 = 0, x_1 + x_2 + x_3 - 150 = 0, \\ \text{i.e.} \quad & -10 - x_2 + x_3 = 0, 50 + x_2 + x_3 - 150 = 0 \quad [\because \ x_1 = 50]. \end{aligned}$$

Adding the preceding two equations, we have

$$2x_3 = 10 - 50 + 150 \quad \text{or} \ \ 2x_3 = 110 \ \ \text{or} \ \ x_3 = 55.$$

Now substituting $x_3 = 55$ in $-10 - x_2 + x_3 = 0$, we have

$$\begin{aligned} & -10 - x_2 + 55 = 0 \\ & \text{or} \ \ x_2 = 45. \end{aligned}$$

$$\therefore \quad x_1 = 50, \ \ x_2 = 45, \ \ x_3 = 55.$$

This solution does not satisfy (11.35b); hence, we reject it.
From (11.41b), we have

$$\begin{aligned} & -20 - 2x_2 + 2x_3 = 0, \quad x_3 = 0, \\ \text{i.e.} \quad & -10 - x_2 + x_3 = 0, \quad x_3 = 0. \\ \therefore \ & x_1 = 50, \quad x_2 = -10, \quad x_3 = 0. \end{aligned}$$

This solution does not satisfy the constraints (11.35b) and (11.35c).
From (11.41c), we have

$$\begin{aligned} & x_1 + x_2 - 100 = 0, \quad x_3 = 0 \\ \text{or} \quad & 50 + x_2 - 100 = 0, \quad x_3 = 0 \quad [\because \ x_1 = 50] \\ \text{or} \quad & x_2 = 50, x_3 = 0. \\ \therefore \ & x_1 = 50, \quad x_2 = 50, \quad x_3 = 0. \end{aligned}$$

This solution does not satisfy the constraint (11.35c).
From (11.41d), we have

$$x_1 + x_2 - 100 = 0, \quad x_1 + x_2 + x_3 - 150 = 0.$$

From $x_1 + x_2 - 100 = 0$ we get $50 + x_2 - 100 = 0$ or $x_2 = 50$.
From $x_1 + x_2 + x_3 - 150 = 0$ we get $50 + 50 + x_3 - 150 = 0$ or $x_3 = 50$.

$$\therefore \ x_1 = 50, \quad x_2 = 50, \quad x_3 = 50.$$

This solution satisfies all the constraints (11.35a)–(11.35c).

Now substituting $x_1 = 50, x_2 = 50, x_3 = 50$ in (11.39), we have

$$\lambda_1 = -120 + 2.50 = -20$$
$$\lambda_2 = -20 - 2.50 + 2.50 = -20$$
$$\lambda_3 = -2.50 = -100.$$

Since all these values of $\lambda_i$ satisfy the requirement $\lambda_i \leq 0, \quad i = 1, 2, 3$, then the optimum solution is $x_1 = 50, \quad x_2 = 50, \quad x_3 = 50$, and the minimum value of $f$ is

$$(50)^2 + (50)^2 + (50)^2 + 40 \times 50 + 20 \times 50 - 3000 = 7500.$$

## 11.5 Exercises

(1) Maximize $f(x_1, x_2) = 2x_1 + \beta x_2$
subject to
$g_1(x_1, x_2) = x_1^2 + x_2^2 \leq 5$
$g_2(x_1, x_2) = x_1 - x_2 \leq 2.$

Using the Kuhn-Tucker conditions, find the values of $\beta$ for which $x_1 = 1, x_2 = 2$ will be the optimal solution.

2. Solve the following problems using the Kuhn-Tucker conditions:

(i) Minimize $f(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 5)^2$
subject to $-x_1^2 + x_2 \leq 4, \quad -(x_1 - 2)^2 + x_2 \leq 3.$

(ii) Maximize $Z = 10x_1 - x_1^2 + 10x_2 - x_2^2$
subject to $x_1 + x_2 \leq 14, \quad -x_1 + x_2 \leq 6.$

(iii) Maximize $z = 12x_1 + 21x_2 + 2x_1x_2 - 2x_1^2 - 2x_2^2$
subject to $x_2 \leq 8, \quad x_1 + x_2 \leq 10.$

(iv) Maximize $Z = 2x_1 + 3x_2 - 2x_1^2$
subject to $x_1 + 4x_2 \leq 4, \quad x_1 + x_2 \leq 2.$

(v) Minimize $Z = 6 - 6x_1 - 2x_1^2 - 2x_1x_2 + 2x_2^2$
subject to $x_1 + x_2 \leq 2.$

(vi) Maximize $z = 10x_1 + 25x_2 - 10x_1^2 - x_2^2 - 4x_1x_2$
subject to $x_1 + 2x_2 \leq 10, \quad x_1 + x_2 \leq 9.$

# Chapter 12
# Quadratic Programming

## 12.1 Objective

The main objective of this chapter is to discuss:

- Solution procedures of quadratic programming problems
- Wolfe's modified simplex method
- Beale's method for solving quadratic programming problems.

## 12.2 Introduction

Quadratic programming deals with non-linear programming problems of optimizing (maximizing or minimizing) the quadratic function subject to a set of linear inequality constraints. The general form of a quadratic programming problem (QPP) may be written as follows:

$$\text{Optimize } f(x) = \sum_{j=1}^{n} c_j x_j + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} \mathrm{d}_{jk} x_j x_k$$

subject to the constraints

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad i = 1, 2, \ldots, m$$

$$\text{and} \quad x_j \ge 0, \quad j = 1, 2, \ldots, n.$$

© Springer Nature Singapore Pte Ltd. 2019
A. K. Bhunia et al., *Advanced Optimization and Operations Research*, Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_12

In matrix notation, it is written as

$$\text{Optimize}\, f(x) = \left\langle c^{\mathrm{T}}, x \right\rangle + \frac{1}{2}\langle Qx, x\rangle$$

subject to the constraints

$$Ax \leq b \;\text{ and }\; x \geq 0,$$

where $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$, $c = (c_1, c_2, \ldots, c_n)$ $b = (b_1, b_2, \ldots, b_n)^{\mathrm{T}}$, $A = \left(a_{ij}\right)$ is an $m \times n$ real matrix and $Q = \left(\mathrm{d}_{jk}\right)$ is an $n \times n$ real symmetric matrix, i.e. $\mathrm{d}_{jk} = \mathrm{d}_{kj}$ for all $j$ and $k$.

The matrix $Q$ can be taken as non-null otherwise the QPP is a linear programing problem.

The term $\langle Qx, x\rangle$ defines a quadratic form where $Q$ is a symmetric matrix. The quadratic form $\langle Qx, x\rangle$ is said to be positive definite (negative definite) if $\langle Qx, x\rangle > 0(<0)$ for $x \neq 0$ and positive semi-definite (negative semi-definite) if $\langle Qx, x\rangle \geq 0$ $(\leq 0)$ for all $x$ such that there is one $x \neq 0$ satisfying $\langle Qx, x\rangle = 0$.

It may easily be verified that:

(i) If $\langle Qx, x\rangle$ is positive semi-definite (negative semi-definite), then it is convex (concave) in $x$ over all of $R^n$

(ii) If $\langle Qx, x\rangle$ is positive definite (negative definite), then it is strictly convex (strictly concave) in $x$ over all of $R^n$.

These results help the students in determining whether the quadratic objective function $f(x)$ is concave (convex) and the implication of the same on the sufficiency of the Kuhn-Tucker conditions for constrained maxima (minima) of $f(x)$.

The feasibility of the problem can be tested by Phase I of the simplex method. The solution of a general constrained optimization problem can be classified as follows:

(i) No feasible solution
(ii) An unbounded solution
(iii) An optimal solution.

**Theorem 1** *The quadratic programming problem cannot have an unbounded solution when the matrix $Q$ is positive definite.*

*Proof* Let $x \neq 0$. The objective function can be written as

$$f(x) = \langle Qx, x\rangle\left(\frac{1}{2} + \left\langle c^{\mathrm{T}}, x\right\rangle / \langle Qx, x\rangle\right). \tag{12.1}$$

Let $x$ be any point on the hypersphere $|x| = r$, where $|x|^2 = \langle x, x\rangle$. Then $x = r\hat{x}$, where $|\hat{x}| = 1$, and then $\langle Qx, x\rangle = r^2\langle Q\hat{x}, \hat{x}\rangle$.

Let $m$ be the minimum value of $\langle Q\hat{x}, \hat{x}\rangle$. Since $Q$ is positive definite,

$$\begin{aligned} &\langle Qx, x\rangle \geq r^2 m > 0 \\ &\therefore\ \langle Qx, x\rangle \to +\infty \ \text{ as } \ |x| = r \to \infty. \end{aligned} \tag{12.2}$$

Now, let $M$ be the maximum value of $\langle c^{\mathrm{T}}, \hat{x}\rangle / \langle Q\hat{x}, \hat{x}\rangle$.

Then, $\frac{\langle c^{\mathrm{T}}, x\rangle}{\langle Qx, x\rangle} = \frac{1}{r}\frac{\langle c^{\mathrm{T}}, \hat{x}\rangle}{\langle Q\hat{x}, \hat{x}\rangle} \leq \frac{M}{r}$,

and therefore

$$\frac{\langle c^{\mathrm{T}}, x\rangle}{\langle Qx, x\rangle} \to 0 \ \text{ as } \ r \to \infty. \tag{12.3}$$

From (12.1), (12.2) and (12.3), we get

$$f(x) \to \infty \ \text{ and } \ |x| \to \infty.$$

Hence, for finite $x, f(x)$ will be finite. This means that the QPP cannot have an unbounded solution.

Now we shall show that the QPP has a unique optimal solution provided the problem has a feasible solution and the matrix $Q$ is positive definite. If $Q$ is positive semi-definite, the solution could be an unbounded one. But the difficulties in this case can be avoided if we consider the perturbed matrix $Q + \varepsilon I$, where $(\varepsilon > 0)$ is substantially less than the magnitude of any element of $Q$, since this matrix is positive definite when $Q$ is positive semi-definite.

## 12.3 Solution Procedure of QPP

Let us consider a QPP in the following form:

$$\text{Maximize } \ z = f(x) = \sum_{j=1}^{n} c_j x_j + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} \mathrm{d}_{jk} x_j x_k$$

subject to the constraints

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \ \text{ and } \ x_j \geq 0, \quad i = 1, 2, \ldots, m, \ \ j = 1, 2, \ldots, n,$$

where $\mathrm{d}_{jk} = \mathrm{d}_{kj}$ for all $j$, $k$, $b_i \geq 0$ for $i = 1, 2, \ldots, m$ and the matrix $Q = \left(\mathrm{d}_{jk}\right)_{n\times n}$ is symmetric negative definite.

Now, introducing slack variables $q_i^2$ in the $i$th constraint $(i = 1, 2, \ldots, m)$ and $r_j^2$ in the $j$th non-negativity constraint $(j = 1, 2, \ldots, n)$, the problem becomes

$$\text{Maximize } z = f(x) = \sum_{j=1}^{n} c_j x_j + \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} \mathrm{d}_{jk} x_j x_k$$

subject to the constraints

$$\sum_{j=1}^{n} a_{ij} x_j + q_i^2 = b_i, \quad i = 1, 2, \ldots, m$$
$$-x_j + r_j^2 = 0, \quad j = 1, 2, \ldots, n.$$

Now the Lagrangian function is given by

$$L(x_1, x_2, \ldots, x_n, q_1, q_2, \ldots, q_m, r_1, r_2, \ldots, r_n, \lambda_1, \lambda_2, \ldots, \lambda_m, \mu_1, \mu_2, \ldots, \mu_n) = f(x) - \sum_{i=1}^{m} \lambda_i \left( \sum_{j=1}^{n} a_{ij} x_j + q_i^2 - b_i \right) - \sum_{j=1}^{n} \mu_j \left( -x_j + r_j^2 \right),$$

where $\lambda_i \geq 0, i = 1, 2, \ldots, m$ and $\mu_j \geq 0, j = 1, 2, \ldots, n$.

The necessary conditions for the maximum of $L$ are given by

$$\frac{\partial L}{\partial x_j} = 0 \Rightarrow \frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} \lambda_i a_{ij} + \mu_j = 0, \quad j = 1, 2, \ldots, n \tag{12.4}$$

$$\frac{\partial L}{\partial q_i} = 0 \Rightarrow 2\lambda_i q_i = 0, \quad i = 1, 2, \ldots, m \tag{12.5}$$

$$\frac{\partial L}{\partial r_j} = 0 \Rightarrow 2\mu_j r_j = 0, \quad j = 1, 2, \ldots, n \tag{12.6}$$

$$\frac{\partial L}{\partial \lambda_i} = 0 \Rightarrow \sum_{j=1}^{n} a_{ij} x_j + q_i^2 - b_i = 0, \quad i = 1, 2, \ldots, m \tag{12.7}$$

$$\frac{\partial L}{\partial \mu_j} = 0 \Rightarrow -x_j + r_j^2 = 0, \quad j = 1, 2, \ldots, n. \tag{12.8}$$

These conditions are known as Kuhn-Tucker conditions.

By defining a set of new variables $s_i$ as $s_i = q_i^2 \geq 0, i = 1, 2, \ldots, m$, the equations in (12.5) become

$$\lambda_i s_i = 0, \quad i = 1, 2, \ldots, m.$$

Then the equations in (12.7) reduce to

$$\sum_{j=1}^{n} a_{ij} x_j + s_i - b_i = 0, \quad i = 1, 2, \ldots, m.$$

Now from (12.6) and (12.8), we have

$$\mu_j x_j = 0$$
$$\text{and} \quad x_j \geq 0, \quad j = 1, 2, \ldots, n.$$

The equations in (12.4) can be written as

$$c_j + \sum_{k=1}^{n} d_{jk} x_k - \sum_{i=1}^{m} \lambda_i a_{ij} + \mu_j = 0, \quad j = 1, 2, \ldots, n.$$

Thus, the necessary conditions can be summarized as follows:

$$c_j + \sum_{k=1}^{n} d_{jk} x_k - \sum_{i=1}^{m} \lambda_i a_{ij} + \mu_j = 0, \quad j = 1, 2, \ldots, n \tag{12.9}$$

$$\sum_{j=1}^{n} a_{ij} x_j + s_i = b_i, \quad i = 1, 2, \ldots, m \tag{12.10}$$

$$\lambda_i s_i = 0, \quad i = 1, 2, \ldots, m \tag{12.11}$$

$$\mu_j x_j = 0, \quad j = 1, 2, \ldots, n \tag{12.12}$$

$$x_j, \mu_j \geq 0, \quad j = 1, 2, \ldots, n \tag{12.13}$$

$$s_i, \lambda_i \geq 0, \quad i = 1, 2, \ldots, m. \tag{12.14}$$

These conditions, except $\lambda_i s_i = 0$ and $\mu_j x_j = 0$, are linear constraints involving 2 $(n + m)$ variables. Thus, the solution of the original QPP can be obtained by finding a non-negative solution to the set of $(n + m)$ linear equations given by the equations in (12.9) and (12.10), which also satisfies the $(n + m)$ equations stated in (12.11) and (12.12).

Since $Q$ is a negative definite matrix, $f(x)$ is a strictly concave function and the feasible space is convex (because of linear equations), any local maximum of the problem will be the global maximum. Again, from (12.9) to (12.14), it is observed that there are $2(n+m)$ equations with $2(n+m)$ variables in the necessary conditions. Hence, the solution of the system (12.9)–(12.14) must be unique. Thus, the feasible solution of the system (12.9)–(12.14), if it exists, must give the optimum solution of the QPP directly. The solution of the system (12.9)–(12.14) can be obtained by using Phase I of the two-phase simplex method. The only restriction here is that the satisfaction of non-linear relations (12.11) and (12.12) has to be maintained all the time. Since our objective is just to find a feasible solution to the system, there is no necessity for Phase II computations. For the application of Phase I of the two-phase method, to get the initial basis as an identity matrix, we have to introduce artificial variables $A_j$ into Eq. (12.9) so that

$$c_j + \sum_{k=1}^{n} d_{jk}x_k - \sum_{i=1}^{m} \lambda_i a_{ij} + \mu_j + A_j = 0, \quad j = 1, 2, \ldots, n.$$

Hence, the equivalent optimization QPP becomes

$$\text{Maximize } z' = -\sum_{j=1}^{n} A_j$$

subject to the constraints

$$\begin{aligned}
&c_j + \sum_{k=1}^{n} d_{jk}x_k - \sum_{i=1}^{m} \lambda_i a_{ij} + \mu_j + A_j = 0, \quad j = 1, 2, \ldots, n \\
&\sum_{j=1}^{n} a_{ij}x_j + s_i = b_i, \quad i = 1, 2, \ldots, m \\
&x_j, \mu_j, A_j \geq 0, \quad j = 1, 2, \ldots, n \;\; s_i, \lambda_i \geq 0, \;\; i = 1, 2, \ldots, m.
\end{aligned}$$

While solving this problem, we have to consider the following additional conditions known as complementary slackness conditions:

$$\begin{aligned}
\lambda_i s_i &= 0, \quad i = 1, 2, \ldots, m \\
\mu_j x_j &= 0, \quad j = 1, 2, \ldots, n
\end{aligned}$$

Thus, when deciding whether to introduce $s_i$ into the basic solution, we have to ensure first that either $\lambda_i$ is not in the basic solution or $\lambda_i$ will be removed when $s_i$ enters the basic solution. Similar care has to be taken regarding the variables $\mu_j$ and $x_j$.

## 12.4 Wolfe's Modified Simplex Method

Let the QPP be the following:

$$\text{Maximize } z = f(x) = \sum_{j=1}^{n} c_j x_j + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} \mathrm{d}_{jk} x_j x_k$$

subject to the constraints

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad i = 1, 2, \ldots, m \text{ and } x_j \ge 0, \; j = 1, 2, \ldots, n,$$

where $\mathrm{d}_{jk} = \mathrm{d}_{kj}$ for all $j$ and $k$, $b_i \ge 0$ for all $i$.

Also, assume that the quadratic form $\sum_{j=1}^{n}\sum_{k=1}^{n} \mathrm{d}_{jk} x_j x_k$ is negative semi-definite.

Then the iterative procedure for the solution of a modified simplex method is as follows:

Step 1. Check whether the given QPP is in the maximization form or not. If it is in the minimization form, convert it into the maximization form by appropriate adjustment in $f(x)$.

Step 2. Convert the inequality constraints into equations by introducing the slack variables $q_i^2$ in the $i$th constraint ($i$ = 1, 2, …, $m$) and the slack variables $r_j^2$ in the $j$th non-negativity constraint ($j$ = 1, 2, …, $n$).

Step 3. Construct the Lagrangian function

$$L(x_1, x_2, \ldots, x_n,\; q_1, q_2, \ldots, q_m, r_1, r_2, \ldots, r_n,\; \lambda_1, \lambda_2, \ldots, \lambda_m,\; \mu_1, \mu_2, \ldots, \mu_n)$$
$$= f(x) - \sum_{i=1}^{m} \lambda_i \left( \sum_{j=1}^{n} a_{ij} x_j + q_i^2 - b_i \right) - \sum_{j=1}^{n} \mu_j \left( -x_j + r_j^2 \right),$$

where $\lambda_i \ge 0, i = 1, 2, \ldots, m$ and $\mu_j \ge 0, \; j = 1, 2, \ldots, n.$

Step 4. Derive the Kuhn-Tucker conditions by differentiating $L$ partially with respect to $x_1, x_2, \ldots, x_n,\; q_1, q_2, \ldots, q_m,\; r_1, r_2, \ldots, r_n, \lambda_1, \lambda_2, \ldots, \lambda_m, \mu_1, \mu_2, \ldots, \mu_n$ and then equating to zero.

Step 5. Introduce the non-negative artificial variables $A_j,\; j = 1, 2, \ldots, n$ in the first $n$ equations obtained from $\frac{\partial L}{\partial x_j} = 0,\; j = 1, 2, \ldots, n$ and construct the new objective function $z' = -\sum_{j=1}^{n} A_j$, and form an LPP with this new objective function.

Step 6. Use two-phase simplex method to obtain an optimum solution to the LPP formed in Step 5 satisfying the complementary slackness conditions.

Step 7. The optimal solution obtained in Step 6 is an optimal solution to the given QPP also.

Step 8. Stop.

## 12.5 Advantage and Disadvantage of Wolfe's Method

The advantage in using the Wolfe method is that only a slight modification has to be made in the simplex method to take care of the restricted basis entry rule. However, the disadvantage is that the number of rows of the coefficient matrix increases from $m$ to $m + n$.

*Example 1* Solve the following problem by Wolfe's method:

$$\begin{aligned} &\text{Minimize } z = x_1^2 - x_1x_2 + 2x_2^2 - x_1 - x_2 \\ &\text{subject to} \\ &\qquad 2x_1 + x_2 \le 1 \\ &\qquad x_1 \ge 0,\ x_2 \ge 0. \end{aligned}$$

**Solution**

$$\begin{aligned} &\text{Maximize } z' = -x_1^2 + x_1x_2 - 2x_2^2 + x_1 + x_2 \\ &\text{subject to} \\ &\qquad 2x_1 + x_2 \le 1 \\ &\qquad -x_1 \le 0, -x_2 \le 0 \\ &\text{where } z' = -z \end{aligned}$$

Now introducing slack variables $q_1^2$, $r_1^2$ and $r_2^2$, we have the reduced problem as

$$\begin{aligned} &\text{Maximize } z' = -x_1^2 + x_1x_2 - 2x_2^2 + x_1 + x_2 \\ &\text{subject to} \\ &\qquad 2x_1 + x_2 + q_1^2 = 1 \\ &\qquad -x_1 + r_1^2 = 0, -x_2 + r_2^2 = 0. \end{aligned}$$

Now the Lagrange function $L$ is given by

$$\begin{aligned} &L(x_1, x_2, q_1, q_2, r_1, r_2, \lambda_1, \lambda_2, \mu_1, \mu_2,) \\ &\quad = -x_1^2 + x_1x_2 - 2x_2^2 + x_1 + x_2 - \lambda_1\left(2x_1 + x_2 + q_1^2 - 1\right) \\ &\quad - \mu_1\left(-x_1 + r_1^2\right) - \mu_2\left(-x_2 + r_2^2\right). \end{aligned}$$

Clearly, $-x_1^2 + x_1x_2 - 2x_2^2 = -\left[x_1^2 + x_2^2 - 2x_1x_2\right] - x_1x_2 - x_2^2 = -(x_1 - x_2)^2 - x_1x_2 - x_2^2$ is a negative semi-definite quadratic form.

Hence, $z' = -x_1^2 + x_1x_2 - 2x_2^2 + x_1 + x_2$ is a concave function in $x_1, x_2$. Hence, the maximum of $L$ will be the maximum of $z'$ and vice versa.

Now the necessary conditions for maxima of $L$ (and hence of $z'$) are given by

$$\frac{\partial L}{\partial x_1} = 0 \Rightarrow -2x_1 + x_2 + 1 - 2\lambda_1 + \mu_1 = 0 \tag{12.15}$$

$$\frac{\partial L}{\partial x_2} = 0 \Rightarrow x_1 - 4x_2 + 1 - \lambda_1 + \mu_2 = 0 \tag{12.16}$$

$$\frac{\partial L}{\partial \lambda_1} = 0 \Rightarrow 2x_1 + x_2 + q_1^2 - 1 = 0 \tag{12.17}$$

$$\frac{\partial L}{\partial \mu_1} = 0 \Rightarrow -x_1 + r_1^2 = 0 \tag{12.18}$$

$$\frac{\partial L}{\partial \mu_2} = 0 \Rightarrow -x_2 + r_2^2 = 0 \tag{12.19}$$

$$\frac{\partial L}{\partial q_1} = 0 \Rightarrow \lambda_1 q_1 = 0 \tag{12.20}$$

$$\frac{\partial L}{\partial r_1} = 0 \text{ or } \mu_1 r_1 = 0 \text{ or } \mu_1 x_1 = 0 \text{ as } -x_1 + r_1^2 = 0 \tag{12.21}$$

$$\frac{\partial L}{\partial r_2} = 0 \text{ or } \mu_2 r_2 = 0 \text{ or } \mu_2 x_2 = 0 \text{ as } -x_2 + r_2^2 = 0 \tag{12.22}$$

Now letting $q_1^2 = s_1 \geq 0$, from (12.17) we get

$$2x_1 + x_2 + s_1 - 1 = 0, \tag{12.23}$$

and from (12.20), we get

$$\lambda_1 s_1 = 0, \tag{12.24}$$

where $x_1,\ x_2,\ s_1,\ \mu_1,\ \mu_2,\ \lambda_1 \geq 0$.

Now the equivalent problem of the given QPP is as follows:

$$\left.\begin{aligned} 2x_1 - x_2 + 2\lambda_1 - \mu_1 &= 1 \\ -x_1 + 4x_2 + \lambda_1 - \mu_2 &= 1 \\ 2x_1 + x_2 + s_1 &= 1 \end{aligned}\right\} \tag{12.25}$$

$$\begin{aligned} &x_1, x_2, s_1, \lambda_1, \mu_1, \mu_2 \geq 0 \\ &\lambda_1 s_1 = 0, \quad \mu_1 x_1 = 0, \quad \mu_2 x_2 = 0. \end{aligned} \tag{12.26}$$

A solution $x_j, j = 1, 2$ of (12.25) and satisfying (12.26) shall necessarily be an optimal one for maximizing $L$ (i.e. maximizing $z$).

Now to determine the solution to the simultaneous equation (12.25), we introduce two artificial variables $A_1(\geq 0)$ and $A_2(\geq 0)$ in the first two equations of (12.25) and construct the dummy objective function $z'' = -A_1 - A_2$. Then the equivalent problem of the given QPP becomes

$$\begin{aligned} &\text{Maximize } z'' = -A_1 - A_2 \\ &\text{subject to} \\ &\qquad 2x_1 - x_2 + 2\lambda_1 - \mu_1 + A_1 = 1 \\ &\qquad -x_1 + 4x_2 + \lambda_1 - \mu_2 + A_2 = 1 \\ &\qquad 2x_1 + x_2 + s_1 = 1 \\ &x_1, x_2 \geq 0, \lambda_1 \geq 0, \mu_1, \mu_2, s_1, A_1, A_2 \geq 0, \end{aligned}$$

satisfying the complementary slackness conditions

$$\lambda_1 s_1 = 0, \quad \mu_1 x_1 = 0, \quad \mu_2 x_2 = 0.$$

To solve the problem, we shall apply two-phase method.

Here the initial basic feasible solution (BFS) is $A_1 = 1, \;\; A_2 = 1, \;\; s_1 = 1$.

Now we construct the initial table of Phase I.

| | | $c_j \rightarrow$ | 0 | 0 | 0 | 0 | 0 | −1 | −1 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x_1'$ | $x_2'$ | $\lambda_1'$ | $\mu_1'$ | $\mu_2'$ | $A_1'$ | $A_2'$ | $s_1'$ | Mini ratio |
| −1 | $A_1$ | 1 | 2 | −1 | 2 | −1 | 0 | 1 | 0 | 0 | – |
| −1 | $A_2$ | 1 | −1 | 4 | 1 | 0 | −1 | 0 | 1 | 0 | 1/4 |
| 0 | $s_1$ | 1 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| $z_B' = -2$ | | | −1 | −3 | −3 | 1 | 1 | 0 | 0 | 0 | |

Here all $\Delta_j \ngeq 0$. Hence, the current BFS is not optimal. From this table, we observe that either $x_2'$ or $\lambda_1'$ as $(\Delta_2 = -3,\ \Delta_3 = -3)$ can be introduced into the basic solution. Arbitrarily, we choose $x_2'$ as the entering vector, since by the complementary condition, $s_1$ is already in the basis.

| | | $c_j \rightarrow$ | 0 | 0 | 0 | 0 | 0 | −1 | −1 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x_1'$ | $x_2'$ | $\lambda_1'$ | $\mu_1'$ | $\mu_2'$ | $A_1'$ | $A_2'$ | $s_1'$ | Mini ratio |
| −1 | $A_1$ | 5/4 | 7/4 | 0 | 9/4 | −1 | −1/4 | 1 | 1/4 | 0 | 5/7 |
| 0 | $x_2$ | 1/4 | −1/4 | 1 | 1/4 | 0 | −1/4 | 0 | 1/4 | 0 | – |
| 0 | $s_1$ | 3/4 | 9/4 | 0 | −1/4 | 0 | 1/4 | 0 | −1/4 | 1 | 1/3 |
| $z_B' = -5/4$ | | | −7/4 | 0 | −9/4 | 1 | 1/4 | 0 | 3/4 | 0 | |

Here all $\Delta_j \ngeq 0$. Hence, the current BFS is not optimal. $\Delta_j$ corresponding to $\lambda_1'$ is most negative, but it does not enter into the basis, since by the complementary condition, $s_1$ is already in the basis.

| | | $c_j \rightarrow$ | 0 | 0 | 0 | 0 | 0 | −1 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x_1'$ | $x_2'$ | $\lambda_1'$ | $\mu_1'$ | $\mu_2'$ | $A_1'$ | $s_1'$ | Mini ratio |
| −1 | $A_1$ | 2/3 | 0 | 0 | 22/9 | −1 | −4/9 | 1 | −7/9 | 3/11 |
| 0 | $x_2$ | 1/3 | 0 | 1 | 2/9 | 0 | −2/9 | 0 | 1/9 | 3/2 |
| 0 | $x_1$ | 1/3 | 1 | 0 | −1/9 | 0 | 1/9 | 0 | 4/9 | – |
| $z_B' = -2/3$ | | | 0 | 0 | −22/9 | 1 | 4/9 | 0 | 7/9 | |

| | | $c_j \rightarrow$ | 0 | 0 | 0 | 0 | 0 | −1 | 0 |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x_1'$ | $x_2'$ | $\lambda_1'$ | $\mu_1'$ | $\mu_2'$ | $A_1'$ | $s_1'$ |
| 0 | $\lambda_1$ | 3/4 | 0 | 0 | 1 | −9/22 | −2/11 | 9/22 | −7/22 |
| 0 | $x_2$ | 3/11 | 0 | 1 | 0 | −1 | −2 | 1 | −2 |
| 0 | $x_1$ | 4/11 | 1 | 0 | 0 | −1 | 2 | 1 | 9 |
| $z_B' = 0$ | | | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

The optimal solution is $x_1 = 4/11, x_2 = 3/11$ and Max $z' = 5/11$.

*Example 2* Solve the following problem by Wolfe's method:

$$\text{Minimize } z = -10x_1 - 25x_2 + 10x_1^2 + x_2^2 + 4x_1x_2$$

subject to

$$x_1 + 2x_2 \leq 10$$
$$x_1 + x_2 \leq 9$$
$$\text{and } x_1 \geq 0,\ x_2 \geq 0.$$

**Solution**

This is a minimization problem. Now converting the problem into a maximization problem and then converting all the inequality constraints, including the non-negativity constraints, into $\leq$ type, we get the reduced problem as follows:

$$\begin{aligned}
&\text{Maximize } z' = 10x_1 + 25x_2 - 10x_1^2 - x_2^2 - 4x_1x_2\\
&\text{subject to}\\
&\qquad\qquad x_1 + 2x_2 \leq 10\\
&\qquad\qquad x_1 + x_2 \leq 9\\
&\qquad\qquad -x_1 \leq 0, \quad -x_2 \leq 0\\
&\text{where } z' = -z.
\end{aligned}$$

Now introducing slack variables $q_1^2, q_2^2, r_1^2$ and $r_2^2$, we have the reduced problem as

$$\begin{aligned}
&\text{Maximize } z' = 10x_1 + 25x_2 - 10x_1^2 - x_2^2 - 4x_1x_2\\
&\text{subject to}\\
&\qquad\qquad x_1 + 2x_2 + q_1^2 = 10\\
&\qquad\qquad x_1 + x_2 + q_2^2 = 9\\
&\qquad\qquad -x_1 + r_1^2 = 0, -x_2 + r_2^2 = 0.
\end{aligned}$$

Now the Lagrange function $L$ is given by

$$\begin{aligned}
&L(x_1, x_2, q_1, q_2, r_1, r_2, \lambda_1, \lambda_2, \mu_1, \mu_2)\\
&\quad = 10x_1 + 25x_2 - 10x_1^2 - x_2^2 - 4x_1x_2 - \lambda_1\left(x_1 + 2x_2 + q_1^2 - 10\right)\\
&\qquad - \lambda_2\left(x_1 + x_2 + q_2^2 - 9\right)\\
&\qquad - \mu_1\left(-x_1 + r_1^2\right) - \mu_2\left(-x_2 + r_2^2\right).
\end{aligned}$$

Clearly,

$-10x_1^2 - x_2^2 - 4x_1x_2$ is a negative semi-definite quadratic expression.
Hence, $z' = 10x_1 + 25x_2 - 10x_1^2 - x_2^2 - 4x_1x_2$ is a concave function in $x_1, x_2$.
Hence, the maximum of $L$ will be the maximum of $z'$ and vice versa.
Now the necessary conditions for maxima of $L$ (and hence of $z'$) are given by

$$\frac{\partial L}{\partial x_1} = 0 \quad \text{or} \quad 10 - 20x_1 - 4x_2 - \lambda_1 - \lambda_2 + \mu_1 = 0 \tag{12.27}$$

$$\frac{\partial L}{\partial x_2} = 0 \quad \text{or} \quad 25 - 2x_2 - 4x_1 - 2\lambda_1 - \lambda_2 + \mu_2 = 0 \tag{12.28}$$

$$\frac{\partial L}{\partial \lambda_1} = 0 \quad \text{or} \;\; x_1 + 2x_2 + q_1^2 - 10 = 0 \tag{12.29}$$

$$\frac{\partial L}{\partial \lambda_2} = 0 \quad \text{or} \;\; x_1 + x_2 + q_2^2 - 9 = 0 \tag{12.30}$$

$$\frac{\partial L}{\partial \mu_1} = 0 \quad \text{or} \quad -x_1 + r_1^2 = 0 \tag{12.31}$$

$$\frac{\partial L}{\partial \mu_2} = 0 \quad \text{or} \quad -x_1 + r_2^2 = 0 \tag{12.32}$$

$$\frac{\partial L}{\partial q_1} = 0 \quad \text{or} \;\; \lambda_1 q_1 = 0 \tag{12.33}$$

$$\frac{\partial L}{\partial q_2} = 0 \quad \text{or} \;\; \lambda_2 q_2 = 0 \tag{12.34}$$

$$\frac{\partial L}{\partial r_1} = 0 \quad \text{or} \;\; \mu_1 r_1 = 0 \text{ or } \mu_1 x_1 = 0 \text{ as } -x_1 + r_1^2 = 0 \tag{12.35}$$

$$\frac{\partial L}{\partial r_2} = 0 \quad \text{or} \;\; \mu_2 r_2 = 0 \text{ or } \mu_2 x_2 = 0 \text{ as } -x_2 + r_2^2 = 0 \tag{12.36}$$

Now letting $q_1^2 = s_1 \geq 0$, we have from (12.29)

$$x_1 + 2x_2 + s_1 - 10 = 0. \tag{12.37}$$

From (12.33),

$$\lambda_1 s_1 = 0. \tag{12.38}$$

Let $q_2^2 = s_2 \geq 0$; then we have from (12.30) and (12.34)

$$x_1 + x_2 + s_2 - 9 = 0 \tag{12.39}$$

$$\lambda_2 s_2 = 0 \tag{12.40}$$

where $x_1, x_2, s_1, s_2, \mu_1, \mu_2, \lambda_1, \lambda_2 \geq 0$.

Now the equivalent problem of the given QPP is as follows:

$$\left.\begin{aligned} 20x_1 + 4x_2 + \lambda_1 + \lambda_2 - \mu_1 &= 10 \\ 4x_1 + 2x_2 + 2\lambda_1 + \lambda_2 - \mu_2 &= 25 \\ x_1 + 2x_2 + s_1 &= 10 \\ x_1 + x_2 + s_2 &= \;\; 9 \end{aligned}\right\} \tag{12.41}$$

$$
\begin{aligned}
&x_1, x_2 \geq 0, \quad s_1, s_2 \geq 0, \quad \lambda_1, \lambda_2 \geq 0, \quad \mu_1, \mu_2 \geq 0 \\
&\lambda_1 s_1 = 0, \quad \lambda_2 s_2 = 0, \quad \mu_1 x_1 = 0, \quad \mu_2 x_2 = 0.
\end{aligned} \tag{12.42}
$$

A solution $x_j,\ j = 1, 2$ of (12.41) satisfying (12.42) shall necessarily be an optimal one for maximizing $L$ (i.e. maximizing $z$).

Now to determine the solution to the simultaneous Eqs. (12.41), we introduce two artificial variables $A_1(\geq 0)$ and $A_2(\geq 0)$ in the first two equations of (12.41) and construct the dummy objective function $z'' = -A_1 - A_2$.

Then the equivalent problem of the given QPP becomes

$$
\begin{aligned}
&\text{Maximize } z'' = -A_1 - A_2 \\
&\text{subject to} \\
&\qquad 20x_1 + 4x_2 + \lambda_1 + \lambda_2 - \mu_1 + A_1 = 10 \\
&\qquad 4x_1 + 2x_2 + 2\lambda_1 + \lambda_2 - \mu_2 + A_2 = 25 \\
&\qquad x_1 + 2x_2 + s_1 = 10 \\
&\qquad x_1 + x_2 + s_2 = 9 \\
&x_1, x_2 \geq 0, \quad \lambda_1, \lambda_2 \geq 0, \quad \mu_1, \mu_2 \geq 0, \quad s_1, s_2 \geq 0, \quad A_1, A_2 \geq 0,
\end{aligned}
$$

satisfying the complementary slackness conditions

$$\lambda_1 s_1 = 0, \quad \lambda_2 s_2 = 0, \quad \mu_1 x_1 = 0, \quad \mu_2 x_2 = 0.$$

To solve the problem we shall apply the two-phase method.

Here the initial BFS is $A_1 = 10,\ A_2 = 25,\ s_1 = 10,\ s_2 = 9$.

Now we construct the initial table of Phase I.

| | | $c_j$ | 0 | 0 | 0 | 0 | 0 | 0 | −1 | −1 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x_1'$ | $x_2'$ | $\lambda_1'$ | $\lambda_2'$ | $\mu_1'$ | $\mu_2'$ | $A_1'$ | $A_2'$ | $s_1'$ | $s_2'$ | Mini ratio |
| −1 | $A_1$ | 10 | 20 | 4 | 1 | 1 | −1 | 0 | 1 | 0 | 0 | 0 | 1/2 |
| −1 | $A_2$ | 25 | 4 | 2 | 2 | 1 | 0 | −1 | 0 | 1 | 0 | 0 | 25/4 |
| 0 | $s_1$ | 10 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 10 |
| 0 | $s_2$ | 9 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 9 |
| $z_B'' = -35$ | | | −24 | −6 | −3 | −2 | 1 | 1 | 0 | 0 | 0 | 0 | $\leftarrow \Delta_j$ |

Here all $\Delta_j \not\geq 0$. Hence, the current BFS is not optimal. Here $x_1'$ is introduced into the basis as $\mu_1 = 0$ with $A_1'$ as the outgoing (leaving) vector.

Now we construct the next table, dropping the $A_1'$ column.

| | | $c_j$ | 0 | 0 | 0 | 0 | 0 | 0 | −1 | −1 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x_1'$ | $x_2'$ | $\lambda_1'$ | $\lambda_2'$ | $\mu_1'$ | $\mu_2'$ | $A_1'$ | $A_2'$ | $s_1'$ | $s_2'$ | Mini ratio |
| 0 | $x_1$ | 1/2 | 1 | 1/5 | 1/20 | 1/20 | −1/20 | 0 | 1/20 | 0 | 0 | 0 | 5/2 |
| −1 | $A_2$ | 23 | 0 | 6/5 | 9/5 | 4/5 | 1/5 | −1 | −1/5 | 1 | 0 | 0 | 115/6 |
| 0 | $s_1$ | 19/2 | 0 | 9/5 | −1/20 | −1/20 | 1/20 | 0 | −1/20 | 0 | 1 | 0 | 95/18 |
| 0 | $s_2$ | 17/2 | 0 | 4/5 | −1/20 | −1/20 | 1/20 | 0 | −1/20 | 0 | 0 | 1 | 85/8 |
| $z_B'' = -23$ | | | 0 | −6/5 | −9/5 | −4/5 | −1/5 | 1 | 6/5 | 0 | 0 | 0 | $\leftarrow \Delta_j$ |

Here all $\Delta_j \ngeq 0$. Hence, the current BFS is not optimal. Again, $\Delta_j$ corresponding to $\lambda_1'$ is most negative, but it does not enter into the basis, since by the complementary condition, $s_1'$ is already in the basis.

So the entering vector is $x_2'$, as $\mu_2'$ is not present in the basis and $x_1'$ is the leaving vector.

| | | $c_j$ | 0 | 0 | 0 | 0 | 0 | 0 | −1 | −1 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x_1'$ | $x_2'$ | $\lambda_1'$ | $\lambda_2'$ | $\mu_1'$ | $\mu_2'$ | $A_1'$ | $A_2'$ | $s_1'$ | $s_2'$ | Mini ratio |
| 0 | $x_2$ | 5/2 | 5 | 1 | 1/4 | 1/4 | −1/4 | 0 | 1/4 | 0 | 0 | 0 | |
| −1 | $A_2$ | 20 | −6 | 0 | 3/2 | 1/2 | 1/2 | −1 | −1/2 | 1 | 0 | 0 | 40 |
| 0 | $s_1$ | 5 | −9 | 0 | −1/2 | −1/2 | 1/2 | 0 | −1/2 | 0 | 1 | 0 | 10 |
| 0 | $s_2$ | 13/2 | −4 | 0 | −1/4 | −1/4 | 1/4 | 0 | −1/4 | 0 | 0 | 1 | 26 |
| $z_B'' = -20$ | | | 6 | 0 | −3/2 | −1/2 | −1/2 | 1 | 3/2 | 0 | 0 | 0 | $\leftarrow \Delta_j$ |

Here all $\Delta_j \ngeq 0$. Hence, the current BFS is not optimal. $\Delta_j$ corresponding to $\lambda_1'$ is most negative, but it does not enter into the basis, as by the complementary condition, $s_1'$ is already in the basis.

Now $\lambda_2'$ does not enter into the basis, as by the complementary condition, $s_2'$ is already in the basis.

Now $\mu_1'$ is the entering vector, since by the complementary condition, $\mu_1 x_1 = 0$ and $s_1'$ is the leaving vector.

| | | $c_j$ | 0 | 0 | 0 | 0 | 0 | 0 | −1 | −1 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x_1'$ | $x_2'$ | $\lambda_1'$ | $\lambda_2'$ | $\mu_1'$ | $\mu_2'$ | $A_1'$ | $A_2'$ | $s_1'$ | $s_2'$ | Mini ratio |
| 0 | $x_2$ | 5 | 1/2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1/2 | 0 | |

(continued)

(continued)

| | | $c_j$ | 0 | 0 | 0 | 0 | 0 | 0 | −1 | −1 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| −1 | $A_2$ | 15 | 3 | 0 | 2 | 1 | 0 | −1 | 0 | 1 | −1 | 0 | 15/2 |
| 0 | $\mu_1$ | 10 | −18 | 0 | −1 | −1 | 1 | 0 | 0 | 0 | 2 | 0 | |
| 0 | $s_2$ | 4 | 1/2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | −1/2 | 1 | |
| $z''_B = -15$ | | | −3 | 0 | −2 | −1 | 0 | 1 | | 0 | 1 | 0 | $\leftarrow \Delta_j$ |

Here all $\Delta_j \ngeq 0$. Hence, the BFS is not optimal. From the table we observe that $x'_1$ is most negative, but it does not enter into the basis, since by the complementary condition, $\mu'_1$ is already in the basis. So by the complementary condition, $\lambda'_1$ is entering the basis and $A'_2$ is the leaving vector.

| | | $c_j$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x'_1$ | $x'_2$ | $\lambda'_1$ | $\lambda'_2$ | $\mu'_1$ | $\mu'_2$ | $s'_1$ | $s'_2$ | Mini ratio |
| 0 | $x_2$ | 5 | 1/2 | 1 | 0 | 0 | 0 | 0 | 1/2 | 0 | |
| 0 | $\lambda_1$ | 15/2 | 3/2 | 0 | 1 | 1/2 | 0 | −1/2 | −1/2 | 0 | |
| 0 | $\mu_1$ | 35/2 | 33/2 | 0 | 0 | −1/2 | 1 | −1/2 | 3/2 | 0 | |
| 0 | $s_2$ | 4 | 1/2 | 0 | 0 | 0 | 0 | 0 | −1/2 | 1 | |
| $z''_B = 0$ | | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $\leftarrow \Delta_j$ |

Since $\Delta_j \geq 0$, the current BFS is optimal. The optimal solution is $x_1 = 0,\ x_2 = 5,$

$$\lambda_1 = \frac{15}{2}, \quad \lambda_2 = 0, \quad \mu_1 = \frac{35}{2}, \quad \mu_2 = 0, \quad s_1 = 0, \quad s_2 = 4.$$

This solution satisfies the complementary slackness conditions, $\lambda_1 s_1 = 0 = \lambda_2 s_2$ and $\mu_1 x_1 = 0 = \mu_2 x_2$.

Since $z''_B = 0$, the current solution is also feasible,

$$\begin{aligned}\text{and} \quad \text{Minimize } z &= -10x_1 - 25x_2 + 10x_1^2 + x_2^2 + 4x_1x_2 \\ &= -10.0 - 25.5 + 10.0 + 5^2 + 4.0 \\ &= -125 + 25 \\ &= -100.\end{aligned}$$

*Example 3* Apply Wolfe's method for solving the following QPP:

$$\text{Maximize } z = 4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2$$
subject to
$$x_1 + 2x_2 \leq 2 \text{ and } x_1, x_2 \geq 0.$$

**Solution**

The given problem is a maximization problem. Now, converting all the inequality constraints including the non-negativity constraints, into $\leq$ type, we have

$$x_1 + 2x_2 \leq 2, \quad -x_1 \leq 0 \text{ and } -x_2 \leq 0.$$

Now introducing slack variables $q_1^2, r_1^2$ and $r_2^2$, we have the reduced problem as

$$\text{Maximize } z = 4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2$$
subject to
$$\begin{aligned} &x_1 + 2x_2 + q_1^2 = 2 \\ &-x_1 + r_1^2 = 0 \\ &-x_2 + r_2^2 = 0 \end{aligned}$$

Now, the Lagrangian function $L$ is given by

$$\begin{aligned} L = L(x_1, x_2, q_1, \mu_1, \mu_2, \lambda_1, r_1, r_2) &= 4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2 \\ &- \lambda_1\left(x_1 + 2x_2 + q_1^2 - 2\right) - \mu_1\left(-x_1 + r_1^2\right) - \mu_2\left(-x_2 + r_2^2\right). \end{aligned}$$

As $-2x_1^2 - 2x_1x_2 - 2x_2^2$ represents a negative semi-definite quadratic form, then $z = 4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2$ is concave in $x_1$, $x_2$. Therefore, the maxima of $L$ will be the maxima of $z$ and vice versa.

Now, the necessary conditions for the maxima of $L$ (and hence of $z$) are

$$\frac{\partial L}{\partial x_1} = 0 \text{ or } 4 - 4x_1 - 2x_2 - \lambda_1 + \mu_1 = 0$$

$$\frac{\partial L}{\partial x_2} = 0 \text{ or } 6 - 2x_1 - 4x_2 - 2\lambda_1 + \mu_2 = 0$$

$$\frac{\partial L}{\partial \lambda_1} = 0 \text{ or } x_1 + 2x_2 + q_1^2 = 2$$

$$\frac{\partial L}{\partial \mu_1} = 0 \text{ or } -x_1 + r_1^2 = 0$$

$$\frac{\partial L}{\partial \mu_2} = 0 \text{ or } -x_2 + r_2^2 = 0$$

$$\frac{\partial L}{\partial q_1} = 0 \text{ or } \lambda_1 q_1 = 0 \Rightarrow \lambda_1 q_2^2 = 0$$

$$\frac{\partial L}{\partial r_1} = 0 \text{ or } \mu_1 r_1 = 0 \text{ or } \mu_1 r_1^2 = 0 \text{ or } \mu_1 x_1 = 0 \text{ as } -x_1 + r_1^2 = 0$$

$$\frac{\partial L}{\partial r_2} = 0 \text{ or } \mu_2 r_2 = 0 \text{ or } \mu_2 r_2^2 = 0 \text{ or } \mu_2 x_2 = 0 \text{ as } -x_2 + r_2^2 = 0.$$

Now, letting $q_1^2 = s_1 \geq 0$, we have $\lambda_1 s_1 = 0$ as $\lambda_1 q_2^2 = 0$.

Also, $x_1 + 2x_2 + s_1 = 2$ and finally, $x_1, x_2, s_1, \lambda_1, \mu_1, \mu_2 \geq 0$.

Hence, to determine all the variables, the equivalent problem of the given QPP is as follows:

$$\left.\begin{aligned} 4x_1 + 2x_2 + \lambda_1 - \mu_1 &= 4 \\ 2x_1 + 4x_2 + 2\lambda_1 - \mu_2 &= 6 \\ x_1 + 2x_2 + s_1 &= 2 \end{aligned}\right\}, \tag{12.43}$$

where all variables are non-negative and

$$\lambda_1 s_1 = 0,\ \mu_1 x_1 = 0\ \mu_2 x_2 = 0. \tag{12.44}$$

A solution $x_j$, $j = 1, 2$, of (12.43) and satisfying (12.44) shall necessarily be an optimal one for maximizing $L$ (i.e. maximizing $z$). Now to determine the solution to the simultaneous equations in (12.43), we introduce two artificial variables $A_1(\geq 0)$ and $A_2(\geq 0)$ in the first two equations of (12.43) and construct the dummy objective function $z' = -A_1 - A_2$.

Then the equivalent problem of the given QPP becomes

$$\begin{aligned} &\text{Maximize } z' = -A_1 - A_2 \\ &\text{subject to} \\ &\qquad 4x_1 + 2x_2 + \lambda_1 - \mu_1 + A_1 = 4 \\ &\qquad 2x_1 + 4x_2 + 2\lambda_1 - \mu_2 + A_2 = 6 \\ &\qquad x_1 + 2x_2 + s_1 = 2 \\ &x_1, x_2 \geq 0,\ \lambda_1 \geq 0,\ \mu_1, \mu_2, s_1, A_1, A_2 \geq 0, \end{aligned}$$

satisfying the complementary slackness conditions

$$\lambda_1 s_1 = 0, \quad \mu_1 x_1 = \mu_2 x_2 = 0.$$

To solve the problem, we shall apply the two-phase method.

Here the initial BFS is $A_1 = 4,\ A_2 = 6,\ s_1 = 0$.

Now we construct the initial table of Phase I.

| | | $c_j$ | 0 | 0 | 0 | 0 | 0 | −1 | −1 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x'_1$ | $x'_2$ | $\lambda'_1$ | $\mu'_1$ | $\mu'_2$ | $A'_1$ | $A'_2$ | $s'_1$ | Mini ratio |
| −1 | $A_1$ | 4 | 4 | 2 | 1 | −1 | 0 | 1 | 0 | 0 | 4/4 = 1 |
| −1 | $A_2$ | 6 | 2 | 4 | 2 | 0 | −1 | 0 | 1 | 0 | 6/2 = 3 |
| 0 | $s_1$ | 2 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 2/1 = 2 |
| $z'_B = -10$ | | | −6 | −6 | −3 | 1 | 1 | 0 | 0 | 0 | $\leftarrow \Delta_j$ |

Here all $\Delta_j \ngeq 0$. Hence, the current BFS is not optimal.

From the preceding table, we observe that either $x_1$ or $x_2$ (as $\Delta_1 = -6,\ \Delta_2 = -6$) can be introduced into the basic solution. Arbitrarily, we choose $x_1$ as the entering basic variable, as $\mu_1 = 0$ (for satisfying $\mu_1 x_1 = 0$). Therefore, $A_1$ is the leaving basic variable. Now we construct the next table, dropping the $A_1$ column.

| | | $c_j$ | 0 | 0 | 0 | 0 | 0 | −1 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x'_1$ | $x'_2$ | $\lambda'_1$ | $\mu'_1$ | $\mu'_2$ | $A'_2$ | $s'_1$ | Mini Ratio |
| 0 | $x_1$ | 1 | 1 | ½ | ¼ | −¼ | 0 | 0 | 0 | 2 |
| −1 | $A_2$ | 4 | 0 | 3 | 3/2 | ½ | −1 | 1 | 0 | 4/3 |
| 0 | $s_1$ | 1 | 0 | 3/2 | −¼ | ¼ | 0 | 0 | 1 | 2/3 |
| $Z'_B = -4$ | | | 0 | −3 | −3/2 | −½ | 1 | 0 | 0 | $\leftarrow \Delta_j$ |

Since $\mu_2 = 0$, $x_2$ is introduced as the basic variable with $s_1$ as the leaving basic variable. Now, we construct the next table.

| | | $c_j$ | 0 | 0 | 0 | 0 | 0 | −1 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x'_1$ | $x'_2$ | $\lambda'_1$ | $\mu'_1$ | $\mu'_2$ | $A'_2$ | $s'_1$ | Mini Ratio |
| 0 | $x_1$ | 1 | 1 | ½ | ¼ | −¼ | 0 | 0 | 0 | 2 |
| −1 | $A_2$ | 4 | 0 | 3 | 3/2 | ½ | −1 | 1 | 0 | 4/3 |
| 0 | $s_1$ | 1 | 0 | 3/2 | −¼ | ¼ | 0 | 0 | 1 | 2/3 |
| $Z'_B = -4$ | | | 0 | −3 | −3/2 | −½ | 1 | 0 | 0 | $\leftarrow \Delta_j$ |

Since $s_1 = 0$, $\lambda_1$ is introduced as the basic variable. $A_2$ is the leaving basic variable. Now we construct the next table, dropping the $A_2$ column.

| | | $c_j$ | 0 | 0 | 0 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x_1'$ | $x_2'$ | $\lambda_1'$ | $\mu_1'$ | $\mu_2'$ | $s_1'$ | |
| 0 | $x_1$ | 1/3 | 1 | 0 | 0 | −1/3 | 1/6 | 0 | |
| 0 | $\lambda_1$ | 1 | 0 | 0 | 1 | 0 | −1/2 | −1 | |
| 0 | $x_2$ | 5/6 | 0 | 1 | 0 | 1/6 | −1/12 | ½ | |
| $Z_B' = -2$ | | | 0 | 0 | 0 | 0 | 0 | 0 | $\leftarrow \Delta_j$ |

Here all $\Delta_j \ngeq 0$. Hence, the optimal solution for Phase I has been obtained, and the solution is $x_1 = \frac{1}{3}$, $x_2 = \frac{5}{6}$, $\lambda_1 = 1, \mu_1 = 0, \mu_2 = 0, s_1 = 0$. Thus, the optimal solution of the original QPP is given by

$$x_1 = \frac{1}{3}, \quad x_2 = \frac{5}{6} \quad \text{and} \quad z_{\max} = 4\left(\frac{1}{3}\right) + \left(\frac{5}{6}\right) - 2\left(\frac{1}{3}\right)^2 - 2\left(\frac{1}{3}\right)\left(\frac{5}{6}\right) - 2\left(\frac{5}{6}\right)^2$$
$$= \frac{25}{6}.$$

*Example 4* Apply Wolfe's method for solving the following problem:

$$\begin{aligned} &\text{Maximize } z = 2x_1 + x_2 - x_1^2 \\ &\text{subject to} \\ &\qquad 2x_1 + 3x_2 \le 6 \\ &\qquad 2x_1 + x_2 \le 4 \ \text{ and } \ x_1, x_2 \ge 0. \end{aligned}$$

**Solution**

The given problem is a maximization problem. Now, converting all the inequality constraints, including the non-negativity constraints, into $\le$ type, we have

$$\begin{aligned} &2x_1 + 3x_2 \le 6 \\ &2x_1 + x_2 \le 4 \\ &-x_1 \le 0, -x_2 \le 0. \end{aligned}$$

Now introducing slack variables $q_1^2, r_1^2$ and $r_2^2$, we have the reduced problem as

$$\text{Maximize } z = 2x_1 + x_2 - x_1^2$$
$$2x_1 + 3x_2 + q_1^2 = 6$$
$$2x_1 + x_2 + q_2^2 = 4$$
$$-x_1 + r_1^2 = 0$$
$$-x_2 + r_2^2 = 0$$

The Lagrangian function $L$ is given by

$$\begin{aligned} &L(x_1, x_2, q_1, q_2, r_1, r_2, \lambda_1, \lambda_2, \mu_1, \mu_2) \\ &\quad = \left(2x_1 + x_2 - x_1^2\right) - \lambda_1\left(2x_1 + 3x_2 + q_1^2 - 6\right) - \lambda_2\left(2x_1 + x_2 + q_2^2 - 4\right) \\ &\qquad - \mu_1\left(-x_1 + r_1^2\right) - \mu_2\left(-x_2 + r_2^2\right). \end{aligned}$$

As $-x_1^2$ represents a negative semi-definite quadratic form, $z = 2x_1 + x_2 - x_1^2$ is concave in $x_1$, $x_2$. Hence, the maxima of $L$ will be the maxima of $z$ and vice versa.

Now, the necessary conditions for maxima of $L$ (and hence of $z$) are

$$\frac{\partial L}{\partial x_1} = 2 - 2x_1 - 2\lambda_1 - 2\lambda_2 + \mu_1 = 0 \text{ or } 2x_1 + 2\lambda_1 + 2\lambda_2 - \mu_1 = 2$$

$$\frac{\partial L}{\partial x_2} = 1 - 3\lambda_1 - \lambda_2 + \mu_2 = 0 \text{ or } 3\lambda_1 + \lambda_2 - \mu_2 = 1$$

$$\frac{\partial L}{\partial \lambda_1} = -\left(2x_1 + 3x_2 + q_1^2 - 6\right) = 0 \text{ or } 2x_1 + 3x_2 + s_1 = 6 \;\left[\text{putting } q_1^2 = s_1\right]$$

$$\frac{\partial L}{\partial \lambda_2} = -\left(2x_1 + x_2 + q_2^2 - 4\right) = 0 \text{ or } 2x_1 + x_2 + s_2 = 4 \;\left[\text{putting } q_2^2 = s_2\right]$$

$$\frac{\partial L}{\partial q_1} = -2\lambda_1 q_1 = 0 \text{ or } \lambda_1 q_1^2 = 0 \text{ or } \lambda_1 s_1 = 0$$

$$\frac{\partial L}{\partial q_2} = -2\lambda_2 q_2 = 0 \text{ or } \lambda_2 q_2^2 = 0 \text{ or } \lambda_2 s_2 = 0$$

$$\frac{\partial L}{\partial \mu_1} = -\left(-x_1 + r_1^2\right) = 0 \text{ or } x_1 = r_1^2$$

$$\frac{\partial L}{\partial \mu_2} = -\left(-x_2 + r_2^2\right) = 0 \text{ or } x_2 = r_2^2$$

$$\frac{\partial L}{\partial r_1} = -2\mu_1 r_1 = 0 \text{ or } \mu_1 r_1 = 0 \text{ or } \mu_1 r_1^2 = 0 \text{ or } \mu_1 x_1 = 0$$

$$\frac{\partial L}{\partial r_2} = -2\mu_2 r_2 = 0 \text{ or } \mu_2 r_2 = 0 \text{ or } \mu_2 r_2^2 = 0 \text{ or } \mu_2 x_2 = 0$$

$$\text{and } x_1, x_2 \geq 0,\ \lambda_1, \lambda_2 \geq 0,\ \mu_1, \mu_2, s_1, s_2 \geq 0.$$

Hence, to determine all the variables, the equivalent problem of the given QPP is as follows:

$$\left.\begin{aligned} 2x_1 + 2\lambda_1 + 2\lambda_2 - \mu_1 &= 2 \\ 3\lambda_1 + \lambda_2 - \mu_2 &= 1 \\ 2x_1 + 3x_2 + s_1 &= 6 \\ 2x_1 + x_2 + s_2 &= 4 \end{aligned}\right\}, \tag{12.45}$$

where all variables are non-negative and

$$\lambda_1 s_1 = 0, \quad \lambda_2 s_2 = 0, \quad \mu_1 x_1 = 0, \quad \mu_2 x_2 = 0. \tag{12.46}$$

A solution $x_j$, $j = 1, 2$, of (12.45) and satisfying (12.46) shall necessarily be an optimal one for maximizing $L$ (i.e. maximizing $z$). Now to determine the solution to the simultaneous Eq. (12.45), we introduce two artificial variables $A_1(\geq 0)$ and $A_2(\geq 0)$ in the first two equations of (12.45) and construct the dummy objective function $z' = -A_1 - A_2$.

Then the equivalent problem of the given QPP becomes

$$\begin{aligned} &\text{Maximize } z' = -A_1 - A_2 \\ &\text{subject to} \\ &\qquad 2x_1 + 2\lambda_1 + 2\lambda_2 - \mu_1 + A_1 = 2 \\ &\qquad 3\lambda_1 + \lambda_2 - \mu_2 + A_2 = 1 \\ &\qquad 2x_1 + 3x_2 + s_1 = 6 \\ &\qquad 2x_1 + x_2 + s_2 = 4 \\ &\text{and } x_1, x_2 \geq 0,\ \lambda_1, \lambda_2 \geq 0,\ \mu_1, \mu_2, s_1, s_2, A_1, A_2 \geq 0, \end{aligned}$$

satisfying the complementary slackness conditions

$$\lambda_1 s_1 = 0, \quad \lambda_2 s_2 = 0, \quad \mu_1 x_1 = 0, \quad \mu_2 x_2 = 0.$$

To solve the problem, we shall apply the two-phase method.

Here the initial BFS is $A_1 = 2,\ A_2 = 1,\ s_1 = 6, s_2 = 4$.

Now we construct the initial table of Phase I.

| | $c_j$ | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | −1 | −1 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x_1'$ | $x_2'$ | $\lambda_1'$ | $\lambda_2'$ | $\mu_1'$ | $\mu_2'$ | $s_1'$ | $s_2'$ | $A_1'$ | $A_2'$ | Mini ratio |
| −1 | $A_1$ | 2 | 2 | 0 | 2 | 2 | −1 | 0 | 0 | 0 | 1 | 0 | 1 |
| −1 | $A_2$ | 1 | 0 | 0 | 3 | 1 | 0 | −1 | 0 | 0 | 0 | 1 | – |
| 0 | $s_1$ | 6 | 2 | 3 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 3 |
| 0 | $s_2$ | 4 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 2 |
| $z_B' = -3$ | | | −2 | 0 | −5 | −3 | 1 | 1 | 0 | 0 | 0 | 0 | $\leftarrow \Delta_j$ |

Here all $\Delta_j \ngeq 0$. Hence, the current BFS is not optimal. $\Delta_j$ corresponding to $\lambda_1'$ is most negative and corresponding to $\lambda_2'$ is second most negative, but the corresponding variables $\lambda_1, \lambda_2$ will not be considered as basic variables, since $s_1$ and $s_2$ are basic variables. So the entering basic variable will be $x_1$, as $\mu_1$ is not a basic variable and $A_1$ is the leaving basic variable. Now we construct the next table, dropping the $A_1'$ column.

| | $c_j$ | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | −1 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x_1'$ | $x_2'$ | $\lambda_1'$ | $\lambda_2'$ | $\mu_1'$ | $\mu_2'$ | $s_1'$ | $s_2'$ | $A_2'$ | Mini ratio |
| 0 | $x_1$ | 1 | 1 | 0 | 1 | 1 | $-\frac{1}{2}$ | 0 | 0 | 0 | 0 | – |
| −1 | $A_2$ | 1 | 0 | 0 | 3 | 1 | 0 | −1 | 0 | 0 | 1 | – |
| 0 | $s_1$ | 4 | 0 | 3 | −2 | −2 | 1 | 0 | 1 | 0 | 0 | 4/3 |
| 0 | $s_2$ | 2 | 0 | 1 | −2 | −2 | 1 | 0 | 0 | 1 | 0 | 2 |
| $z_B' = -1$ | | | 0 | 0 | −3 | −1 | 0 | 1 | 0 | 0 | 0 | $\leftarrow \Delta_j$ |

Here all $\Delta_j \ngeq 0$. Hence, the current BFS is not optimal. From this table, we observe that $\Delta_1$ corresponding to $\lambda_1$ is most negative and $\lambda_2$ is second most negative. However, the corresponding variables $\lambda_1, \lambda_2$ will not be considered as basic variables, as $s_1$ and $s_2$ are already basic variables. So the basic variable will be $x_2$ (as $\mu_2$ is not the basic variable) and $s_1$ is the leaving basic variable. Now, we construct the next table.

| | $c_j \rightarrow$ | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | −1 | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x_1'$ | $x_2'$ | $\lambda_1'$ | $\lambda_2'$ | $\mu_1'$ | $\mu_2'$ | $s_1'$ | $s_2'$ | $A_2'$ | |
| 0 | $x_1$ | 1 | 1 | 0 | 1 | 1 | $-\frac{1}{2}$ | 0 | 0 | 0 | 0 | 1 |
| −1 | $A_2$ | 1 | 0 | 0 | 3 | 1 | 0 | −1 | 0 | 0 | 1 | 1/3 |

(continued)

(continued)

| | $c_j \rightarrow$ | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | −1 | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x'_1$ | $x'_2$ | $\lambda'_1$ | $\lambda'_2$ | $\mu'_1$ | $\mu'_2$ | $s'_1$ | $s'_2$ | $A'_2$ | |
| 0 | $x_2$ | 4/3 | 0 | 1 | $-\frac{2}{3}$ | $-\frac{2}{3}$ | 1/3 | 0 | 1/3 | 0 | 0 | – |
| 0 | $s_2$ | 2/3 | 0 | 0 | $-\frac{4}{3}$ | $-\frac{4}{3}$ | 2/3 | 0 | $-\frac{1}{3}$ | 1 | 0 | – |
| $z'_B = -1$ | | | 0 | 0 | −3 | −1 | 0 | 1 | 0 | 0 | 0 | $\leftarrow \Delta_j$ |

Here all $\Delta_j \ngeq 0$. Hence, the current solution is not optimal.

From the preceding table, it is observed that $\Delta_j$ corresponding to $\lambda_1$ is most negative. Hence, $\lambda_1$ is the new basic variable, since $s_1$ is not in the basis and $A_2$ is the leaving basic variable.

| | $c_j$ | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | Basic variables | $x_B$ | $x'_1$ | $x'_2$ | $\lambda'_1$ | $\lambda'_2$ | $\mu'_1$ | $\mu'_2$ | $s'_1$ | $s'_2$ | |
| 0 | $x_1$ | 2/3 | 1 | 0 | 0 | 2/3 | $-\frac{1}{2}$ | 1/3 | 0 | 0 | |
| 0 | $\lambda_1$ | 1/3 | 0 | 0 | 1 | 1/3 | 0 | $-\frac{1}{3}$ | 0 | 0 | |
| 0 | $x_2$ | 14/9 | 0 | 1 | 0 | $-\frac{4}{9}$ | 1/3 | $-\frac{2}{9}$ | 1/3 | 0 | |
| 0 | $s_2$ | 10/9 | 0 | 0 | 0 | $-\frac{8}{9}$ | 2/3 | $-\frac{4}{9}$ | $-\frac{1}{3}$ | 1 | |
| $z'_B = 0$ | | | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $\leftarrow \Delta_j$ |

From the table, it is clear that $\Delta_j \geq 0$. Hence, the current solution is optimal. The optimal solution is $x_1 = \frac{2}{3}, x_2 = \frac{14}{9}, \lambda_1 = \frac{1}{3}, \lambda_2 = 0, \mu_1 = 0, \mu_2 = 0, s_1 = 0, s_2 = \frac{10}{9}$. This solution also satisfies the complementary slackness conditions

$$\lambda_1 s_1 = 0 = \lambda_2 s_2 \text{ and } \mu_1 x_1 = 0 = \mu_2 x_2.$$

Again, since $z_B = 0$, the current solution is also feasible.

Now, Max $z = \frac{2.2}{3} + \frac{14}{9} - \frac{4}{9} = \frac{12+14-4}{9} = \frac{22}{9}$.

Therefore, the optimal solution of the original QPP is given by $x_1 = \frac{2}{3}, x_2 = \frac{14}{9}$ and $z_{\max} = \frac{22}{9}$.

## 12.6 Beale's Method

In the year 1959, Beale developed an alternative technique for solving QPPs without using Kuhn-Tucker conditions. In this technique, the classical calculus results have been used to partition the variables into basic and non-basic ones. In each iteration, all the basic variables and the objective function are expressed in terms of non-basic variables.

## 12.7 Beale's Algorithm for QPPs

Let the general QPP be the following:

$$\text{Maximize } z = \langle c, x\rangle + \frac{1}{2}\langle Qx, x\rangle$$

subject to the constraints

$$Ax \leq \text{or} \geq \text{or} = b \ \text{ and } \ x \geq 0,$$

where $x^{\mathrm{T}} \in R^n$, $A$ is an $m \times n$ matrix, $b$ is an $m \times 1$ matrix and $Q$ is an $n \times n$ symmetric matrix.

Beale's iterative procedure for solving this problem can be summarized as follows:

Step 1. Convert the minimization problem into a maximization problem, if required. Introduce slack and/or surplus variables in the inequality constraints, if any, and put the QPP in the standard form.

Step 2. Select arbitrarily $m$ number of variables as basic variables; the remaining variables are considered as non-basic variables. Let the basic variables be denoted by $x_{Bi}$ ($i$ = 1, 2, …, $m$) and the non-basic variables as $x_{NB_k}, k = 1, 2, \ldots, n - m$.

Step 3. Express each basic variable $x_{B_i}$ entirely in terms of the non-basic variables, i.e. the $x_{NB_i}$'s (and $u_i$'s, if any), using the given constraints (as well as any additional ones).

Step 4. Express the objective function $z$, also in terms of the non-basic variable $x_{NB_i}$'s (and $u_i$'s, if any).

Step 5. Examine the partial derivatives of $z$ with respect to the non-basic variables at the point $x_{NB} = 0$ (and $u$ = 0):

(i) If $\left.\frac{\partial z}{\partial x_{NB_k}}\right|_{\substack{x_{NB} = 0 \\ u = 0}} \leq 0$ for each $k = 1, 2, \ldots, n - m$

and $\left.\frac{\partial z}{\partial u_i}\right|_{\substack{x_{NB} = 0 \\ u = 0}} = 0$ for each $i$ = 1,2, …, $m$,

the current basic solution is optimal. Go to Step 7.

(ii) If $\left.\frac{\partial z}{\partial x_{NB_k}}\right|_{\substack{x_{NB} = 0 \\ u = 0}} > 0$ for each least one $k$, choose the most positive one. The corresponding non-basic variables will enter the basis.

(iii) If $\left.\frac{\partial z}{\partial x_{NB_k}}\right|_{\substack{x_{NB}=0\\u=0}} < 0$ for each $k = 1, 2, \ldots, n-m$

but $\left.\frac{\partial z}{\partial u_i}\right|_{\substack{x_{NB}=0\\u=0}} \neq 0$ for some $i = r$,

introduce a new non-basic variable $u_j$, defined by $u_j = \frac{\partial z}{\partial u_r}$, and treat $u_r$ as a basic variable. Go to Step 3.

Step 6. Let $x_{NB_i} = x_k$ be the entering variable identified in Step 5 (ii). Compute the largest non-negative values of $x_k$ from the non-negativity conditions of each basic variable and the value of $x_k$ from $\frac{\partial z}{\partial x_k} = 0$ using other non-basic variables as zero.

Step 7. Find the minimum value of $x_k$.

(i) If the minimum value of $x_k$ occurs from the non-negative condition of a basic variable, then that variable will be the leaving variable from the basis and $x_k$ will be the new basic variable.

(ii) If the minimum value of $x_k$ occurs from $\frac{\partial z}{\partial x_k} = 0$, introduce an additional non-basic variable, called a free variable, defined by $u_i = \frac{\partial z}{\partial x_k}$ ($u_i$ is unrestricted).
This relation becomes an additional constraint equation.

Step 8. Go to Step 3 and repeat the procedure until an optimal basic solution is reached.

Step 9. Determine the optimal values of $x_B$ and $z$ by setting $x_{NB} = 0$ in their expressions obtained in Step 3 and Step 4.

Step 10. Stop.

*Example-5* Use Beale's method to solve the following QPP:

$$\text{Maximize } z = 4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2$$
$$\text{subject to } x_1 + 2x_2 \leq 2 \text{ and } x_1, x_2 \geq 0.$$

**Solution** The given problem is

$$\text{Maximize } z = 4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2$$
$$\text{subject to } x_1 + 2x_2 \leq 2 \text{ and } x_1, x_2 \geq 0.$$

Now introducing slack variable $x_3 \geq 0$ in the given constraint, the given problem reduces to

$$\text{Maximize } z = 4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2$$
$$\text{subject to } x_1 + 2x_2 + x_3 = 2 \text{ and } x_1, x_2, x_3 \geq 0.$$

Let us consider $x_2$ as basic variables $x_B$ and the remaining two variables $x_1, x_3$ as non-basic variables $x_{NB}$.

Now expressing $x_B$ and $z$ in terms of $x_{NB}$, we have

$$x_2 = 1 - \frac{1}{2}(x_1 + x_3)$$

$$\text{and } z = 4x_1 + 6\left(1 - \frac{x_1}{2} - \frac{x_3}{2}\right) - 2x_1^2 - 2x_1\left\{1 - \frac{1}{2}(x_1 + x_3)\right\} - 2\left\{1 - \frac{1}{2}(x_1 + x_3)\right\}^2$$

$$\therefore \frac{\partial z}{\partial x_1} = 4 - \frac{6}{2} - 4x_1 - 2x_1\left(-\frac{1}{2}\right) - 2\left\{1 - \frac{1}{2}(x_1 + x_3)\right\} - 4\left\{1 - \frac{1}{2}(x_1 + x_3)\right\}\left(-\frac{1}{2}\right)$$
$$= 4 - 3 - 4x_1 + x_1 - 2 + x_1 + x_3 + 2 - x_1 - x_3$$
$$= 1 - 3x_1$$

$$\frac{\partial z}{\partial x_3} = 0 - \frac{6}{2} - 2x_1\left(-\frac{1}{2}\right) - 4\left\{1 - \frac{1}{2}(x_1 + x_3)\right\}\left(-\frac{1}{2}\right)$$
$$= -3 + x_1 + 2 - x_1 - x_3 = -1 - x_3$$

$$\therefore \left.\frac{\partial z}{\partial x_1}\right|_{x_{NB}=0} = 1 - 3.0 = 1 \quad \text{and} \quad \left.\frac{\partial z}{\partial x_3}\right|_{x_{NB}=0} = 1 - 0 = -1.$$

The values of the partial derivatives of $z$ with respect to $x_1$ and $x_3$ at $x_{NB} = 0$ indicate that $z$ would be increased if $x_1$ is increased.

Therefore, $x_1$ will be the new basic variable. Now we have to find the value of $x_1$.

Since $x_2 = 1 - \frac{1}{2}(x_1 + x_3)$ and $x_3 = 0$, then $x_1 = 2$ if $x_2$ is selected as the non-basic variable.

Again, $\frac{\partial z}{\partial x_1} = 0$ with $x_3 = 0$ implies

$$1 - 3x_1 = 0 \text{ or } x_1 = \frac{1}{3}.$$

Hence, the new value of $x_1$ is given by

$$x_1 = \text{Min}\{2, \tfrac{1}{3}\} = \tfrac{1}{3}, \text{ which corresponds to } \tfrac{\partial z}{\partial x_1} = 0 \text{ with } x_3 = 0.$$

Hence, $x_2$ cannot be removed from the basis.

Now, we introduce a non-basic free variable $u_1$ and the new constraint

$$u_1 = \frac{\partial z}{\partial x_1}, \text{ i.e. } u_1 = 1-3x_1.$$

Hence, the current basic variables will be $x_2$, $x_1$, and the remaining variables $x_3$ and $u_1$ will be non-basic variables.

$$\therefore \; x_B = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \text{ and } x_{NB} = \begin{pmatrix} x_3 \\ u_1 \end{pmatrix}.$$

Now expressing $x_B$ and $z$ in terms of $x_{NB}$, we have

$$\begin{aligned}
x_1 &= (1-u_1)/3 \\
x_2 &= 1-(x_1+x_3)/2 = 1-\frac{1}{6}(1-u_1)-\frac{1}{2}x_3 = \frac{5}{6}+\frac{u_1}{6}-\frac{x_3}{2} \\
\text{and } z &= 4(1-u_1)/3+5+u_1-3x_3-2(1-u_1)^2/9-2(1-u_1)\left(\frac{5}{6}+\frac{u_1}{6}-\frac{x_3}{2}\right)/3 \\
&\quad -2\left(\frac{5}{6}+\frac{u_1}{6}-\frac{x_3}{2}\right)^2 \\
\therefore \; \frac{\partial z}{\partial x_3} &= -3+\frac{1}{3}+\frac{5}{6}+\frac{u_1}{6}-\frac{x_3}{2} \\
\frac{\partial z}{\partial u_1} &= -\frac{4}{3}+1+4(1-u_1)/9+2\left(\frac{5}{6}+\frac{u_1}{6}-\frac{x_3}{2}\right)/3-2(1-u_1)/9 \\
&\quad -2\left(\frac{5}{6}+\frac{u_1}{6}-\frac{x_3}{2}\right)/3.
\end{aligned}$$

Now evaluating these partial derivatives at $x_{NB}=0$, i.e. at $x_3 = 0$ and $u_1 = 0$, we have

$$\left.\frac{\partial z}{\partial x_3}\right|_{x_{NB}=0} = \text{negative} \quad \text{and} \quad \left.\frac{\partial z}{\partial u_1}\right|_{x_{NB}=0} = 0,$$

which indicate that the current $x_B$ provides the optimal solution.

As $x_3 = 0$, $u_1 = 0$, $x_1 = \frac{1}{3}, x_2 = \frac{5}{6}$

and Max $z = 4\cdot\frac{1}{3}+6\cdot\frac{5}{6}-2\cdot\left(\frac{1}{3}\right)^2-2\left(\frac{1}{3}\right)\left(\frac{5}{6}\right)-2\left(\frac{5}{6}\right)^2 = \frac{25}{6}$.

*Example 6* Use Beale's method to solve the following QPP:

Maximize $z = 2x_1 + 3x_2 - 2x_2^2$
subject to the constraints $x_1 + 4x_2 \le 4$, $x_1, + x_2 \le 2$ and $x_1, x_2 \ge 0$.

**Solution**
Introducing slack variables $x_3(\ge 0)$ and $x_4(\ge 0)$ into the given constraints, we get the reduced problem as

$$\begin{aligned}&\text{Maximize } z = 2x_1 + 3x_2 - 2x_2^2\\&\text{subject to } x_1 + 4x_2 + x_3 = 4\\&\qquad x_1 + x_2 + x_4 = 2 \text{ and } x_1, x_2, x_3, x_4 \ge 0.\end{aligned}$$

Let us take $x_1$ and $x_2$ as basic variables and the remaining variables $x_3$ and $x_4$ as non-basic variables.

$$\therefore\ x_B = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ and } x_{NB} = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix}.$$

Now expressing $x_B$ and $z$ in terms of $x_{NB}$, we have

$$\begin{aligned}x_2 &= (2 - x_3 + x_4)/3\\x_1 &= (4 + x_3 - 4x_4)/3\\\text{and } z &= 2(4 + x_3 - 4x_4)/3 + (2 - x_3 + x_4) - 2(2 - x_3 + x_4)^2/9\\\therefore\ \frac{\partial z}{\partial x_3} &= \frac{2}{3} - 1 - \frac{4}{9}(2 - x_3 + x_4)(-1) = -\frac{1}{3} + \frac{4}{9}(2 - x_3 + x_4)\\\text{and } \frac{\partial z}{\partial x_4} &= -\frac{8}{3} + 1 - \frac{4}{9}(2 - x_3 + x_4) = -\frac{5}{3} - \frac{4}{9}(2 - x_3 + x_4).\end{aligned}$$

Now evaluating these partial derivatives at $x_{NB} = 0$, i.e. at $x_3 = 0$, $x_4 = 0$, we have

$$\left.\frac{\partial z}{\partial x_3}\right|_{x_{NB}=0} = -\frac{1}{3} + \frac{8}{9} = \frac{5}{9} \text{ and } \left.\frac{\partial z}{\partial x_4}\right|_{x_{NB}=0} = -\frac{23}{9}.$$

The values of the partial derivatives of $z$ with respect to $x_3$ and $x_4$ at $x_{NB} = 0$ indicate that $z$ would be increase if $x_3$ is increased.

Therefore, $x_3$ will be the new basic variable. Now we have to find the value of $x_3$.

Since $x_1 = (4 + x_3 - 4x_4)/3$ and $x_4 = 0$, then $x_3 = -4$ if $x_1$ is selected as the non-basic variable. This is impossible, as $x_3 \ge 0$.

Again, since $x_2 = (2 - x_3 + x_4)/3$ with $x_4 = 0$, then $x_3 = 2$ if $x_2$ is selected as the non-basic variable. On the other hand, $\frac{\partial z}{\partial x_3} = 0$ with $x_4 = 0$ gives

$$-\frac{1}{3} + \frac{4}{9}(2 - x_3) = 0 \Rightarrow \frac{4}{9}(2 - x_3) = \frac{1}{3} \Rightarrow x_3 = \frac{5}{4}.$$

Hence, the new value of $x_3$ is given by

$x_3 = \min\{2, \frac{5}{4}\} = \frac{5}{4}$, which corresponds to $\frac{\partial z}{\partial x_3} = 0$ with $x_4 = 0$.

Hence, we cannot remove $x_1$ and $x_2$ from the basis.

Now, we introduce a non-basic free variable $u_1$ and the new constraint

$$u_1 = \frac{\partial z}{\partial x_3} \quad \text{i.e.,}\, u_1 = -\frac{1}{3} + \frac{4}{9}(2 - x_3 + x_4).$$

Hence, the current basic variables will be $x_1$, $x_2$, $x_3$, and the remaining variables $x_4$ and $u_1$ will be non-basic variables.

$$\therefore\ x_B = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \text{ and } x_{NB} = \begin{pmatrix} x_4 \\ u_1 \end{pmatrix}.$$

Now expressing $x_B$ and $z$ in terms of $x_{NB}$, we have

$$x_3 = \frac{9}{4}\left(\frac{5}{9} - u_1 + \frac{4}{9}x_4\right) = \frac{5}{4} - \frac{9}{4}u_1 + x_4$$

$$x_1 = \frac{1}{3}(4 + x_3 - 4x_4) = \frac{1}{3}\left(4 + \frac{5}{4} - \frac{9}{4}u_1 + x_4 - 4x_4\right) = \frac{1}{3}\left(\frac{21}{4} - \frac{9}{4}u_1 - 3x_4\right) = \frac{7}{4} - \frac{3}{4}u_1 - x_4$$

$$x_2 = \frac{1}{3}(2 - x_3 + x_4) = \frac{1}{3}\left(2 - \frac{5}{4} + \frac{9}{4}u_1 - x_4 + x_4\right) = \frac{1}{3}\left(\frac{3}{4} + \frac{9}{4}u_1\right) = \frac{1}{4}(1 + 3u_1)$$

$$\text{and } z = 2\left(\frac{7}{4} - \frac{3}{4}u_1 - x_4\right) + \frac{3}{4}(1 + 3u_1) - \frac{1}{8}(1 + 3u_1)^2.$$

Now, $\frac{\partial z}{\partial x_4} = -2$ and $\frac{\partial z}{\partial u_1} = -\frac{3}{2} + \frac{9}{4} - \frac{3}{4}(1 + 3u_1)$.

Now evaluating these partial derivatives at $x_{NB} = 0$, i.e. at $x_4 = 0$, $u_1 = 0$, we have

$$\left.\frac{\partial z}{\partial x_4}\right|_{x_{NB}=0} = -2 \text{ and } \left.\frac{\partial z}{\partial u_1}\right|_{x_{NB}=0} = -\frac{3}{2} + \frac{9}{4} - \frac{3}{4} = 0.$$

These partial derivatives indicate that the current $x_B$ provides the optimal solution.

Since $x_4 = 0,\ u_1 = 0,$ then $x_1 = \frac{7}{4},\ x_2 = \frac{1}{4}$ and $z = \frac{7}{2} + \frac{3}{4} - \frac{1}{8} = \frac{33}{8}$.

Hence, the optimal solution is $x_1 = \frac{7}{4}$, $x_2 = \frac{1}{4}$ and $\max z = \frac{33}{8}$.

*Example 7* Solve the following QPP by Beale's method:

$$\begin{aligned}
&\text{Maximize } z = 10x_1 + 25x_2 - 10x_1^2 - x_2^2 - 4x_1x_2\\
&\text{subject to } \quad \begin{aligned} x_1 + 2x_2 &\le 10\\ x_1 + x_2 &\le 9 \end{aligned}\\
&\qquad\qquad \text{and } x_1, x_2 \ge 0.
\end{aligned}$$

**Solution**

Introducing slack variables $x_3(\geq 0)$ and $x_4(\geq 0)$ into the given problem, we convert it into

$$\begin{aligned}
&\text{Maximize } z = 10x_1 + 25x_2 - 10x_1^2 - x_2^2 - 4x_1x_2\\
&\text{subject to}\\
&\qquad x_1 + 2x_2 + x_3 = 10\\
&\qquad x_1 + x_2 + x_4 = 9 \ \text{ and } \ x_1, x_2, x_3, x_4 \ge 0
\end{aligned}$$

Let us consider $x_1, x_2$ as basic variables and $x_3, x_4$ as non-basic variables.

Now expressing $x_B$ and $z$ in terms of $x_{NB}$, we have

$$x_1 = 8 + x_3 - 2x_4 \ \text{ and } \ x_2 = 1 - x_3 + x_4$$

and

$$\begin{aligned}
z = {} & 10(8 + x_3 - 2x_4) + 25(1 - x_3 + x_4)\\
& - 10(8 + x_3 - 2x_4)^2 - (1 - x_3 + x_4)^2 - 4(8 + x_3 - 2x_4)(1 - x_3 + x_4).
\end{aligned}$$

Now

$$\begin{aligned}
\frac{\partial z}{\partial x_3} = {} & 10 - 25 - 20(8 + x_3 - 2x_4) + 2(1 - x_3 + x_4)\\
& - 4(1 - x_3 + x_4) + 4(1 + x_3 - 2x_4)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial z}{\partial x_4} = {} & -20 + 25 - 20(-2)(8 + x_3 - 2x_4) - 2(1 - x_3 + x_4) + 8(1 - x_3 + x_4)\\
& - 4(8 + x_3 - 2x_4).
\end{aligned}$$

$$\left.\frac{\partial z}{\partial x_3}\right|_{x_{NB}=0} = -145 \text{ and } \left.\frac{\partial z}{\partial x_4}\right|_{x_{NB}=0} = 299.$$

The values of the partial derivatives of $z$ with respect to $x_3$ and $x_4$ at $x_{NB} = 0$ indicate that $z$ will increase if $x_4$ increases.

Since $2x_4 = (8 + x_3 - x_1)$, if $x_4$ is increased to a value greater than 4, then $x_1$ will become negative and $x_3 = 0$.

Now

$$\frac{\partial z}{\partial x_4} = 0$$
$$\text{or } -20 + 25 + 40(8 + x_3 - 2x_4) - 2(1 - x_3 + x_4) + 8(1 - x_3 + x_4) - 4(8 + x_3 - 2x_4) = 0.$$

Setting $x_3 = 0$, we get

$$\begin{aligned}
&5 + 40(8 - 2x_4) - 2(1 + x_4) + 8(1 + x_4) - 4(8 - 2x_4) = 0\\
&\text{or,}\quad 5 + 320 - 80x_4 - 2 - 2x_4 + 8 + 8x_4 - 32 + 8x_4 = 0\\
&\text{or,}\quad -66x_4 = -299\\
&\text{or,}\quad x_4 = \frac{299}{66}
\end{aligned}$$

Hence, the new value of $x_4$ is given by

Min $\left\{4, \frac{299}{66}\right\} = 4$ which corresponds to $\frac{\partial z}{\partial x_4} = 0$ with $x_3 = 0$.

So, the new basic variables are $x_4$ and $x_2$.

Now, expressing $x_B$ and $z$ in terms of $x_{NB}$, we have

$$\begin{aligned}
&x_2 = 5 - \frac{1}{2}(x_1 + x_3)\\
&x_4 = 4 + \frac{1}{2}(x_3 - x_1)\\
&z = 10x_1 + 25\left\{5 - \frac{1}{2}(x_1 + x_3)\right\} - 10x_1^2 - \left\{5 - \frac{1}{2}(x_1 + x_3)\right\}^2 - 4x_1\left\{5 - \frac{1}{2}(x_1 + x_3)\right\},
\end{aligned}$$

where $x_B = (x_2, x_4), x_{NB} = (x_1, x_3)$.

Now,

$$\frac{\partial z}{\partial x_1} = 10 - \frac{25}{2} - 20x_1 - 2\left\{5 - \frac{1}{2}(x_1 + x_3)\right\}\left(-\frac{1}{2}\right) - 4\left\{5 - \frac{1}{2}(x_1 + x_3)\right\}$$
$$- 4x_1\left(-\frac{1}{2}\right)$$

and

$$\frac{\partial z}{\partial x_3} = -\frac{25}{2} - 2\left\{5 - \frac{1}{2}(x_1 + x_3)\right\}\left(-\frac{1}{2}\right) - 4x_1\left(-\frac{1}{2}\right).$$

$$\left.\frac{\partial z}{\partial x_1}\right|_{x_{NB}=0} = -\frac{35}{2} \text{ and } \left.\frac{\partial z}{\partial x_3}\right|_{x_{NB}=0} = -\frac{15}{2}.$$

Both partial derivatives are negative at $x_{NB}$; hence, neither $x_1$ nor $x_3$ can be introduced to increase $z$, and the current $x_B$ provides the optimal solution.

Setting $x_1 = 0, x_3 = 0$, we have $x_2 = 5$ and $x_4 = 4$.

Hence, the optimal solution is given by

$$x_2 = 5 \text{ and } x_4 = 4 \text{ and } \text{Max}\, z = 0 + 25 \times 5 - 0 - 25 - 0 = 100.$$

**Note** In this problem, if we initially consider $x_2, x_4$ as basic variables and $x_1, x_3$ as non-basic variables, then this problem can be solved very easily in one iteration.

## 12.8 Exercises

1. Use Wolfe's method for solving the following problem:
   Maximize $z = 10x_1 - x_1^2 + 10x_2 - x_2^2$, where $x_1 + x_2 \le 14, -x_1 + x_2 \le 6, x_1, x_2 \ge 0$.
2. Use the Kuhn-Tucker conditions to solve the problem:

   Maximize $z = 12x_1 + 21x_2 + 2x_1x_2 - 2x_1^2 - 2x_2^2$
   subject to $x_2 \le 8,\ x_1 + x_2 \le 10$.

3. Use Wolfe's method for solving the following problem:

   Maximize $Z = 2x_1 + 3x_2 - 2x_1^2$
   subject to the constraints $x_1 + 4x_2 \le 4,\ x_1 + x_2 \le 2, x_1,\ x_2 \ge 0$.

4. Use Beale's method to solve the following problem:

$$\text{Minimize } Z = 6 - 6x_1 - 2x_1^2 - 2x_1x_2 + 2x_2^2$$
$$\text{such that } x_1 + x_2 \leq 2 \ \text{ and } \ x_1, x_2 \geq 0.$$

5. Use Wolfe's method for solving the following problem:

$$\text{Maximize } z = 10x_1 + 25x_2 - 10x_1^2 - x_2^2 - 4x_1x_2$$
$$\text{subject to } x_1 + 2x_2 \leq 10, x_1 + x_2 \leq 9 \ \text{ and } \ x_1, x_2 \geq 0.$$

6. Use Beale's method to solve the following problem:

$$\text{Minimize } z = x_1^2 - x_1x_2 + 2x_2^2 - x_1 - x_2$$
$$\text{subject to}$$
$$2x_1 + x_2 \leq 1$$
$$\text{and } x_1 \geq 0, \ x_2 \geq 0.$$

# Chapter 13
# Game Theory

## 13.1 Objectives

The objectives of this chapter are to discuss:

- The process of finding the optimal strategies in conflict and competitive situations.
- Different types of finite and infinite games and their solution procedures.

## 13.2 Introduction

In most practical problems, there arises a situation where there are two or more opposite parties (particularly players) with conflicting of interests and actions. This type of situation is called a competitive situation. A large variety of competitive situations are commonly observed in daily life, e.g. in political campaigns, elections, advertisements, military battles, etc.

A competitive situation is called a game if the following properties are satisfied:

(i) There is a conflict of interests between the participants.
(ii) The number of competitors (participants), called players, is finite.
(iii) Each of the participants has a finite/infinite set of possible courses of action.
(iv) The rules governing these choices are specified and known to the players; a play of the game results when each of the players chooses a single course of action from a list of courses available to him.
(v) The outcome of the game is affected by the choices made by all the players.
(vi) The outcome for all specific sets of choices by all the players is known in advance and defined.

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*, Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_13

(vii) Every play, i.e. combination of courses of action, determines an outcome (which may be money or points) which determines a set of payments (positive, negative or zero), one to each player.

## 13.3 Cooperative and Non-cooperative Games

Games in which the objective of each participant (player) is to achieve the largest possible individual gain (the 'payoff') are called non-cooperative games.

Games in which the actions of the players are directed to maximize the gain of 'collectives' without subsequent subdivision of the gain among the players within the condition are called cooperative games.

Here we shall consider only non-cooperative games.

Let $I$ denote the set of all players. We shall assume that $I$ is finite. Let $I = \{1, 2, \ldots, n\}$, where $n$ denotes the number of players in the game. Let each player $i \in I$ have at his disposal a certain set $S_i$ of the available actions, which are called strategies. Each player has at least two distinct strategies.

The process of the game consists of each one of the players choosing a certain strategy $s_i \in S_i$. Thus, as a result of each 'round' of the game, a system of strategies $(s_1, \ldots, s_n) \equiv s$ is put together. This system is called a situation. The set of all strategies is denoted by $S = S_1 \times S_2 \times \cdots \times S_n$. In each situation the players attain certain gains (payoffs). The payoff of player $i$ in situation $s$ is denoted by $H_i(s)$. The function $H_i$ defined on the set of all situations is called the payoff function of player $i$. Thus,

$H_i : S \to R$, where $R$ is a set of real numbers.

**Definition** A system denoted by $\Gamma = \langle I, \{S_i\}_{i \in I}, \{H_i\}_{i \in I} \rangle$, where $I$ and $S_i (i \in I)$ are sets and $H_i$'s are functions defined on the set $S = \prod_{i=1}^{n} S_i$ taking on real values, is called a non-cooperative game.

**Definition** A non-cooperative game $\Gamma = \langle I, \{S_i\}_{i \in I}, \{H_i\}_{i \in I} \rangle$ is called a constant-sum game if there exists a constant $C$ such that $\sum_{i \in I} H_i(s) = C$ for each and every situation $s \in S$.

***Admissible situation and equilibrium situation***

Let $s = (s_1, \ldots, s_{i-1}, s_i, s_{i+1}, \ldots, s_n)$ be an arbitrary situation in the game $\Gamma$, and let $s_i$ be a strategy of player $i$. We form a new situation that differs from the situation $s$ only in that the strategy $s_i$ for player $i$ is now replaced by $s_i'$. Denote this as $s \| s_i' = (s_1, \ldots, s_{i-1}, s_i', s_{i+1}, \ldots, s_n)$. If $s_i$ and $s_i'$ coincide, then obviously $s \| s_i' = s$.

***Admissible situation***

A situation $s = (s_1, \ldots, s_n)$ in a game $\Gamma = \langle I, \{s_i\}, \{H_i\} \rangle$ is called an admissible situation for player $i$ if for any other strategy $s'_i$ of this player we have

$$H_i\left(s \| s'_i\right) \leq H_i(s).$$

The term 'admissible' can thus be justified by the fact that if in situation $s$ there exists a strategy $s'_i$ for player $i$ such that $H_i\left(s \| s'_i\right) > H_i(s)$, then player $i$, knowing that the situation $s$ will materialize, may choose the strategy $s'_i$ at the last moment and as a result of this choice end up with a large payoff. In this sense the situation may be viewed as admissible for player $i$.

***Equilibrium situation***

An admissible situation for all players is called an equilibrium situation; i.e.

$$H_i\left(s \| s'_i\right) \leq H_i(s)$$

is satisfied for any player $i$ and for any strategy $s'_i \in S_i$.

***Equilibrium strategy***

An equilibrium strategy is one that appears in at least one equilibrium situation of the game.

***Solution of the game***

The process of determination of an equilibrium situation in a non-cooperative game is often referred to as the solution of the game.

***Strategic equivalence of games***

Let two non-cooperative games with the same set of players and the same set of strategies be given (i.e. the games differ only in their payoff functions) as

$$\Gamma_1 = \left\langle I, \{S_i\}_{i \in I}, \left\{H_i^{(1)}\right\}_{i \in I} \right\rangle$$

$$\Gamma_2 = \left\langle I, \{S_i\}_{i \in I}, \left\{H_i^{(2)}\right\}_{i \in I} \right\rangle.$$

**Definition** The games $\Gamma_1$ and $\Gamma_2$ are called strategically equivalent if a positive number $k$ and numbers $C_i$ (for each of the players $i \in I$) exist such that for any situations

$$H_i^{(1)}(s) = k H_i^{(2)}(s) + C_i.$$

We denote this fact by the symbol $\Gamma_1 \sim \Gamma_2$.

***Properties of strategic equivalence***

(i) Reflexivity: $\Gamma \sim \Gamma$; this can easily be proved choosing $k = 1$ and $C_i = 0$.
(ii) Symmetry: $\Gamma_1 \sim \Gamma_2 \Rightarrow \Gamma_2 \sim \Gamma_1$.

*Proof* Since $\Gamma_1 \sim \Gamma_2$, we can write $H_i^{(1)} = k\,H_i^{(2)}(s) + C_i \quad k > 0, \forall\, i \in I$

$$\begin{aligned}\text{or } H_i^{(2)}(s) &= \frac{1}{k}H_i^{(1)}(s) - \frac{C_i}{k}\\ &= k' H_i^{(1)}(s) + C_i' \text{ where } k' = \frac{1}{k}, \quad C_i' = -\frac{C_i}{k}.\end{aligned}$$

Hence, by definition, $\Gamma_2 \sim \Gamma_1$.

(iii) Transitivity: If $\Gamma_1 \sim \Gamma_2$ and $\Gamma_2 \sim \Gamma_3$, then $\Gamma_1 \sim \Gamma_3$.
This can easily be proved using the definition.

*Note* The basic difference between two strategically equivalent games lies in the amount of initial capital possessed by the players and in the relative units in which the payoffs are measured. This is indicated by the coefficient $k$. It is therefore natural to suppose that the rational behaviour of players in distinct strategically equivalent games should be the same.

**Theorem 13.1** *Strategically equivalent games possess the same equilibrium situation.*

*Proof* Let $\Gamma_1 \sim \Gamma_2$, where $\Gamma_1 = \left\langle I,\ \{S_i\}_{i\in I},\ \left\{H_i^{(1)}\right\}_{i\in I}\right\rangle$

$$\Gamma_2 = \left\langle I, \{S_i\}_{i\in I}, \left\{H_i^{(2)}\right\}_{i\in I}\right\rangle$$

and let $s$ be an equilibrium situation in $\Gamma_1$. This means that $\forall i \in I$ and $s_i^0 \in S_i$, we have the inequality
$H_i^{(1)}\left(s \parallel s_i^0\right) \le H_i^{(1)}(s)$, where $s = (s_1, \ldots, s_{i-1}, s_i, s_{i+1}, \ldots, s_n)$, $s \parallel s_i^0 = (s_1, \ldots, s_{i-1}, s_i^0, s_{i+1}, \ldots, s_n)$.

Now since $\Gamma_1 \sim \Gamma_2, H_i^{(1)} = k\,H_i^{(2)} + C_i (k > 0) \quad \forall i \in I$ and $s_i \in S_i$

$$\begin{aligned}&\text{i.e. } H_i^{(1)}\left(s \parallel s_i^0\right) \le k\,H_i^{(2)} + C_i\\ &\text{or } k\,H_i^{(2)}\left(s \parallel s_i^0\right) + C_i \le H_i^{(2)}(s) + C_i \text{ as } H_i^{(1)}\left(s \parallel s_i^0\right) = k\,H_i^{(2)}\left(s \parallel s_i^0\right) + C_i\\ &\text{or } H_i^{(2)}\left(s \parallel s_i^0\right) \le H_i^{(2)}(s) \quad \forall i \in I \quad \text{and } s_i^0 \in S_i \text{ since } k > 0\end{aligned}$$

which implies $s$ is an equilibrium situation in $\Gamma_2$.

***Zero-sum game***

A non-cooperative game $\Gamma$ is called a zero-sum game if for each situation $s$, $\sum_{i\in I} H_i(s) = 0$.

**Theorem 13.2** *Any non-cooperative constant-sum game is strategically equivalent to a certain zero-sum game.*

*Proof* Let $\Gamma = \langle I, \{S_i\}_{i\in I}, \{H_i\}_{i\in I}\rangle$ be a constant-sum game. Then $\sum_{i\in I} H_i(s) = C$, $C$ being a constant. Now let us consider the game $\Gamma' = \left\langle I, \{S_i\}_{i\in I}, \left\{H'_i\right\}_{i\in I}\right\rangle$, where $H'_i = H_i - C_i$ and $\sum_{i\in I} C_i = C$. Then obviously $\Gamma \sim \Gamma'$ when $k = 1$.

But $\sum_i H'_i(s) = \sum_i H_i(s) - \sum_i C_i = C - C = 0$,

So $\Gamma'$ is a zero-sum game.

## 13.4 Antagonistic Game

The game $\Gamma = \langle I, \{S_i\}_{i\in I}, \{H_i\}_{i\in I}\rangle$ is called antagonistic if there are only two players in the game (i.e. $I = \{1, 2\}$), and the values of the payoff function for these players in each situation are the same in absolute value but are of opposite sign.

Thus, for an antagonistic game, $I = \{1, 2\}$ and $H_2(s) = -H_1(s)$. Hence, $H_1(s) + H_2(s) = 0$, $s = (s_1, s_2) \in S_1 \times S_2$. Thus, an antagonistic game is also a two-person zero-sum game.

Thus, to define an antagonistic game, it is sufficient to stipulate the payoff function of one of the players only. It is usually denoted by the triplet $\langle X, Y, M\rangle$, where $X$, $Y$ are the sets of strategies for players 1 and 2 respectively and $M$ is the payoff function for player 1 which is a real-valued function such that $M : X \times Y \to R$.

*Note* It is obvious that each non-cooperative two-person constant zero-sum game is strategically equivalent to a certain antagonistic game.

***Equilibrium situation for an antagonistic game***

Let $\langle X, Y, M\rangle$ be an antagonistic game, $H_1 = M, H_2 = -M$.

Hence, $s_1 \equiv x \in X,\ s_2 \equiv y \in Y,\ s \equiv (x, y) \in X \times Y$.

If $(x, y)$ is an equilibrium situation, then

$$
\begin{aligned}
&H_1(x,y) \geq H_1(x',y)\forall x' \in X \text{ and } H_2(x,y) \geq H_2(x,y')\forall y' \in Y,\\
\text{i.e. } &M(x,y) \geq M(x',y)\forall x' \in X \text{ and } -M(x,y) \geq -M(x,y')\forall y' \in Y,\\
\text{i.e. } &M(x',y) \leq M(x,y) \leq M(x',y)\forall x' \in X, y' \in Y.
\end{aligned}
$$

Changing the notation $(x, y)$ to $(x^*, y^*)$ and $(x', y')$ to $(x, y)$, we see that $(x^*, y^*)$ will be an equilibrium situation for an antagonistic game $\langle X, Y, M \rangle$ if

$$M(x, y^*) \leq M(x^*, y^*) \leq M(x^*, y) \ \forall x \in X, \ y \in Y.$$

The point $(x^*, y^*)$ is called the saddle point of $M(x, y)$.

## 13.5 Matrix Games or Rectangular Games

Antagonistic games in which each player possesses a finite number of strategies are called matrix games or rectangular games.

When there are two competitors playing a game, it is called a two-person game. In a game, if the number of competitors is more than two, say $n$, the game is referred to as an $n$-person game.

If, in a game, the algebraic sum of the payments to all the competitors is zero for every possible outcome of the game, then the game is said to be a zero-sum game. Otherwise, it is said to be non-zero-sum game.

A game with only two players in which the gains of one player are exactly equal to the losses of another player is called a two-person zero-sum game. It is also called a rectangular game because their payoff matrix is in a rectangular form.

**Some basic terms**

**Player**: The competitors in the game are known as players. A player may be an individual or a group of individuals or an organization.

**Strategy**: A strategy for a player is defined as a set of rules or alternative courses of action available to him/her in advance, by which the player decides the course of action that he should adopt. A strategy may be of two types: (i) pure strategy, (ii) mixed strategy.

**Pure strategy**: If the players select the same strategy each time, then it is referred to as a pure strategy. In this case, each player knows exactly what the other player is going to do, and the objective of the players is to maximize the gains or to minimize the losses.

**Payoff matrix**: The payoff is the outcome of playing the game. When the players select their particular strategies, the payoffs (gains or losses) can be represented in the form of a matrix called a payoff matrix. Since the game is zero sum, the gain of one player is exactly equal to the loss of the other and vice versa. In other words, one player's payoff table would contain the same amounts as in the payoff table of the other player with the sign changed. Thus, it is sufficient to construct payoff only for one of the players.

Let player $A$ have $m$ strategies, say, $A_1, A_2, \ldots, A_m$ and player $B$ have $n$ strategies, say, $B_1, B_2, \ldots, B_\mathrm{n}$. Here, it is assumed that each player has his choices from among the pure strategies. Also, it is assumed that player $A$ is always the gainer and player $B$ is always the loser. That is, all payoffs are assumed in terms of player $A$. Let $a_{ij}$ be the payoff which is the gain of player $A$ from player $B$ if player $A$ chooses strategy $A_i$ whereas player $B$ chooses $B_j$. Then the payoff matrix of player $A$ is

$$\begin{array}{cc} & \text{Player } B \\ & \begin{array}{cccc} B_1 & B_2 & \cdots & B_n \end{array} \\ \text{Player } A \begin{array}{c} A_1 \\ A_2 \\ \vdots \\ A_m \end{array} & \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \end{array} \text{ i.e., } \left[a_{ij}\right]_{mxn}$$

The payoff matrix of player $B$ is $[-a_{ij}]_{m\times n}$.

*Example 1* Consider a two-person coin tossing game. Each player tosses an unbiased coin simultaneously. Player $B$ pays Rs. 7.00 to player $A$ if the outcomes of both tosses are heads and Rs. 4.00 if both the outcomes are tails; otherwise, player $A$ pays Rs. 3.00 to player $B$. This two-person game is a zero-sum game, since the winnings of one player are the losses of the other. Each player has choices from among two pure strategies $H$ and $T$.

In that case, $A$'s payoff matrix will be

$$\begin{array}{cc} & \text{Player } B \\ & \begin{array}{cc} H & T \end{array} \\ \text{Player } A \begin{array}{c} H \\ T \end{array} & \begin{pmatrix} 7 & -3 \\ -3 & 4 \end{pmatrix} \end{array}$$

**Maximin-minimax principle or maximin-minimax criteria of optimality**

This principle is used for the selection of optimal strategies by two players.

It states that 'If a player lists his worst possible outcomes of all potential strategies then he will choose that strategy which corresponds to the best of these worst outcomes'.

Let player $A$'s payoff matrix be

$$\begin{array}{cc} & \text{Player } B \\ & \begin{array}{cccccc} B_1 & B_2 & \cdots & B_j & \cdots & B_n \end{array} \\ \text{Player } A \begin{array}{c} A_1 \\ A_2 \\ \vdots \\ A_i \\ \vdots \\ A_m \end{array} & \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} \end{array}.$$

If player $A$ takes the strategy $A_i$, then surely he will get at least $a_{ij}(i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n)$ for taking any strategy by the opponent player $B$.

Thus, by the maximin-minimax criteria of optimality, player $A$ will choose that strategy which corresponds to the best of these worst outcomes:

$$\underset{j}{\text{Min}}\, a_{1j},\ \underset{j}{\text{Min}}\, a_{2j}, \ldots,\ \underset{j}{\text{Min}}\, a_{mj}.$$

Thus, the maximin value for player $A$ is given by $\text{Max}_i \left[\text{Min}_j\, a_{ij}\right]$.

Similarly, player $B$ will choose that strategy which corresponds to the best (minimum) of the worst outcomes (maximum losses):

$$\underset{i}{\text{Max}}\, a_{i1}, \underset{i}{\text{Max}}\, a_{i2}, \ldots, \underset{i}{\text{Max}}\, a_{in}.$$

Thus, the minimax value for player $B$ is given by $\text{Min}_j \left[\text{Max}_i\, a_{ij}\right]$.

$$\text{Let } \underset{i}{\text{Max}} \left[\underset{i}{\text{Min}}\ a_{ij}\right] = a_{pq} \tag{13.1}$$

$$\text{and } \underset{j}{\text{Min}} \left[\underset{i}{\text{Max}}\, a_{ij}\right] = a_{rs}. \tag{13.2}$$

From (13.1), it follows that $a_{pq}$ is the minimum element in the $p$th row,

$$\text{i.e. } a_{pq} \leq a_{ps}, \tag{13.3}$$

where $a_{ps}$ is another element of the $p$th row.

From (13.2), it follows that $a_{rs}$ is the maximum element in the $s$th column,

$$\text{i.e. } a_{ps} \leq a_{rs}, \tag{13.4}$$

where $a_{ps}$ is another element of the $s$th column.

From (13.3) and (13.4), we have $a_{pq} \le a_{rs}$.

$$\operatorname*{Max}_{i}\left[\operatorname*{Min}_{j} a_{ij}\right] \le \operatorname*{Min}_{j}\left[\operatorname*{Max}_{i}\ a_{ij}\right],$$

i.e. maximin for $A\ \le$ minimax for $B$.

Therefore, $\mathrm{Max}_i\left[\mathrm{Min}_j\, a_{ij}\right]$; i.e. maximin for $A$ is called the lower value of the game and is denoted by $\mathrm{Min}_j\left[\mathrm{Max}_i\ a_{ij}\right]$, and minimax for $B$ is called the upper value of the game and is denoted by $\underline{v}$.

If $v$ is the value of the game, then it will always satisfy the inequality

$$\text{maximin for } A \le v \le \text{minimax for } B.$$

If for a game, $\bar{v} = \underline{v} = a_{lk}$, then the game possesses a solution given by:

(i) The optimal strategy for player $A$ is the strategy $A_l$.
(ii) The optimal strategy for player $B$ is the strategy $B_k$.
(iii) The value of the game is $v = a_{lk}$.

For the case of pure strategy, such a game is called a game with a saddle point.

**Saddle point**: A saddle point is a position in the payoff matrix where the maximum of row minima coincides with the minimum of the column maxima. The cell entry (or payoff) at the saddle point position is called the value of the game.

A game for which maximin for $A$ = minimax for $B$ is called a game with a saddle point. Thus, in a game with a saddle point, the players use pure strategies; i.e. they choose the same course of action throughout the game.

*Note*

(i) A game is said to be fair if $\bar{v} = \underline{v} = 0$.
(ii) A game is said to be strictly determinable if maximin value = minimax value = the value of the game $\neq 0$, i.e. $\underline{v} = v = \bar{v}(\neq 0)$.
(iii) Not all payoff matrices possess a saddle point, and so the determination of the optimal solution is not an easy task.

*Example 2* Solve the game whose payoff matrix is given by

$$\begin{array}{ccc} & & \text{Player } B \\ & & \begin{array}{ccc} B_1 & B_2 & B_3 \end{array} \\ \text{Player } A & \begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} & \begin{bmatrix} 1 & 3 & 3 \\ 0 & -4 & -3 \\ 1 & 5 & -1 \end{bmatrix} \end{array}$$

**Solution** To find the saddle point, the row minima and column maxima are selected and displayed on the right side of the corresponding row and in the bottom of the corresponding column respectively.

$$\begin{array}{ccc} & & \text{Player } B & \\ & & \begin{matrix} B_1 & B_2 & B_3 \end{matrix} & \text{Row minima} \\ & \begin{matrix} A_1 \\ A_2 \\ A_3 \end{matrix} & \begin{bmatrix} 1 & 3 & 3 \\ 0 & -4 & -3 \\ 1 & 5 & -1 \end{bmatrix} & \begin{matrix} 1 \\ -4 \\ -1 \end{matrix} \\ \text{Player } A & & & \\ \text{Column maxima} & & \begin{matrix} 1 & 5 & 3 \end{matrix} & \end{array}$$

Hence, the maximin value of the game

is $\underline{v} = \text{Max}\{1, -4, -1\} = 1$ whose location is in the cell (1, 1), and the minimax value of the payoff matrix is $\overline{v} = \text{Min}\{1, 5, 3\} = 1$ whose locations are in the cells (1, 1) and (3, 1).

Hence, the payoff matrix has a saddle point at the position (1, 1). The solution of the game is given by:

(i) The optimal strategy for player $A$ is $A_1$.
(ii) The optimal strategy for player $B$ is $B_1$.
(iii) The value of the game is 1.

*Example 3* For what value of λ is the game with the following payoff matrix strictly determinable?

$$\begin{array}{ccc} & & \text{Player } B \\ & & \begin{matrix} B_1 & B_2 & B_3 \end{matrix} \\ \text{Player } A & \begin{matrix} A_1 \\ A_2 \\ A_3 \end{matrix} & \begin{bmatrix} \lambda & 6 & 2 \\ -1 & \lambda & -7 \\ -2 & 4 & \lambda \end{bmatrix} \end{array}$$

**Solution** Ignoring the value of λ, we shall find the maximin and minimax values of the payoff. For this purpose, we have

$$\begin{array}{cccc} & & \text{Player } B & \\ & & \begin{matrix} B_1 & B_2 & B_3 \end{matrix} & \text{Row minima} \\ \text{Player } A & \begin{matrix} A_1 \\ A_2 \\ A_3 \end{matrix} & \begin{bmatrix} \lambda & 6 & 2 \\ -1 & \lambda & -7 \\ -2 & 4 & \lambda \end{bmatrix} & \begin{matrix} 2 \\ -7 \\ -2 \end{matrix} \\ \text{Column maxima} & & \begin{matrix} -1 & 6 & 2 \end{matrix} & \end{array}$$

$\therefore \underline{v} =$ Max{2, −7, −2} = 2 and Min{−1, 6, 2} = −1.

We know that the game is strictly determinable if $\underline{v} = \bar{v} = v \neq 0$.

Hence, $-1 \leq \lambda \leq 2 (\lambda \neq 0)$.

***Mixed strategy***

If the payoff matrix does not have a saddle point, then there exists no equilibrium situation of the form $(i^*, j^*)$; i.e. an optimal strategy of the players cannot be a single (pure) strategy.

Let us consider the matrix game whose payoff matrix is $A = \left[a_{ij}\right]_{m \times n}$. By a mixed strategy for player 1, we shall mean an ordered $m$-tuple $(x_1, \ldots, x_m)$ or the vector $\xi = (x_1, \ldots, x_m)'$, where $x_i \geq 0$ $i = 1, 2, \ldots, m$ and $\sum_{i=1}^{m} x_i = 1$. The components $x_1, x_2, \ldots, x_m$ may be thought of as the relative frequencies (or probabilities) with which player 1 chooses the strategies (pure) 1, 2, …, $m$. It should be noted that $e_i^m = (0, \ldots, 0, 1, 0, \ldots, 0)', i = 1, 2, \ldots, m$ represents the pure strategy of player 1, and $\xi = \sum_{i=1}^{m} e_i^m x_i$.

Similarly, a mixed strategy for player 2 is denoted by $\eta = (y_1, y_2, \ldots, y_n)'$, where $y_j \geq 0$, $j = 1, 2, \ldots, n$ and $\sum_{j=1}^{n} y_j = 1$. Note that $e_j^n = (0, 0, \ldots, 0, 1, 0, \ldots, 0)'$ represents the pure strategy of player 2, and $\eta = \sum_{j=1}^{n} e_j^m y_j$.

Let $S_m = \left\{\xi : x_i \geq 0, \sum_{i=1}^{m} x_i = 1\right\}$, then $S_m \in E_m$. This $S_m$ is the space of mixed strategies for player 1. Note that a pure strategy is a particular case of a mixed strategy.

Similarly, let $S_n = \left\{\eta : y_j \geq 0, \sum_{j=1}^{n} y_j = 1\right\}$, then $S_n \in E_n$. Thus, $S_n$ is the space of mixed strategies for player 2.

***Mixed extension of a matrix game without a saddle point***

Let players 1 and 2 choose their mixed strategies as $\xi = (x_1, x_2, \ldots, x_m)'$, $\eta = (y_1, y_2, \ldots, y_n)'$ independently of one another in a matrix game with payoff matrix $A = \left[a_{ij}\right]_{m \times n}$.

**Definition** A pair $(\xi, \eta)$ of mixed strategies for the players in a matrix game is called a situation in mixed strategies.

In a situation $(\xi, \eta)$ of mixed strategies, each usual situation $(i, j)$ in pure strategies becomes a random event occurring with probabilities $x_i$, $y_j$. Since in the situation $(i, j)$ player 1 receives a payoff $a_{ij}$, the expected value of his payoff under $(\xi, \eta)$ is equal to

$$E(\xi, \eta) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_i y_j.$$

We thus arrive at a new game which can be described as follows.

**Definition** A mixed extension of a matrix game is an antagonistic game, denoted by $\langle S_m, S_n, E\rangle$, in which the set of strategies for the players is the set of their mixed strategies in the original game, and the payoff function for player 1 is defined by

$$E(\xi, \eta) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}\, x_i\, y_j.$$

Recalling the definition of saddle point of an antagonistic game, we observe that the situation $(\xi^*, \eta^*)$ in a mixed extension of a matrix game will be a saddle point (i.e. an equilibrium situation), provided that for any $\xi \in S_m$ and $\eta \in S_n$ the following double inequality is satisfied:

$$E(\xi,\, \eta^*) \le E(\xi^*,\, \eta^*) \le E(\xi^*, \eta) \quad \forall \xi \in S_m, \eta \in S_n.$$

**Theorem 13.3** *Let E* $(\xi,\eta)$ *be such that both* $\operatorname*{Min}_{\eta\in S_n} \operatorname*{Max}_{\xi\in S_m} E(\xi, \eta)$ *and* $\operatorname*{Max}_{\xi\in S_m} \operatorname*{Min}_{\eta\in S_n}$ $E(\xi, \eta)$ *exist; then*

$$\operatorname*{Min}_{\eta\in S_n} \operatorname*{Max}_{\xi\in S_m} E(\xi, \eta) \ge \operatorname*{Max}_{\xi\in S_m} \operatorname*{Min}_{\eta\in S_n} E(\xi, \eta).$$

*Proof* Let $\xi_\circ$ and $\eta_\circ$ be two arbitrarily chosen mixed strategies for players $A$ and $B$ respectively.

Then, for every $\xi \in S_m$, we have

$$\operatorname*{Max}_{\xi\in S_m} E(\xi,\, \eta_\circ) \ge E(\xi_\circ, \eta_\circ), \tag{13.5}$$

and for every $\eta \in S_n$, we have

$$\operatorname*{Min}_{\eta\in S_n} E(\xi_\circ,\, \eta) \le E(\xi_\circ, \eta_\circ). \tag{13.6}$$

Hence, from (13.5) and (13.6), the inequality

$$\operatorname*{Min}_{\eta\in S_n} E(\xi_\circ, \eta) \le \operatorname*{Max}_{\xi\in S_m} E(\xi, \eta_\circ)$$
$$\text{or} \ \operatorname*{Max}_{\xi\in S_m} E(\xi, \eta_\circ) \ge \operatorname*{Min}_{\eta\in S_n} E(\xi_\circ, \eta)$$

holds for all $\xi$ and $\eta$.

Since $\eta_\circ$ is an arbitrarily chosen mixed strategy, the preceding inequality holds for all values of $\eta$. Hence, if $\eta_\circ$ is such a strategy for which $\operatorname{Max}_\xi E(\xi, \eta)$ has the maximum value, the inequality remains true. Therefore, $\underset{\eta}{\text{Min}}\,\underset{\xi}{\text{Max}}\,E(\xi,\eta) \geq \underset{\eta}{\text{Min}}\,E(\xi_\circ,\eta)$.

Again, since $\xi_\circ$ is any strategy, the preceding inequality holds even if we select $\xi_\circ$ which gives the maximum value of $\underset{\eta}{\text{Min}}\,E(\xi,\eta)$.

Therefore, $\underset{\eta}{\text{Min}}\;\underset{\xi}{\text{Max}}\,E(\xi,\eta) \geq \underset{\xi}{\text{Max}}\,\underset{\eta}{\text{Min}}\,E(\xi,\eta)$. This proves the theorem.

### Saddle point of a function

Let $E(\xi,\eta)$ be a function of two variables (vectors) $\xi$ and $\eta$ in $S_m$ and $S_n$ respectively. The point $(\xi_\circ,\eta_\circ)$, $\xi_\circ \in S_m$, $\eta_\circ \in S_n$ is said to be the saddle point of the function $E(\xi,\eta)$ if

$$E(\xi,\eta_\circ) \leq E(\xi_\circ,\eta_\circ) \leq E(\xi_\circ,\eta).$$

Now we shall discuss a theorem regarding the existence of the saddle point of a function.

**Theorem 13.4** *Let $E(\xi,\eta)$ be a function of two variables $\xi \in S_m$ and $\eta \in S_n$ such that* $\underset{\xi}{\text{Max}}\;\underset{\eta}{\text{Min}}\,E(\xi,\eta)$ *and* $\underset{\eta}{\text{Min}}\;\underset{\xi}{\text{Max}}\,E(\xi,\eta)$ *exist. Then the necessary and sufficient condition for the existence of a saddle point $(\xi_\circ,\eta_\circ)$ of $E(\xi,\eta)$ is that*

$$E(\xi_\circ,\eta_\circ) = \underset{\eta}{\text{Max}}\,\underset{\xi}{\text{Min}}\,E(\xi,\eta) = \underset{\eta}{\text{Min}}\,\underset{\xi}{\text{Max}}\,E(\xi,\eta).$$

*Proof* The condition is necessary; i.e. the point $(\xi_\circ,\eta_\circ)$ is the saddle point of $E(\xi,\eta)$.

Hence, from the definition of saddle point, we have

$$E(\xi,\eta_\circ) \leq E(\xi_\circ,\eta_\circ) \leq E(\xi_\circ,\eta) \tag{13.7}$$

for all $\xi \in S_m$ and $\eta \in S_n$.

From (13.7), clearly, $\underset{\xi}{\text{Max}}\,E(\xi,\eta_\circ) \leq E(\xi_\circ,\eta_\circ)$ holds for all $\xi \in S_m$.

$$\text{Hence, } \underset{\eta}{\text{Min}}\;\underset{\xi}{\text{Max}}\,E(\xi,\eta) \leq E(\xi_\circ,\eta_\circ). \tag{13.8}$$

Similarly, from (13.7), we have

$$E(\xi_\circ,\eta_\circ) \leq \underset{\eta}{\text{Min}}\,E(\xi_\circ,\eta),$$

$$\text{i.e. } E(\xi_\circ,\eta_\circ) \leq \underset{\xi}{\text{Max}}\,\underset{\eta}{\text{Min}}\,E(\xi,\eta). \tag{13.9}$$

From (13.8) and (13.9), we have

$$\underset{\eta}{\text{Min}}\,\underset{\xi}{\text{Max}}\,E(\xi,\eta) \leq \underset{\xi}{\text{Max}}\,\underset{\eta}{\text{Min}}\,E(\xi,\eta). \tag{13.10}$$

But, we know that

$$\underset{\eta}{\text{Min}}\,\underset{\xi}{\text{Max}}\,E(\xi,\eta) \geq \underset{\xi}{\text{Max}}\,\underset{\eta}{\text{Min}}\,E(\xi,\eta). \tag{13.11}$$

Hence, from (13.10) and (13.11), we have

$$E(\xi_\circ,\eta_\circ) = \underset{\xi}{\text{Max}}\,\underset{\eta}{\text{Min}}\,E(\xi,\eta) = \underset{\eta}{\text{Min}}\,\underset{\xi}{\text{Max}}\,E(\xi,\eta).$$

The condition is sufficient.

Let the point $(\xi_\circ,\eta_\circ)$ be such that

$$\underset{\xi}{\text{Max}}\,\underset{\eta}{\text{Min}}\,E(\xi,\eta) = \underset{\eta}{\text{Min}}\,\underset{\xi}{\text{Max}}\,E(\xi,\eta). \tag{13.12}$$

Also, let $\underset{\xi}{\text{Max}}\,\underset{\eta}{\text{Min}}\,E(\xi,\eta) = \underset{\eta}{\text{Min}}\,E(\xi_\circ,\eta)$

and $\underset{\eta}{\text{Min}}\,\underset{\xi}{\text{Max}}\,E(\xi,\eta) = \underset{\xi}{\text{Max}}\,E(\xi,\eta_\circ)$.

Hence, from (13.12), we have

$$\underset{\eta}{\text{Min}}\,E(\xi_\circ,\eta) = \underset{\xi}{\text{Max}}\,E(\xi,\eta_\circ). \tag{13.13}$$

Now, from the definition of maximum and minimum, we have

$$\underset{\eta}{\text{Min}}\,E(\xi_\circ,\eta) \leq E(\xi_\circ,\eta_\circ) \tag{13.14}$$

$$\text{and } E(\xi_\circ,\eta_\circ) \leq \underset{\xi}{\text{Max}}\,E(\xi,\eta_\circ). \tag{13.15}$$

From (13.13) and (13.14), we have

$$\underset{\xi}{\text{Max}}\,E(\xi,\eta_\circ) \leq E(\xi_\circ,\eta_\circ),$$

$$\text{which implies } E(\xi,\eta_\circ) \leq E(\xi_\circ,\eta_\circ) \text{ for all } \xi \in S_m. \tag{13.16}$$

From (13.13) and (13.15), we have

$$E(\xi,\eta_\circ) \leq \underset{\eta}{\text{Min}}\,E(\xi_\circ,\eta),$$

$$\text{which implies } E(\xi_\circ,\eta_\circ) \leq E(\xi_\circ,\eta) \text{ for all } \eta \in S_n. \tag{13.17}$$

Combining (13.16) and (13.17), we have

$$E(\xi, \eta_\circ) \le E(\xi_\circ, \eta_\circ) \le E(\xi_\circ, \eta),$$

which implies $(\xi_\circ, \eta_\circ)$ is a saddle point of $E(\xi, \eta)$.

***Fundamental theorem of a matrix game***

**Theorem 13.5** *The quantities* $\operatorname*{Max}_{\xi \in S_m}\left\{\operatorname*{Min}_{\eta \in S_n} E(\xi, \eta)\right\}$ *and* $\operatorname*{Min}_{\eta \in S_n}\left\{\operatorname*{Max}_{\xi \in S_m} E(\xi, \eta)\right\}$ *exist and are equal.*

*Proof* This theorem can easily be proved from Theorem 13.4.

***Value of a matrix game***

The common value of $\operatorname*{Max}_{\xi}\left\{\operatorname*{Min}_{\eta} E(\xi, \eta)\right\}$ and $\operatorname*{Min}_{\eta}\left\{\operatorname*{Max}_{\xi} E(\xi, \eta)\right\}$ is called the value of the matrix game with payoff matrix $A = \left[a_{ij}\right]$ and denoted by $v(A)$ or simply $v$.

**Definition** Equilibrium strategies of players in a matrix game are called their optimal strategies.

**Definition** The solution of a matrix game is the process of determining the value of the game and the pairs of optimal strategies.

***Note***: Thus, if $(\xi^*, \eta^*)$ is an equilibrium situation in mixed strategies of the game $\langle S_m, S_n, E\rangle$, then $\xi^*, \eta^*$ are the optimal strategies for players 1 and 2 respectively in the matrix game with payoff matrix $A = \left[a_{ij}\right]_{m\times n}$. Hence, $\xi^*, \eta^*$ are optimal strategies for players 1 and 2 respectively iff

$$E(\xi, \eta^*) \le E(\xi^*, \eta^*) \le E(\xi^*, \eta) \quad \forall \xi \in S_m, \eta \in S_n.$$

***Note***:

$$\operatorname*{Min}_{\xi} E(\xi, \eta) = E(\xi, \eta^*) \quad \therefore \operatorname*{Max}_{\xi}\left\{\operatorname*{Min}_{\eta} E(\xi, \eta)\right\} = \operatorname*{Max}_{\xi} E(\xi, \eta^*) = E(\xi^*, \eta^*)$$

$$\text{and } \operatorname*{Max}_{\xi} E(\xi, \eta) = E(\xi^*, \eta) \quad \therefore \operatorname*{Min}_{\eta}\left\{\operatorname*{Max}_{\xi} E(\xi, \eta)\right\} = \operatorname*{Min}_{\eta} E(\xi, \eta) = E(\xi^*, \eta^*).$$

***Notation***

The expected payoff to player 1 under $(\xi, \eta)$ is $E(\xi, \eta) = \sum_{i,j} a_{ij}\, x_i\, y_j$.

The expected payoff to player 1 under $\left(\xi, e_j^n\right)$ or simply $(\xi, j)$ is

$$E\left(\xi,\, e_j^n\right) = E(\xi,\, j) = \sum_i a_{ij} x_i.$$

The expected payoff to player 1 under $\left(e_i^m,\, \eta\right)$ or simply $(i,\, \eta)$ is

$$E\left(e_j^m,\, \eta\right) = E(i,\, \eta) = \sum_j a_{ij} y_j.$$

Thus, $E(\xi,\, \eta) = \sum_i E(i,\, \eta) = \sum_j E(\xi,\, j)$.

Also, $E\left(e_i^m,\, e_j^n\right) = E(i,\, j) = a_{ij}$.

***Some properties of the value of a matrix game***

The matrix game is characterized by the payoff matrix $A = \left[a_{ij}\right]_{m \times n}$ and $v$ is its value.

**Theorem 13.6** $v = \underset{\xi}{\text{Max}} \left\{ \underset{j}{\text{Min}}\ E(\xi,\, j) \right\} = \underset{\eta}{\text{Min}} \left\{ \underset{i}{\text{Max}}\ E(i,\, \eta) \right\}$, *and the outer extrema are attained at optimal strategies of the players.*

*Proof* We have, by definition, $v = \underset{\xi}{\text{Max}} \left\{ \underset{\eta}{\text{Min}}\ E(\xi,\, \eta) \right\}$.

For a fixed $\xi$, let $E(\xi,\, j_0) = \underset{j}{\text{Min}}\ E(\xi,\, j)$. Then

$$\begin{aligned} &E(\xi,\, j_0) \le E(\xi,\, j), \quad j = 1,\, 2,\, \ldots,\, n \\ &\text{or } \sum_j E(\xi,\, j_0) y_j \le \sum_j E(\xi,\, j) y_j, \\ &\text{i.e. } E(\xi,\, j_0) \le E(\xi,\, \eta) \text{ for a fixed } \xi \text{ and all } \eta \in S_n \end{aligned}$$

$$\text{Hence, } E(\xi,\, j_0) \le \underset{\eta}{\text{Min}}\ E(\xi,\, \eta). \tag{13.18}$$

Again since $\underset{\eta}{\text{Min}}\ E(\xi,\, \eta) \le E(\xi,\, \eta)\ \ \forall \eta \in S_n$, then setting $\eta = e_{j_0}^n$, we have

$$\underset{\eta}{\text{Min}}\ E(\xi,\, \eta) \le E\left(\xi,\, e_{j_0}^n\right) = E(\xi,\, j_0). \tag{13.19}$$

By (13.18) and (13.19), we have

$$\underset{\eta}{\text{Min}}\ E(\xi,\, \eta) = E(\xi,\, j_0) = \underset{j_0}{\text{Min}}\ E(\xi,\, j).$$

This is valid for any $\xi \in S_m$; hence, taking the maximum of both sides with respect to $\xi$, we have

$$\underset{\xi}{\text{Max}}\left\{\underset{\eta}{\text{Min}}\ E(\xi,\eta)\right\} = \underset{\xi}{\text{Max}}\left\{\underset{j}{\text{Min}}\ E(\xi,j)\right\},$$

where the outer maxima on both sides are attained for the same value of $\xi$. It is known, however, that on the left side the outer extremum is attained at the optimal strategies of player 1. Therefore, the same strategies yield the maximum on the right-hand side. Thus, the first part of the theorem is proved.

The proof of the second part is similar.

**Theorem 13.7**

$$\underset{i}{\text{Max}}\left\{\underset{j}{\text{Min}}\ a_{ij}\right\} \le v \le \underset{j}{\text{Min}}\left\{\underset{i}{\text{Max}}\ a_{ij}\right\}.$$

*Proof* By the preceding theorem, we have $v = \underset{\xi}{\text{Max}}\left\{\underset{j}{\text{Min}}\ E(\xi,j)\right\}\ \forall \xi \in S_m$.

But $\underset{\xi}{\text{Max}}\left\{\underset{j}{\text{Min}}\ E(\xi,j)\right\} \ge \underset{j}{\text{Min}}\ E(\xi,j) \quad \forall \xi \in S_m$

$\therefore v \ge \underset{j}{\text{Min}}\, E(\xi,j) \quad \forall \xi \in S_m$.

Setting $\xi = e_i^m$, we have $v \ge \underset{j}{\text{Min}}\ E\left(e_i^m, j\right) = \underset{j}{\text{Min}}\ E(i,j) = \underset{j}{\text{Min}}\ a_{ij}$, i.e. $v \ge \underset{j}{\text{Min}}\ a_{ij}$.

The left side $v$ is independent of $i$, so taking the maximum with respect to $i$, we obtain

$$v \ge \underset{i}{\text{Max}}\left\{\underset{j}{\text{Min}}\, a_{ij}\right\}.$$

The proof of the second part is similar.

**Theorem 13.8**

(i) *If player 1 possesses a pure optimal strategy* $i^*$, *then*

$$v = \underset{i}{\text{Max}}\left(\underset{j}{\text{Min}}\ a_{ij}\right) = \underset{j}{\text{Min}}\ a_{i^*j}.$$

(ii) *If player 2 possesses a pure optimal strategy* $j^*$, *then*

$$v = \underset{j}{\text{Min}}\left(\underset{i}{\text{Max}}\ a_{ij}\right) = \underset{i}{\text{Max}}\ a_{ij^*}$$

*Proof* We know that

$$\begin{aligned} v &= \underset{\xi}{\text{Max}}\ \underset{j}{\text{Min}}\ E(\xi, j) \\ &= \underset{j}{\text{Min}}\ E\left(e_{i^*}^m, j\right) \text{as } \xi^* = e_{i^*}^m \text{ is optimal.} \end{aligned}$$

The proof of the second part is similar.

***Solution of matrix game***

Let us consider a $2 \times 2$ matrix game whose payoff matrix $A$ is given by $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$.

If $A$ has a saddle point, the solution is obvious.

Let $A$ have no saddle point. Let player 1 have the strategy $\xi = (x_1, x_2)' \equiv (x, 1-x)(0 \le x \le 1)$ and player 2 have the strategy $\eta = (y, 1-y)'(0 \le y \le 1)$.

$$\begin{aligned} \text{Then } E(\xi, \eta) &= \sum_{i=1}^{2} \sum_{j=1}^{2} a_{ij} x_i y_j \\ &= a_{11} xy + a_{12} x(1-y) + a_{21}(1-x)y + a_{22}(1-x)(1-y) \\ &= f(x, y), \text{ say.} \end{aligned}$$

If $\xi^* = (x^*, 1-x^*)'$, $\eta^* = (y^*, 1-y^*)$ are optimal strategies, then from

$$E(\xi, \eta^*) \le E(\xi^*, \eta^*) \le E(\xi^*, \eta) \ \ \forall \xi \in S_1, \eta \in S_2$$

we have $f(x, y^*) \le f(x^*, y^*) \le f(x^*, y) \quad \forall x \in (0, 1),\ y \in (0, 1)$.

From the first part of the inequality, we set that $f(x, y^*)$ regarded as a function of $x$ has a maximum at $x^*$, i.e.

$$\left.\frac{\partial f}{\partial x}\right|_{(x^*, y^*)} = 0.$$

This gives $a_{11}y^* + a_{12}(1-y^*) - a_{21}y^* - a_{22}(1-y^*) = 0$

or $(a_{11} - a_{12} - a_{21} + a_{22})y^* = a_{22} - a_{12}$

or $y^* = \dfrac{a_{22} - a_{12}}{a_{11} + a_{22} - (a_{12} + a_{21})}$ provided that $a_{11} + a_{22} - (a_{12} + a_{21}) \neq 0$.

Similarly, from the second part of the inequality, it is seen that $f(x^*, y)$ regarded as a function of $y$ has a minimum at $y^*$, i.e.

$$\left.\frac{\partial f}{\partial y}\right|_{(x^*,\,y^*)} = 0.$$

This gives $a_{11}x^* - a_{12}x^* + a_{21}(1 - x^*) - a_{22}(1 - x^*) = 0$

$$\text{or } x^* = \frac{a_{22} - a_{21}}{a_{11} + a_{22} - (a_{12} + a_{21})} \text{ provided that } a_{11} + a_{22} - (a_{12} + a_{21}) \neq 0.$$

$$\begin{aligned}\text{Now } v^* &= f(x^*,\, y^*) \\ &= a_{11}x^*y^* + a_{12}x^*(1 - y^*) + a_{21}(1 - x^*)y^* + a_{22}(1 - x^*)(1 - y^*) \\ &= \frac{a_{11}\,a_{22} - a_{12}\,a_{21}}{a_{11} + a_{22} - (a_{12} + a_{21})}.\end{aligned}$$

It can be proved that $a_{11} + a_{22} - (a_{12} + a_{21}) = 0$ implies that $A$ has a saddle point.

***Some properties of optimal strategies in a matrix game***

Let $A = \left[a_{ij}\right]_{m\times n}$ be a payoff matrix of a game whose value is $v$.

**Theorem 13.9** *Let $u$ be an arbitrary real number. Then*

(i) $u \le E(\xi,\, j), \quad j = 1,\, 2,\, \ldots,\, n$ *for some fixed* $\xi$ *implies* $u \le v$.
(ii) $u \ge E(i,\, \eta), \quad i = 1,\, 2,\, \ldots,\, m$ *for some fixed* $\eta$ *implies* $u \ge v$.

*Proof* Let $\eta^*$ be an optimal strategy for player 2. Now we have $u \le E(\xi,\, j)$ $j = 1,\, \ldots,\, n$ for some fixed $\xi$.

$$\therefore \quad u\sum_{j=1}^{n} y_j^* \le \sum_{j=1}^{n} E(\xi,\, j)y_j^*,$$

i.e. $u \le E(\xi,\, \eta^*) \le E(\xi^*,\, \eta^*),$ where $\xi^*$ is an optimal strategy for player 1,

i.e. $u \le v$

The proof of the second part is similar.

**Theorem 13.10** *Let $u$ be a real number such that $E(i,\, \eta_0) \le u \le E(\xi_0,\, j)$ for $i = 1,\, 2,\, \ldots,\, m,\, j = 1,\, 2,\, \ldots,\, n$ and some particular $\xi_0,\, \eta_0$. Then $u$ is the value of the game and $\xi_0,\, \eta_0$ are the optimal strategies for players 1 and 2 respectively.*

*Proof* Given that $E(i,\, \eta_0) \le u, \;\; (i = 1,\, 2,\, \ldots,\, m)$ and for some particular $\eta_0$,

$$\therefore \sum_i E(i,\, \eta_0)x_i^* \le u\sum_i x_i^*$$

$$\text{or } E(\xi_0,\, \eta_0) \le u. \tag{13.20}$$

Again $u \ge E(\xi_0,\, j), \quad j = 1,\, 2,\, \ldots,\, n$ gives similarly

$$u \geq E(\xi_0, \eta_0). \tag{13.21}$$

Hence, (13.20) and (13.21) imply that $u = E(\xi_0,\ \eta_0)$.

$$\begin{aligned}
&\text{Again we have } E(i,\ \eta_0) \leq u = E(\xi_0,\ \eta_0), \quad i = 1,\ 2,\ \ldots,\ m \\
&\text{so that } \sum_i E(i,\ \eta_0)x_i \leq E(\xi_0,\ \eta_0)\sum_i x_i \\
&\text{or } E(\xi,\ \eta_0) \leq E(\xi_0,\ \eta_0) \quad \forall \xi \in S_m.
\end{aligned} \tag{13.22}$$

$$\begin{aligned}
&\text{Again } u = E(\xi_0,\ \eta_0) \leq E(\xi_0, j), \quad j = 1,\ 2,\ \ldots,\ n \\
&\text{so that } E(\xi_0,\ \eta_0)\sum_j y_j \leq \sum_j E(\xi_0, j)y_j \\
&\text{or } E(\xi_0,\ \eta_0) \leq E(\xi_0,\ \eta) \quad \forall \eta \in S_n.
\end{aligned} \tag{13.23}$$

Hence, from (13.22) and (13.23), we have

$$E(\xi,\ \eta_0) \leq E(\xi_0,\ \eta_0) \leq E(\xi_0,\ \eta) \quad \forall \xi \in S_m, \eta \in S_n.$$

Hence, by the definition of optimal strategies, $\xi_0,\ \eta_0$ are optimal strategies and $E(\xi_0,\ \eta_0) = u$ is the value of the game.

**Theorem 13.11** *If* $\underset{i}{\text{Max}}\ E(i,\ \eta_0) \leq \underset{j}{\text{Min}}\ E(\xi_0,\ j)$ *is valid for some fixed* $\xi_0$ and $\eta_0$, *then* $\xi_0,\ \eta_0$ *are optimal strategies and the inequality becomes equality.*

*Proof* We have $E(i,\ \eta_0) \leq \underset{i}{\text{Max}}\ E(i,\ \eta_0) \quad \forall i = 1,\ 2,\ \ldots,\ m$ [by the definition of maximum]

$\leq \underset{j}{\text{Min}}\ E(\xi_0,\ j)$ [by the given condition]

$\leq E(\xi_0, j) \quad \forall j = 1,\ 2,\ \ldots,\ n$ [by the definition of minimum]

$\forall i = 1,\ 2,\ \ldots,\ m$ and $j = 1,\ 2,\ \ldots,\ n$.

Let $u_1 = \text{Max}_i\ E(i,\ \eta_0)$.

Then $E(i,\ \eta_0) \leq u_1 \leq E(\xi_0,\ j) \quad \forall i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

Hence, by Theorem 13.10, $u_1$ is the value of the game and $\xi_0,\ \eta_0$ are the optimal strategies.

Again let $u_2 = \underset{j}{\text{Min}}\ E(\xi_0, j)$.

Now by the definition of minima, we have

$$\begin{aligned}
E(\xi_0,\ j) &\geq \underset{j}{\text{Min}}\ E(\xi_0, j) \quad \forall j = 1, 2, \ldots, n \\
&\geq \underset{i}{\text{Max}}\ E(i,\ \eta_0) \text{ [by the given condition]} \\
&\geq E(i,\ \eta_0) \quad \forall i = 1,\ 2,\ \ldots,\ m \text{ [by the definition of maxima]}
\end{aligned}$$

$$\therefore E(\xi_0, j) \geq \underset{j}{\text{Min}}\, E(\xi_0, j) \geq E(i, \eta_0).$$

$$\text{i.e. } E(i, \eta_0) \leq \underset{j}{\text{Min}}\; E(\xi_0, j) \leq E(\xi_0, j)$$

$$\text{or } E(i, \eta_0) \leq u_2 \leq E(\xi_0, j)$$

Hence, by Theorem 13.10, $u_2$ is the value of the game and $\xi_0$, $\eta_0$ are the optimal strategies. But the value of the game is unique. Hence, $u_1 = u_2$; i.e. the inequality becomes an equality.

**Theorem 13.12** *If* $E(i, \eta_0) \leq E(\xi_0, j) \quad \forall i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n,$ *then* $\xi_0, \eta_0$ *are the optimal strategies.*

*Proof* This theorem can be proved from Theorem 13.3.

**Theorem 13.13**

(i) *If for a strategy* $\xi_0$ *of player 1,* $v \leq E(\xi_0, j), \quad \forall j = 1, 2, \ldots, n,$ *then* $\xi_0$ *is an optimal strategy for player 1.*
(ii) *If for a strategy* $\eta_0$ *of player 2,* $v \geq E(i, \eta_0), \quad \forall i = 1, 2, \ldots, m,$ *then* $\eta_0$ *is an optimal strategy for player 2.*

*Proof* Let $\eta^*$ be an optimal strategy for player 2, then $E(\xi, \eta^*) \leq v, \ \forall \xi \in S_m$. Setting $\xi = e_i^m$, we have $E(i, \eta^*) \leq v \quad \forall i = 1, 2, \ldots, m.$

But $v \leq E(\xi_0, j)$ (given). Hence, $E(i, \eta^*) \leq v \leq E(\xi_0, j) \quad \forall i = 1, 2, \ldots, m, \forall j = 1, 2, \ldots, n.$

By Theorem 13.2, it follows that $\xi_0$ is an optimal strategy for player 1.

The proof of the second part is similar.

**Theorem 13.14**

(i) *If* $v \leq \underset{j}{\text{Min}}\; E(\xi_0, j)$ *for a fixed* $\xi_0$, $\xi_0$ *is an optimal strategy for player 1.*
(ii) *If* $v \geq \underset{i}{\text{Max}}\; E(i, \eta_0)$ *for a fixed* $\eta_0$, $\eta_0$ *is an optimal strategy for player 2.*

*Proof*

(i) $v \leq \underset{j}{\text{Min}}\; E(\xi_0, j) \leq E(\xi_0, j) \quad \forall j = 1, 2, \ldots, n..$

Then by Theorem 13.5, this theorem can be proved.

**Corollary**

(i) *A necessary and sufficient condition for the optimality of strategy* $\xi_0$ *for player 1 is* $v = \underset{j}{\text{Min}}\, E(\xi_0, j)$.
(ii) *A similar condition for player 2 is* $v = \underset{i}{\text{Max}}\, E(i, \eta_0)$.

## 13.6 Dominance

In a game, if one course of action (pure strategy) of one player is better than or as good as another, for all possible courses of action of the opponent, then the first course of action is said to dominate the second. In this situation, the dominated (or inferior) course of action can simply be discarded from the payoff matrix, and a reduced matrix is obtained. In this case, the optimal strategies for the reduced matrix are also optimal for the original matrix with zero probability for discarded strategies. When there is no saddle point in a payoff matrix, then the size of the game can be reduced by this process, which is known as dominance.

For example, let us consider a game with the following payoff matrix:

$$\begin{array}{cc} & \textit{Player } B \\ & \begin{array}{ccc} B_1 & B_2 & B_3 \end{array} \\ \textit{Player } A \quad \begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} & \begin{array}{|ccc|} \hline 4 & 9 & 4 \\ 7 & 4 & 8 \\ 8 & 5 & 9 \\ \hline \end{array} \end{array}$$

From this payoff matrix, it is clear that the second row is inferior to the third row for the maximizing player $A$, as every element of the third row is either greater than or equal to the corresponding element of the second row. Then by deleting the second row, we get the reduced payoff matrix as follows:

$$\begin{array}{cc} & \begin{array}{ccc} B_1 & B_2 & B_3 \end{array} \\ \begin{array}{c} A_1 \\ A_3 \end{array} & \begin{array}{|ccc|} \hline 4 & 7 & 4 \\ 8 & 5 & 9 \\ \hline \end{array} \end{array}$$

Again, from player $B$'s point of view, the third strategy $(B_3)$ is dominated by $B_1$, as every element of the third column is either greater than or equal to the corresponding element of the first column. Then by deleting the third column, we get the reduced payoff matrix as follows:

$$\begin{array}{cc} & \begin{array}{cc} B_1 & B_2 \end{array} \\ \begin{array}{c} A_1 \\ A_3 \end{array} & \begin{array}{|cc|} \hline 4 & 7 \\ 8 & 5 \\ \hline \end{array} \end{array}$$

Now, if $\left(x_1^*,\ x_3^*\right)$ and $\left(y_1^*,\ y_2^*\right)$ are the optimal strategies for this reduced matrix, then $\left(x_1^*,\ 0,\ x_3^*\right)$ and $\left(y_1^*,\ y_2^*,\ 0\right)$ are the optimal strategies for the original payoff matrix.

**Theorem 13.15** *If the ith row of the payoff matrix of an $m \times n$ rectangular game is dominated by its rth row, then the deletion of the ith row from the matrix does not change the optimal strategies of the player whose objective is to maximize the gain.*

*Proof* Let $A$ and $B$ be the players of a rectangular game whose payoff matrix is given by the matrix $\left[a_{ij}\right]_{m\times n}$.

According to the given condition of the theorem,

$$a_{ij} \leq a_{rj} \text{ for } j = 1, 2, \ldots, n \quad \text{and} \quad i \neq r, \tag{13.24}$$

and for at least one $j$, $a_{ij} \neq a_{rj}$.

Let $\eta^* = \left(y_1^*, y_2^*, \ldots, y_n^*\right)$ be the optimal mixed strategy for player $B$. Then from (13.24), we have

$$\begin{gathered}\sum_{j=1}^{n} a_{ij}\, y_j^* < \sum_{j=1}^{n} a_{rj}\, y_j^*, \\ \text{i.e. } E(A_i,\, \eta^*) < E(A_r,\, \eta^*),\end{gathered} \tag{13.25}$$

where $A_i$ and $A_r$ are the $i$th and $r$th pure strategies of $A$.

Let $v$ be the value of the game.

Hence, from (13.25), we have

$$v \geq E(A_r,\, \eta^*) > E(A_i,\, \eta^*). \tag{13.26}$$

Let $\xi^* = \left(x_1^*, x_2^*, \ldots, x_m^*\right)$ be an optimal strategy for player $A$.

If possible, let $x_i^* > 0$.

Then, from (13.26), we have

$$x_i^* v > x_i^* E(A_i, \eta_0).$$

Since $(\xi_0,\, \eta_0)$ is the optimal solution of the game problem, we have

$$\begin{aligned} v = E(\xi^*,\, \eta^*) &= \sum_{k=1}^{m} x_k^*\, E(A_k,\, \eta^*) \\ &= x_i^*\, E(A_i,\, \eta^*) + \sum_{\substack{k=1 \\ k\neq i}}^{m} x_k^*\, E(A_k,\, \eta^*) < x_i^* v + v \sum_{\substack{k=1 \\ k\neq i}}^{m} x_k^* = v \sum_{k=1}^{m} x_k^* = v \end{aligned}$$

i.e. $v < v$.

This is a contradiction to our assumption that $x_i^* > 0$ and hence $x_i^* = 0$.

Hence, the deletion of the $i$th row does not alter the optimal solution.

**Theorem 13.16** *If the jth column of the payoff matrix of an $m \times n$ rectangular game dominates its kth column, then the deletion of the jth column from the matrix does not change the optimal strategy of the player whose objective is to minimize the loss.*
[The proof of this theorem is similar to that of the preceding theorem.]

**Theorem 13.17** *If the ith row of the payoff matrix of an $m \times n$ rectangular game is strictly dominated by a convex combination of the other rows of the matrix, then the deletion of the ith row from the matrix does not affect the optimal strategy of the player whose objective is to maximize the gain.*

*Proof* Let $A$ and $B$ be the players of a rectangular game whose payoff matrix is given by the matrix $\left[a_{ij}\right]_{m\times n}$.

Let $(x_1, x_2, \ldots, x_m)$ be the mixed strategy of player $A$ where $\sum_k x_k = 1$.

According to the given condition,

$$\begin{aligned} &x_i = 0 \text{ and } a_{ij} \le \sum_{\substack{k=1\\k\neq i}}^{m} a_{kj}x_k, \quad j = 1, 2, \ldots, n \\ &\text{or } a_{ij} \le \sum_{k=1}^{m} a_{kj}x_k \end{aligned} \tag{13.27}$$

in which strict inequality holds for at least one $j$.

Let $\eta^* = \left(y_1^*, y_2^*, \ldots, y_n^*\right)$ be the optimal strategy for player $B$, whose objective is to minimize the loss. Hence, from (13.27), we have

$$\begin{aligned} &\sum_{j=1}^{n} a_{ij}\, y_j^* < \sum_{j=1}^{n}\sum_{k=1}^{m} a_{kj}\, x_k\, y_j^* \\ &\text{or } E(A_i, \eta^*) < \sum_{j=1}^{n}\sum_{k=1}^{m} a_{kj}\, x_k\, y_j^* \le v, \end{aligned} \tag{13.28}$$

where $A_i$ is the $i$th pure strategy of $A$ and $v$ is the value of the game.

Let $\xi^* = \left(x_1^*, x_2^*, \ldots, x_m^*\right)$ be the optimal strategy of player $A$, whose objective is to maximize the gain.

If possible, let $x_i^* \neq 0$.

From (13.28), we have

$$x_i^* E(A_i, \eta^*) < x_i^* v, \quad \text{since } x_i^* \neq 0.$$

Hence, we have

$$v = E(\xi^*, \eta^*) = \sum_{j=1}^{n} \sum_{k=1}^{m} a_{kj} x_k^* y_j^*$$

$$= x_i^* E(A_i, \eta^*) + \sum_{\substack{k=1 \\ k \neq i}}^{m} x_k^* \sum_{j=1}^{n} a_{kj} y_j^*$$

$$= x_i^* E(A_i, \eta^*) + \sum_{\substack{k=1 \\ k \neq i}}^{m} x_k^* E(A_k, \eta^*) < x_i^* v + v \sum_{\substack{k=1 \\ k \neq i}}^{m} x_k^* = v \sum_{k=1}^{m} x_k^* = v.$$

This is a contradiction to our assumption that $x_i^* \neq 0$. Thus, $x_i^* = 0$.

Hence, the deletion of the $i$th row does not affect the optimal solution.

**Theorem 13.18** *If the jth column of the matrix strictly dominates a convex combination of other columns, then the deletion of the jth column of the matrix does not affect the optimal strategy of the player whose objective is to minimize the loss.* [The proof of this theorem is similar to that of Theorem 13.3.]

**Theorem 13.19** *$R_1$ and $R_2$ are the subsets of the rows and $C_1$ and $C_2$ are the subsets of the columns of the payoff matrix $\left[a_{ij}\right]_{m \times n}$ of a matrix game.*

(i) *If a convex combination of the rows in $R_1$ strictly dominates a convex combination of the rows in $R_2$, then there exists a row in $R_2$ which, if discarded, does not change the optimal strategy of the player whose objective is to maximize the gain.*
(ii) *If a convex combination of the columns in $C_1$ dominates a convex combination of the columns in $C_2$, then there exists a column in $C_1$ which, if discarded, does not change the optimal strategy of the player whose objective is to minimize the loss.*

***The Rules of dominance***

Sometimes, in a rectangular game, it is seen that one or more pure strategies of a player are inferior to at least one of the remaining strategies. In such a case, this inferior strategy is never used. In other words, we can say that this inferior pure strategy is dominated by a superior pure strategy. In such cases of dominance, we can reduce the size of the payoff matrix by removing the pure strategies which are dominated by other strategies. Hence, the rules of dominance are used to reduce the size of the payoff matrix. These rules are especially used for solving a two-person zero-sum game without a saddle point.

***Rule 1***: If each element in a row (say, the $i$th row, i.e. $R_i$) of a payoff matrix is either less than or equal to the corresponding element in another row (say, the $j$th row, i.e. $R_j$), then the $i$th strategy is dominated by the $j$th strategy, and that row (i.e. the $i$th row) can be deleted from the payoff matrix.

In other words, player $A$ will never use that strategy, because if player $A$ chooses that strategy, then he will gain less payoff.

***Rule 2***: If each element in a column (say, the $i$th column, i.e. $C_i$) of a payoff matrix is either greater than or equal to the corresponding element in another column (say, the $j$th column, i.e. $C_j$), then the $i$th strategy is dominated by the $j$th strategy, and that column (i.e. the $i$th column) can be deleted from the payoff matrix.

***Rule 3***: If the $i$th row is dominated by a convex combination of some/all other rows, then the $i$th row is deleted from the payoff matrix. If the $j$th column dominates a convex combination of some/all other columns, then the $j$th column is deleted from the payoff matrix.

***Rule 4***: If $R_1$, $R_2$ are the subsets of the rows of the payoff matrix, and if a convex combination of rows in $R_1$ dominates a convex combination of the rows in $R_2$, then there exists a row in $R_2$ which is deleted.

If $C_1$, $C_2$ are the subsets of the columns of the payoff matrix, and if a convex combination of the columns in $C_1$ dominates a convex combination of the columns in $C_2$, then there exists a column in $C_1$ which is deleted.

Rules 3 and 4 are called the modified dominance property. We note that the element or probability value corresponding to the deleted strategy by the rules of dominance is taken as zero.

*Example 4* Solve the game whose payoff matrix is given by

$$\text{Player } A \quad \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \begin{array}{c} \text{Player } B \\ \begin{array}{cccc} B_1 & B_2 & B_3 & B_4 \end{array} \\ \begin{bmatrix} 1 & 2 & -2 & 2 \\ 3 & 1 & 2 & 3 \\ -1 & 3 & 2 & 1 \\ -2 & 2 & 0 & -3 \end{bmatrix} \end{array}$$

**Solution** First of all, we shall compute the maximin and minimax values in the case of a pure strategy.

| | $B_1$ | $B_2$ | $B_3$ | $B_4$ | Row minima |
|---|---|---|---|---|---|
| $A_1$ | 1 | 2 | −2 | 2 | −2 |
| $A_2$ | 3 | 1 | 2 | 3 | 1 |
| $A_3$ | −1 | 3 | 2 | 1 | −1 |
| $A_4$ | −2 | 2 | 0 | −3 | −3 |
| Column maxima | 3 | 3 | 2 | 3 | |

In the case of a pure strategy, maximin value = 1 and minimax value = 2. As maximin value $\neq$ minimax value, this game has no saddle point for a pure strategy.

Hence, this game can be solved for the mixed strategy. For this purpose, we shall try to reduce the size of the payoff matrix by using the dominance rule.

Since every element of the fourth row is less than the corresponding elements of the third row, then from player $A$'s point of view, the fourth strategy, i.e. $A_4$, is dominated by the third strategy ($A_3$); thus, we can delete the fourth row. In this situation, the optimal strategy will not be affected. Now deleting the fourth row, we get the reduced payoff matrix as follows:

$$\begin{array}{c c} & \begin{array}{cccc} B_1 & B_2 & B_3 & B_4 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} & \begin{bmatrix} 1 & 2 & -2 & 2 \\ 3 & 1 & 2 & 3 \\ -1 & 3 & 2 & 1 \end{bmatrix} \end{array}$$

Again, from player $B$'s point of view, the fourth strategy ($B_4$) is dominated by $B_1$, as every element of the fourth column is either greater than or equal to the corresponding element of the first column. Then by deleting the fourth column, we get the reduced payoff matrix as follows:

$$\begin{array}{c c} & \begin{array}{ccc} B_1 & B_2 & B_3 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} & \begin{bmatrix} 1 & 2 & -2 \\ 3 & 1 & 2 \\ -1 & 3 & 2 \end{bmatrix} \end{array}$$

From the reduced payoff matrix, it is seen that none of the pure strategies of players $A$ and $B$ is inferior to any of their other strategies. However, the convex combination due to strategies $A_2$ and $A_3$ (i.e. the average of the second and third rows, 1, 2, 2) is superior to the payoff due to strategy $A_1$. Thus, strategy $A_1$ may be deleted, giving us the reduced payoff matrix as follows:

$$\begin{array}{c c} & \begin{array}{ccc} B_1 & B_2 & B_3 \end{array} \\ \begin{array}{c} A_2 \\ A_3 \end{array} & \begin{bmatrix} 3 & 1 & 2 \\ -1 & 3 & 2 \end{bmatrix} \end{array}$$

In this reduced matrix, the convex combination due to strategies $B_1$ and $B_2$ is superior to the payoff due to strategy $B_3$. Thus, strategy $B_3$ may be deleted, and we will get the reduced $2 \times 2$ payoff matrix as follows:

$$\begin{array}{c c} & \begin{array}{cc} B_1 & B_2 \end{array} \\ \begin{array}{c} A_2 \\ A_3 \end{array} & \begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix} \end{array}$$

Now, we have to solve this reduced game whose payoff matrix is of $2 \times 2$ order.

Clearly, this game has no saddle point and cannot be reduced further. Therefore, the optimal strategies will be mixed strategies.

Let player $A$ choose his strategies $A_2$ and $A_3$ with probabilities $x_2$ and $x_3$; those of $B$ are $y_1$ and $y_2$.

$\therefore x_2 + x_3 = 1$, $y_1 + y_2 = 1$ and $x_2$, $x_3$, $y_1$, $y_2$ >0 and $v$ is the value of the game.

To determine the optimal values of $x_2$ and $x_3$, we have

$$3x_2 - x_3 = x_2 + 3x_3 = v,$$

which implies $x_2 = \frac{2}{3} \therefore x_1 = 1 - x_2 = \frac{1}{3}$.

To determine the optimal values of $y_1$ and $y_2$, we have

$$3y_1 + y_2 = -y_1 + 3y_2$$

which implies $y_1 = \frac{1}{3} \therefore y_2 = \frac{2}{3}$.

The value of the game $v = 3x_2 - x_3 = 3 \cdot \frac{2}{3} - \frac{1}{3} = \frac{5}{3}$.

Hence, the solution of the game is as follows:

Optimal strategies are $\xi^* = \left(0, \frac{2}{3}, \frac{1}{3}, 0\right)$, $\eta^* = \left(\frac{1}{3}, \frac{2}{3}, 0, 0\right)$, and the value of the game is $\frac{5}{3}$.

***An important result***

If a constant is added to all the elements of a matrix game, the optimal strategies remain unaltered, and the value of the game is increased by this constant.

This result can easily be proved.

## 13.7 Graphical Solution of 2 × *n* or *m* × 2 Games

Any rectangular game of order 2 × $n$ or $m$ × 2 can be reduced to a 2 × 2 game by using a graphical method, and the reduced game of order 2 × 2 can be solved by an algebraic method.

Let us consider a 2 × $n$ game without a saddle point. Its payoff matrix is as follows:

$$\begin{array}{cc} & \text{Player } B \\ & \begin{array}{cccc} B_1 & B_2 & \cdots & B_n \end{array} \\ \text{Player } A \begin{array}{c} A_1 \\ A_2 \end{array} & \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \end{bmatrix} \end{array}$$

Let $\xi$ = ($x_1$, $x_2$) and $\eta$ = ($y_1$, $y_2$, …, $y_n$) be the mixed strategies for players $A$ and $B$ respectively. When player $B$ uses his pure strategy $B_j$, then the expected gain of player $A$ is given by

$$E(\xi, B_j) = a_{1j}x_1 + a_{2j}x_2 = a_{1j}x_1 + a_{2j}(1 - x_1),\ j = 1, 2, \ldots, n \tag{13.29}$$

as $x_1 + x_2 = 1$.

Clearly, both $x_1$ and $x_2$ must lie in the open interval (0, 1) [because if either $x_1 = 1$ or $x_2 = 1$, the game reduces to a game of pure strategy, which is against our assumption]. Hence $Ej$ ($\xi$) is a linear function of either $x_1$ or $x_2$. Considering $E_j$ as a linear function of $x_1$ (say), we have

$$\begin{aligned} E_j\left(\xi_2, B_j\right) &= a_{2j} \text{ for } x_1 = 0 \\ &= a_{1j} \text{ for } x_1 = 1 \end{aligned}$$

Hence, $E_j(\xi, B_j)$ represents a line segment joining the points (0, $a_{2j}$) and (1, $a_{1j}$).

Now player $A$ expects a least possible gain $v$. Therefore, $E_j(\xi, B_j) \geq v$ for all $j$. Now considering the strict equations for inequalities, and with the help of a graphical method, we shall find out two particular moves or choices of $B$ which will maximize the minimum gain of $A$.

Let us draw two parallel lines $x_1 = 0$ and $x_1 = 1$ at unit distance apart. Now draw $n$ line segments joining the points (0, $a_{2j}$) and (1, $a_{1j}$), $j$ = 1, 2, …, $n$. The lower envelope (or lower boundary) of these line segments (indicated by a thick line segment) will give the minimum expected gain of $A$ as a function of $x_1$. Now the highest point of the lower envelope will give the maximum of minimum gain of $A$. The line segments passing through the point corresponding to $B$'s two pure moves, say, $B_k$ and $B_l$, are the critical moves for $B$ which will maximize the minimum expected gain of $A$. Now the 2 × 2 payoff matrix corresponding to $A$'s moves $A_1$, $A_2$ and $B$'s moves $B_k$, $B_l$ will produce the required result. Thus, solving the 2 × 2 game algebraically, we can find the value of the game.

If there are more than two line segments passing through the highest (maximin) point, there are ties for the optimum mixed strategies for player $B$. Thus, any two such lines with opposite sign slopes will determine an alternative optimum for player $B$.

Again, if there is more than one maximin point, alternative optima exist corresponding to these points.

*Example 5* Solve the following 2 × 4 game graphically:

$$\text{Player } A \quad \begin{array}{c} \\ A_1 \\ A_2 \end{array} \begin{array}{c} \text{Player } B \\ \begin{array}{cccc} B_1 & B_2 & B_3 & B_4 \end{array} \\ \begin{bmatrix} 1 & 3 & -3 & 7 \\ 2 & 5 & 4 & -6 \end{bmatrix} \end{array}$$

*Solution* Clearly, the given problem does not possess any saddle point in the case of pure strategy. Hence, this problem can be solved with a mixed strategy.

Let player $A$ play with mixed strategy $\xi = (x_1, x_2)$, where $x_1 + x_2 = 1$ and both $x_1$ and $x_2$ lie in the open interval (0, 1).

Then player $A$'s expected gains against $B$'s pure moves are given by

| B′s Pure strategy | A′s expected gain |
|---|---|
| $B_1$ | $E(x_1, B_1) = x_1 + 2x_2 = x_1 + 2(1 - x_1)$ |
| $B_2$ | $E(x_1, B_2) = 3x_1 + 5(1 - x_1)$ |
| $B_3$ | $E(x_1, B_3) = -3x_1 + 4(1 - x_1)$ |
| $B_4$ | $E(x_1, B_4) = -7x_1 + 6(1 - x_1)$ |

Draw two vertical lines $x_1 = 0$ and $x_1 = 1$ at unit distance apart. Mark the lines $x_1 = 0$ and $x_1 = 1$ by using the same scale as given in Fig. 13.1. Now draw the line segments for the expected gain equations between the two vertical lines $x_1 = 0$ and $x_1 = 1$. These line segments represent $A$'s expected gain due to $B$'s pure move. We denote these line segments as $B_1$, $B_2$, $B_3$, $B_4$. Since player $A$ wishes to maximize his minimum expected gain, the highest point of intersection $H$ of two line segments $B_3$ and $B_4$ represents the maximin expected value of the game for player $A$. Hence, the solution to the original game reduces to that of a simpler game with the following $2 \times 2$ payoff matrix:

$$\begin{array}{ccc} & & \text{Player } B \\ & & \begin{array}{cc} B_3 & B_4 \end{array} \\ \text{Player } A & \begin{array}{c} A_1 \\ A_2 \end{array} & \begin{bmatrix} -3 & 7 \\ 4 & -6 \end{bmatrix} \end{array}$$

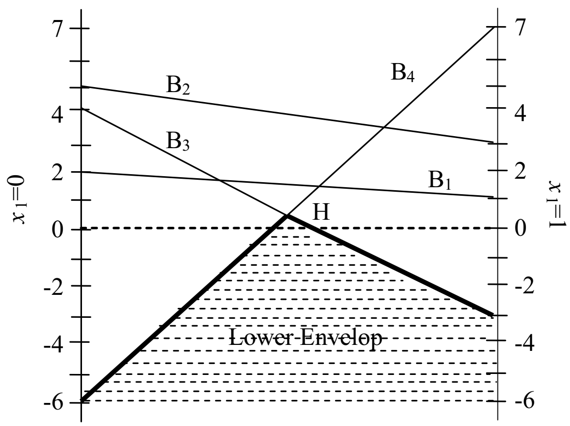



**Fig. 13.1** Graphical representation of 2 × 4 game

Now, if $\xi = (x_1, x_2)$ and $\eta = (y_3, y_4)$ are the optimum strategies for players $A$ and $B$, then we have

$$x_1 = \frac{a_{22} - a_{21}}{(a_{11} + a_{22}) - (a_{12} + a_{21})} = \frac{-6 - 4}{-3 - 6 - (7+4)} = \frac{-10}{-20} = \frac{1}{2}$$
$$\therefore x_2 = 1 - x_1 = \frac{1}{2}.$$

Again, $y_3 = \frac{-6-7}{-3-6-(7+4)} = \frac{13}{20}$, $y_4 = 1 - y_3 = \frac{7}{20}$ and the value of the game is given by

$$v = \frac{a_{11}a_{22} - a_{12}a_{21}}{(a_{11} + a_{22}) - (a_{12} + a_{21})} = \frac{18 - 28}{-20} = \frac{1}{2}.$$

Hence, the solution of the game is as follows:

(i) Optimal strategies $\xi^* = \left(\frac{1}{2}, \frac{1}{2}\right)$, $\eta^* = \left(0, 0, \frac{13}{20}, \frac{7}{20}\right)$.
(ii) The value of the game is $\frac{1}{2}$.

*Example 6* By the graphical method, solve the game whose payoff matrix is given by

$$\text{Player } A \quad \begin{array}{c} \\ A_1 \\ A_2 \end{array} \overset{\text{Player } B}{\begin{array}{cccc} B_1 & B_2 & B_3 & B_4 \\ \end{array}} \begin{bmatrix} 2 & 2 & 3 & -1 \\ 4 & 3 & 2 & 6 \end{bmatrix}$$

**Solution** In a similar way, draw the graph. Here three lines pass at the highest point of the lower envelope. Thus, accordingly we get three square matrices of order 2, and there are three optimal solutions. But actually we shall have to select a pair of lines which have the slope opposite in sign. Thus, we get two reduced games with a payoff matrix of order 2 × 2 (see Fig. 13.2)

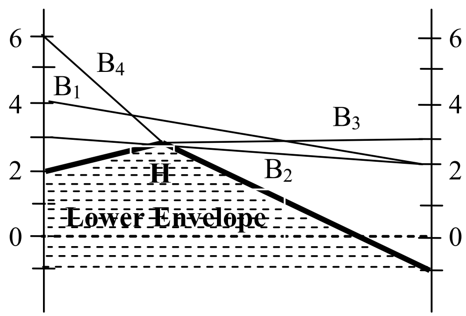



**Fig. 13.2** Graphical representation of 2 × 4 game

$$\begin{array}{cc} & \text{Player } B \\ & \begin{array}{cc} B_2 & B_3 \end{array} \\ \text{Player } A & \begin{array}{c} A_1 \\ A_2 \end{array} \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} \end{array} \quad \text{and} \quad \begin{array}{cc} & \text{Player } B \\ & \begin{array}{cc} B_3 & B_4 \end{array} \\ \text{Player } A & \begin{array}{c} A_1 \\ A_2 \end{array} \begin{pmatrix} 3 & -1 \\ 2 & 6 \end{pmatrix} \end{array}$$

Solving these, we get the value of the game as $\frac{5}{2}$ and the optimal strategies as $\xi^* = \left(\frac{1}{2}, \frac{1}{2}\right)$, $\eta^* = \left(0, \frac{1}{2}, \frac{1}{2}, 0\right)$ for the first case and $\xi^* = \left(\frac{1}{2}, \frac{1}{2}\right)$, $\eta^* = \left(0, 0, \frac{7}{8}, \frac{1}{8}\right)$ for the second case.

Any $m \times 2$ game problem can be solved by using B's mixed strategy $\eta = (y_1, y_2)$, where $y_1 + y_2 = 1$ and both lie in the open interval (0,1), and $A$'s particular critical moves can be determined graphically. In this case, the problem is to determine the minimum of the maximum expected loss of B. Thus, we shall have to select the lowest point (minimax) of the upper envelope, and $A$'s critical moves are those whose corresponding line segments pass through the minimax point. Now selecting a $2 \times 2$ payoff matrix, the value of the game can be determined easily.

*Example 7* Solve the game whose payoff matrix is as follows:

$$\begin{array}{ccc} & & \text{Player } B \\ & & \begin{array}{cc} B_1 & B_2 \end{array} \\ \text{Player } A & \begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} & \begin{bmatrix} 2 & -3 \\ -2 & 5 \\ 0 & -1 \end{bmatrix} \end{array}$$

Here $H$ is the lowest point of the upper envelope. The point $H$ is the point of intersection of the two line segments $A_1$, $A_2$ (see Fig. 13.3). Hence, the payoff matrix of the reduced game will be

$$\begin{array}{cc} & \begin{array}{cc} B_1 & B_2 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \end{array} & \begin{bmatrix} 2 & -3 \\ -2 & 5 \end{bmatrix} \end{array}$$

Now solving the reduced game by the algebraic method, we shall easily find optimal strategies and the value of the game. For this problem, the optimal strategies will be

$$\xi^* = \left(\frac{7}{12}, \frac{5}{12}, 0\right), \eta^* = \left(\frac{2}{3}, \frac{1}{3}\right)$$

and the value of the game is $\frac{1}{3}$.

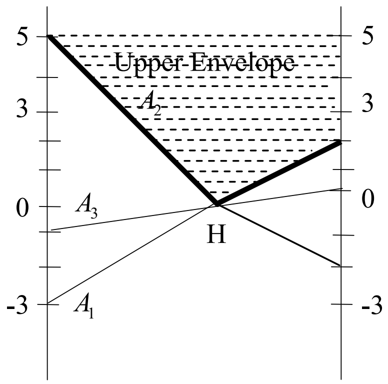



**Fig. 13.3** Graphical representation of 3 × 2 game

***Symmetric game***

A game is said to be symmetric if the payoff matrix remains unchanged on interchanging the positions of two players. Hence, the payoff matrix of any player of a symmetric game is skew-symmetric; i.e. the payoff matrix of any player is a square matrix in which $a_{ij} = -a_{ji}$ and $a_{ii} = 0$, since the gain of one player is equal to the loss of the other.

**Theorem 13.20** *In a symmetric game, the sets of all optimal strategies of players 1 and 2 are the same, and the value of the game is zero.*

*Proof* For a matrix game, the expected payoff for player 1 is given by

$$E(\xi, \eta) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_i y_j,$$

where $\xi = (x_1, x_2, \ldots, x_m)$ and $\eta = (y_1, y_2, \ldots, y_n)$ are the mixed strategies of $A$ and $B$ respectively and $\left[a_{ij}\right]_{m \times n}$ is the payoff matrix.

In a symmetric game, the payoff matrix is skew-symmetric, and $E(\xi, \xi) = 0$ and $E(\eta, \eta) = 0$.

Let $\xi^*$ and $\eta^*$ be the optimal strategies for players 1 and 2 respectively.

Then the value of the game is given by

$$v = E(\xi^*, \eta^*).$$

Again, we know that

$$v = \underset{\xi}{\text{Max}}\ \underset{\eta}{\text{Min}}\ E(\xi, \eta) = \underset{\eta}{\text{Min}}(\xi^*, \eta)$$
$$\therefore E(\xi^*, \eta) \geq v.$$

Setting $\eta = \xi^*$ above, we have $E(\xi^*, \xi^*) \geq v$.

$$\text{But } E(\xi^*,\ \xi^*) = 0;\ \text{therefore } v \le 0. \tag{13.30}$$

$$\text{Again } v = \underset{\eta}{\text{Min}}\ \underset{\xi}{\text{Max}}\ E(\xi,\ \eta) = \underset{\xi}{\text{Max}}\, E(\xi,\ \eta^*)$$
$$\therefore E(\xi,\ \eta^*) \le v.$$

Setting $\xi = \eta^*$, we have $E(\eta^*,\ \eta^*) \le v$.

$$\text{But } E(\eta^*,\ \eta^*) = 0;\ \text{therefore } v \ge 0. \tag{13.31}$$

Hence, from (13.30) and (13.31), we have

$$v = 0$$

and therefore $\xi^* = \eta^*$.

This completes the proof.

## 13.8 Matrix Games and Linear Programming Problems (LPPs)

**Theorem 13.21** *A matrix game is equivalent to a certain LPP.*

*Proof* Let $A = \left[a_{ij}\right]_{m \times n}$ be the payoff matrix and $v$ be the value of the game. It is assumed that $v > 0$. If not, then a suitable constant may be added to each element of $A$ so as to make $v$ positive. This change does not affect the optimal strategies for two players; only the value of the game is increased by this constant.

Let $\xi^*$, $\eta^*$ be the optimal strategies and $v$ (> 0) be the value of the game. Then

$$v \le E(\xi^*, \eta) \quad \forall \eta \in S_n,$$
$$\text{i.e. } v \le E(\xi^*,\ j) \quad \forall j = 1,\ 2,\ \ldots,\ n$$

We can say that $v$ is the largest number $u$ ($v \ \ge\ u$) such that for some $\xi \in S_m$

$$u \le E(\xi, j), \quad j = 1, 2, \ldots, n \tag{13.32}$$

$\xi$ is an optimal strategy for player 1 iff $\xi$ satisfies (13.32) with $u$ replaced by $v$. Now (13.32) can also be written as

$$\sum_{i=1}^{m} a_{ij}\, x_i \ge u \quad j = 1, 2, \ldots, n. \tag{13.33}$$

$$\text{Also, we have } \sum_{i=1}^{m} x_i = 1, \quad x_i \geq 0, \quad i = 1, 2, \ldots, m \tag{13.34}$$

Since $v$ is assumed to be positive, $u$ is also positive, so that from (13.33) and (13.34), we have

$$\sum_{i=1}^{m} a_{ij} \frac{x_i}{u} \geq 1, \quad j = 1,\, 2,\, \ldots,\, n$$
$$\text{and } \sum_{i=1}^{m} \frac{x_i}{u} = \frac{1}{u}, \quad \frac{x_i}{u} \geq 0, \quad i = 1,\, 2,\, \ldots,\, m.$$

Now substituting $\bar{x}_i$ for $\frac{x_i}{u}$, we get

$$\sum_{i=1}^{m} a_{ij}\, \bar{x}_i \geq 1, \quad j = 1,\, 2,\, \ldots,\, n$$
$$\text{and } \sum_{i=1}^{m} \bar{x}_i = \frac{1}{u}, \quad \bar{x}_i \geq 0, \quad i = 1,\, 2,\, \ldots,\, m.$$

For the optimal strategy, $v = \underset{\xi}{\text{Max}}\; u$ (we can regard $u$ as a function of $\xi$) or $\frac{1}{v} = \underset{\xi}{\text{Min}} \frac{1}{u} = \text{Min} \sum_{i=1}^{m} \bar{x}_i$.

Thus, finding $v$ and $\xi^*$ is equivalent to solving the following LPP:

$$\left.\begin{aligned} &\text{Minimize } z = \sum_{i=1}^{m} \bar{x}_i \\ &\text{subject to} \\ &\qquad \left.\begin{array}{ll} \sum_{i=1}^{m} a_{ij}\, \bar{x}_i \geq 1, & j = 1,\, 2,\, \ldots,\, n \\ \bar{x}_i \geq 0, & i = 1,\, 2,\, \ldots,\, m \end{array}\right\}. \end{aligned}\right. \tag{13.35}$$

If $z_{\min}$ is the optimal value and $\bar{\xi}^* = \left(x_1^*,\, \ldots,\, x_m^*\right)'$ is an optimal solution of (13.35), then the value of the game is $v = \frac{1}{z_{\text{Min}}}$, and an optimal strategy for player 1 is $\xi^* = v\bar{\xi}^* = \frac{\bar{\xi}^*}{z_{\text{Min}}}$.

Again, we have $v \geq E(\xi,\, \eta^*) \;\; \forall \xi \in S_m$.

Hence, $v \geq E(i,\, \eta^*), \quad i = 1,\, 2,\, \ldots,\, m.$

Thus there exists $w$ such that $(v \;\leq\; w)$ for $\eta \in S_n$,

$$w \geq E(i, \eta) \quad i = 1, \ldots, m. \tag{13.36}$$

$\eta$ is an optimal strategy for player 2 iff $\eta$ satisfies (13.36) with $w$ replaced by $v$. Now (13.36) can be written as

$$\sum_{j=1}^{n} a_{ij}\, y_j \leq w \quad i = 1, 2, \ldots, m. \tag{13.37}$$

$$\text{Also, we have } \sum_{j=1}^{n} y_j = 1, \quad y_j \geq 0, \quad j = 1, 2, \ldots, n. \tag{13.38}$$

Since $v$ is assumed to be positive, we can assume $w > 0$, so that from (13.37) and (13.38), we have

$$\begin{aligned}
&\sum_{j=1}^{n} a_{ij}\, \bar{y}_j \leq 1, \quad i = 1, 2, \ldots, m \\
&\sum_{j=1}^{n} \bar{y}_j = \frac{1}{w}, \quad \bar{y}_j \geq 0, \quad j = 1, 2, \ldots, n,
\end{aligned}$$

where $\bar{y}_j = y_j / w$.

For the optimal strategy,

$$\begin{aligned}
&v = \underset{\eta}{\text{Min}}\ w \ (w \text{ can be regarded as a function of } \eta) \\
\text{or } &\frac{1}{v} = \text{Min}\ \frac{1}{w} = \text{Min} \sum_{j=1}^{n} \bar{y}_j.
\end{aligned}$$

Thus, finding $v$ and $\eta^*$ is equivalent to solving the following LPP:

$$\left.\begin{aligned}
&\text{Maximize } z' = \sum_{j=1}^{n} \bar{y}_j \\
&\text{subject to} \\
&\left.\begin{aligned}
&\sum_{j=1}^{n} a_{ij}\, \bar{y}_j \leq 1, \quad && i = 1, 2, \ldots, m \\
&\bar{y}_j \geq 0, && j = 1, 2, \ldots, n
\end{aligned}\right\}.
\end{aligned}\right. \tag{13.39}$$

If $z'_{\text{Max}}$ is the optimal value and $\bar{\eta}^*$ is an optimal solution of (13.39), then $v = \frac{1}{z'_{\text{Max}}}$ and $\eta^* = v\,\bar{\eta}^* = \frac{\bar{\eta}^*}{z'_{\text{Max}}}$.

Note that (13.39) is the dual of (13.35) and vice versa.

*Example 8* Solve the following game by reducing it to an LPP:

$$\begin{array}{cc} & \begin{array}{ccc} B_1 & B_2 & B_3 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} & \begin{pmatrix} 1 & -1 & 3 \\ 3 & 5 & -3 \\ 6 & 2 & -2 \end{pmatrix} \end{array}$$

**Solution**

First we construct the table for row minimum and column maximum as follows:

$$\begin{array}{rcc} & & \text{Row minima} \\ & \begin{bmatrix} 1 & -1 & 3 \\ 3 & 5 & -3 \\ 6 & 2 & -2 \end{bmatrix} & \begin{array}{c} -1 \\ -3 \\ -2 \end{array} \\ \text{Column maxima} & \begin{array}{ccc} 6 & 5 & 3 \end{array} & \end{array}$$

Hence, maximin value = −1, minimax value = 3. Hence, the value of the game lies between −1 and 3, i.e. $-1 \le v \le 3$. This means that the value of the game may be negative or zero. Thus, we add a constant $k$ = 4 to all the elements of the matrix so that all these elements of the payoff matrix became positive, assuming that the value of the game represented by this new matrix is non-negative and non-zero. The transformed (reduced) matrix is as follows:

$$\begin{array}{cc} & B \\ & \begin{array}{ccc} B_1 & B_2 & B_3 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} & \begin{bmatrix} 5 & 3 & 7 \\ 7 & 9 & 1 \\ 10 & 6 & 2 \end{bmatrix} \end{array}$$

Let $p = (p_1, p_2, p_3)$ and $q = (q_1, q_2, q_3)$ be the mixed strategies of the two players respectively. Hence, the problem of player $B$ is as follows:

$$\text{Maximize } z' = x_1 + x_2 + x_3$$

subject to

$$5x_1 + 3x_2 + 7x_3 \le 1$$
$$7x_1 + 9x_2 + x_3 \le 1$$
$$10x_1 + 6x_2 + 2x_3 \le 1 \text{ and } x_1, x_2, x_3 \ge 0.$$

Now, if max $z' = \frac{1}{v^*}$, the value of the original game is $v = v^* - 4$, and the optimal solution is $q_j = x_j \times v^*$, $j$ = 1, 2, 3.

Introducing the slack variables $x_4$, $x_5$, $x_6$, we get

$$5x_1 + 3x_2 + 7x_3 + x_4 = 1$$

$$7x_1 + 9x_2 + x_3 + x_5 = 1$$

$$10x_1 + 6x_2 + 2x_3 + x_6 = 1.$$

| | | $c_j$ | 1 | 1 | 1 | 0 | 0 | 0 | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 0 | $y_4$ | $x_4 = 1$ | 5 | 3 | $7^*$ | 1 | 0 | 0 | 1/7 |
| 0 | $y_5$ | $x_5 = 1$ | 7 | 9 | 1 | 0 | 1 | 0 | 1/1 |
| 0 | $y_6$ | $x_6 = 1$ | 10 | 6 | 2 | 0 | 0 | 1 | 1/2 |
| $z_B = 0$ | | | $-1$ | $-1$ | $-1$ | 0 | 0 | 0 | $\leftarrow \Delta_j$ |
| | | | | | $\uparrow$ | $\downarrow$ | | | |

| | | $c_j$ | 1 | 1 | 1 | 0 | 0 | 0 | Mini ratio |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 1 | $y_3$ | $x_3 = 1/7$ | 5/7 | 3/7 | 1 | 1/7 | 0 | 0 | 1/3 |
| 0 | $y_5$ | $x_5 = 6/7$ | 44/7 | $60/7^*$ | 0 | $-1/7$ | 1 | 0 | 1/10 |
| 0 | $y_6$ | $x_6 = 5/7$ | 60/7 | 36/7 | 0 | $-2/7$ | 0 | 1 | 5/36 |
| $z_B = 1/7$ | | | $-2/7$ | $-4/7$ | 0 | 1/7 | 0 | 0 | $\leftarrow \Delta_j$ |
| | | | | $\uparrow$ | | | $\downarrow$ | | |

| | | $c_j$ | 1 | 1 | 1 | 0 | 0 | 0 | |
|---|---|---|---|---|---|---|---|---|---|
| $c_B$ | $y_B$ | $x_B$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | |
| 1 | $y_3$ | $x_3 = 1/10$ | 2/5 | 0 | 1 | 3/20 | $-1/20$ | 0 | |
| 1 | $y_2$ | $x_2 = 1/10$ | 11/15 | 1 | 0 | $-1/60$ | 7/60 | 0 | |
| 0 | $y_6$ | $x_6 = 1/5$ | 24/5 | 0 | 0 | $-1/5$ | $-3/5$ | 1 | |
| $z_B = 1/5$ | | | 2/5 | 0 | 0 | 2/15 | 1/5 | 0 | $\leftarrow \Delta_j$ |

Since all $z_j - c_j \geq 0$, the optimal solution is attained, and the optimal solution is $x_1 = 0,\quad x_2 = \frac{1}{10},\quad x_3 = \frac{1}{10}$ and $\text{Max}\, z' = \frac{1}{5}$.

From $z' = \frac{1}{v^*}$, we have $v^* = 5$

$$\therefore q_1 = x_1 v^* = 0, \quad q_2 = x_2 v^* = \frac{1}{10} \times 5 = \frac{1}{2}, \quad q_3 = x_3 v^* = \frac{1}{10} \times 5 = \frac{1}{2}$$

and the value of the game $v = v^* - 4 = 5 - 4 = 1$.

Since $A$'s problem is the dual of $B$'s problem, then using duality theory, we have

$$w_1 = \frac{2}{15}, \quad w_2 = \frac{1}{15}, \quad w_3 = 0 \quad \left[\text{values of } \Delta_4,\ \Delta_5,\ \Delta_6 \text{ if } \Delta_j = z_j - c_j\right]$$

$$\therefore p_1 = x_1 \cdot v^* = \frac{2}{15} \times 5 = \frac{2}{3}, \quad p_2 = \frac{1}{3}, \quad p_3 = 0.$$

Hence, the optimal solution is the following:

(i) Best strategy for $A$ is $p^* = \left(\frac{2}{3}, \frac{1}{3}, 0\right)$.
(ii) Best strategy for $B$ is $q^* = \left(0, \frac{1}{2}, \frac{1}{2}\right)$.
(iii) Value of the game = 1.

## 13.9 Infinite Antagonistic Game

**Definition** An infinite antagonistic game is a game in which at least one of the players possesses an infinite number of strategies.

Let $X$ and $Y$ be the set of strategies for players 1 and 2 respectively and $M(x, y)$ $(x \in X,\ y \in Y)$ be the payoff function of the antagonistic game $\langle X,\ Y,\ M\rangle$. In the case of the matrix game, recall that $X$, $Y$ are discrete sets of strategies (pure) for players 1 and 2 respectively. In that case, we have defined a mixed strategy for player 1 as a vector $\xi$ having $m$ (the number of pure strategies) components, each component diminishing the probability of choosing a pure strategy. A mixed strategy $\eta$ for player 2 was similarly defined. We can extend this idea of mixed strategy when $X$ and $Y$ are infinite.

In this case, a mixed strategy for player 1 will be denoted by a probability distribution function $F(x)$, where $F\left(x_i'\right) - F\left(x_i''\right)$ denotes the probability of choosing $x$ when $x_i'' < x < x_i'$. Similarly, the probability function $G(y)$ will denote a mixed strategy for player 2. Note that $F(x)$ is a non-negative, non-decreasing function and $\int_X \mathrm{d}F(x) = 1$. $G(y)$ is also non-negative and non-decreasing, and $\int_Y \mathrm{d}G(y) = 1$. Here the integration is in the Stieltjes sense.

$$\text{Let } \Delta_X = \left\{ F(x) : 0 \le F(x) \le 1, \quad \int_X \mathrm{d}F(x) = 1 \right\}$$

$$\Delta_Y = \left\{ G(y) : 0 \le G(y) \le 1, \quad \int_Y \mathrm{d}G(y) = 1 \right\}.$$

For a fixed $y$, let player 1 choose $x \in X$ with probability distribution $F(x)$; i.e. the probability of choosing $x$, where $x_i'' < x < x_i'$, is $F\left(x_i'\right) - F\left(x_i''\right)$. Then the expected payoff to player 1 is $M(x,\, y)\{F(x_i') - F(x_i'')\}$. Hence, when $x \in X$, the expected payoff to the player is (for a fixed $y$)

$$\sum_{i=1}^{n} M(x,\, y)\{F(x_i) - F(x_{i-1})\},$$

where $X$ is divided into $n$ intervals with the help of the points $x_i,\ i = 0,\, 1,\, \ldots,\, n$. If we now make $n \to \infty$ in such a way that $\operatorname*{Max}_i |x_i - x_{i-1}| \to 0$, then this becomes the Stieltjes integral with respect to $F(x)$ over the interval $X$;

$$\text{i.e. } E(F,\, y) = J(y) = \int_X M(x,\, y)\,\mathrm{d}F(x).$$

In a similar way, when player 2 chooses $y$ with probability distribution $G(y)$, then the expected payoff to player $j$ is $\sum_{j=1}^{n'} J(y)\{G(y_j) - G(y_{j-1})\}$, where $Y$ is divided into $n'$ intervals with the help of the points $y_0,\, \ldots,\, y_{n'}$. Now making $n' \to \infty$ in such a way that $\operatorname*{Max}_j |y_j - y_{j-1}| \to 0$, the expected payoff to player 1, when his strategy is $F(x)$ and player 2's strategy is $G(y)$, is

$$E(F,\, G) = \int_Y \left( \int_X M(x,\, y)\,\mathrm{d}F(x) \right) \mathrm{d}G(y).$$

*Note*

(i) If $M(x,\, y)$ is a continuous function of $x$, $y$ for $x \in X,\ y \in Y$, then

$$E(F,\, G) = \int_Y \left( \int_X M(x,\, y)\,\mathrm{d}x \right) \mathrm{d}y = \int_X \left( \int_Y M(x,\, y)\,\mathrm{d}y \right) \mathrm{d}x = \int_Y \int_X M(x,\, y)\,\mathrm{d}x\,\mathrm{d}y.$$

(ii) When $F(x)$ is the Heaviside function $H(x - x_0) = \begin{cases} 0 \ \text{ for } \ x < x_0 \\ 1 \ \text{ for } \ x > x_0 \end{cases}$, then $E(F,\, G)$ becomes $\int_Y M(x_0,\, y)\,\mathrm{d}G(y)$ and we denote it by $E(x_0,\, G)$. Thus, $E(x_0,\, G)$ is the expected payoff to player 1 when he chooses a simple strategy $H(x - x_0)$ and player 2 chooses the strategy $G$. Similarly, $E(F,\, y_0) = \int_X M(x,\, y_0)\,\mathrm{d}G(x)$ is the expected payoff to player 1 when he chooses the strategy $F(x)$ and player 2 chooses the simple strategy $H(y - y_0)$.

(iii) Let $\langle X,\, Y,\, M\rangle$ be an infinite antagonistic game. Then this is strategically equivalent to the game $\langle D_X,\, D_Y,\, E\rangle$, where $E = E(F,\, G),\ F \in D_X,\ G \in D_Y$.

***Equilibrium situation***

**Definition** Let $(F^*, G^*)$ be a situation in mixed strategies for the infinite antagonistic game $\langle X, Y, M\rangle$ or $\langle D_X, D_Y, E\rangle$. This situation is called an equilibrium situation in mixed strategies or a saddle point in mixed strategies if for any fixed strategies $F$, $G$ for player 1 and player 2 respectively, the following inequality is satisfied:

$$E(F, G^*) \le E(F^*, G^*) \le E(F^*, G)\, \forall F \in D_X, G \in D_Y.$$

For mixed strategies in infinite antagonistic games, we shall prove the following theorem.

**Theorem 13.22** If $E(x, G) \le u\ \forall x \in X$ and a fixed $G$, then $E(F, G) \le u\ \ \forall F \in D_X$ [for a matrix game: $F(i, \eta) \le u$, $i = 1, 2, \ldots, m$ for a fixed $\eta \in S_n$ implies that $E(\xi, \eta) \le u\ \ \forall \xi \in S_m$].

*Proof* We integrate both sides of $E(x, G) \le u$ with respect to $F(x)$ over $X$ so as to obtain $\int_X E(x, G)\,\mathrm{d}F(x) \le u \int_X \mathrm{d}F(x)$,

$$\text{i.e. } E(F, G) \le u \quad \left[\because \int_X \mathrm{d}F(x) = 1\right].$$

$$\text{Similarly, } E(x, G) \ge u\ \forall x \in X \text{ implies } E(F, G) \ge u\ \forall F \in D_X,$$
$$E(F, y) > u,\ \forall y \in Y \text{ implies } E(F, G) > u,\ \forall G \in D_Y$$
$$\text{and } E(F, y) < u,\ \forall y \in Y \text{ implies } E(F, G) < u,\ \forall G \in D_Y$$

**Theorem 13.23** *In order for the situation $(F^*, G^*)$ to be an equilibrium situation in an infinite antagonistic game $\Gamma = \langle X, Y, M\rangle$, it is necessary and sufficient that for all $x \in X$ and $y \in Y$, the following inequalities are satisfied:*

$$E(x, G^*) \le E(F^*, G^*) \le E(F^*, y).$$

[For a matrix game: $E(i, \eta^*) \le E(\xi^*, \eta^*) \le E(\xi^*, j) \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n$]

This proof is left to the reader.

**Theorem 13.24** *For any mixed strategy $G$ for player 2,*

$$\sup_{x \in X} E(x, G) = \sup_{F \in D_X} E(F, G),$$

*where the supremum on the left side is taken over all the pure strategies for player 1, and on the right side it is taken over all the mixed strategies.*

Similarly, for a fixed strategy $F$ for player 1,

$$\inf_{y \in Y} E(F, y) = \inf_{G \in D_Y} E(F, G).$$

$$\left[\begin{array}{r}\text{For a matrix game: } \operatorname*{Max}_i E(i, \eta) = \operatorname*{Max}_\xi (\xi, \eta) \text{ for some fixed } \eta \\ \text{and } \operatorname*{Min}_j E(\xi, j) = \operatorname*{Min}_\eta E(\xi, \eta) \text{ for some fixed } \xi\end{array}\right]$$

*Proof* Since the set of mixed strategies contains all the pure strategies

$$\left(\text{e.g.}, F(x) = H(x - x_0) = \begin{cases} 1 & x > x_0 \\ 1 & x < x_0 \end{cases}\right),$$

we have $\sup_{x \in X} E(x, G) \leq \sup_{F \in D_X} E(F, G)$.

Now, let if possible, $\sup_{x \in X} E(x, G) < \sup_{F \in D_X} E(F, G)$.

This implies that there exists $F_1 \in D_X$ such that for some $\varepsilon > 0$

$$\sup_{x \in X} E(x, G) < E(F_1, G) - \varepsilon,$$
$$\text{i.e. } E(x, G) < E(F_1, G) - \varepsilon \quad \forall x \in X.$$

Now integrating both sides with respect to $F_1(x)$, we have

$$\int_X E(x, G)\, \mathrm{d}F_1(x) < \{E(F_1, G) - \varepsilon\} \int_X \mathrm{d}F_1(x),$$
$$\text{i.e. } E(F_1, G) < E(F_1, G) - \varepsilon \quad \left[\because \int_X \mathrm{d}F_1(x) = 1\right],$$

which is a contradiction

Hence, $\sup_{x \in X} E(x, G) = \sup_{F \in D_X} E(F, G)$.

The proof of the second result is similar.

***Value of an infinite antagonistic game***

**Definition** If $\sup_F \inf_G E(F, G) = \inf_G \sup_F E(F, G)$, then the common value of these mixed extrema is called the value of the game $\langle X, Y, M \rangle$.

***Some properties of the value of the game and optimal strategies***

Let $v$ be the value of the game $\langle X, Y, M\rangle$.

**Theorem 13.25**

$$\sup_x \inf_y M(x, y) \le v \le \inf_y \sup_x M(x, y)$$

$$\left[\text{For a matrix game: } \underset{i}{\text{Max}}\ \underset{j}{\text{Min}}\ a_{ij} \le u \le \underset{j}{\text{Min}}\ \underset{i}{\text{Max}}\ a_{ij}\right].$$

*Proof* We know that $\inf_{y\in Y} E(F, y) = \inf_{G\in D_Y} E(F, G)$ for any arbitrary mixed strategy $F$ for player 1.

Let us choose a simple strategy $F = H(x - x') = \begin{cases} 0 & \text{for } x < x' \\ 1 & \text{for } x > x' \end{cases}$.

Then

$$\inf_{y\in Y} M(x', y) = E(x', G) \quad \forall x' \in X,$$

i.e. changing $x'$ to $x$, $\inf_{y\in Y} M(x, y) = \inf_{G\in D_Y} E(x, G)$.

$$\text{But } \sup_x \left\{\inf_y M(x, y)\right\} \le \sup_F \left\{\inf_y E(F, y)\right\}$$

[as the set of mixed strategies $D_X$ contains all pure strategies]

$$= \sup_F \left\{\inf_G E(F, G)\right\} = v,$$

$$\text{i.e. } \sup_x \left\{\inf_y M(x, y)\right\} \le v. \tag{13.40}$$

Similarly, we can prove that

$$v \le \inf_y \sup_x M(x, y). \tag{13.41}$$

Hence, from (13.40) and (13.41), we have

$$\sup_x \left\{\inf_y M(x, y)\right\} \le v \le \inf_y \sup_x M(x, y).$$

**Theorem 13.26** *If player 1 possesses a pure optimal strategy $x_0$ and player 2 possesses an arbitrary strategy (in general mixed), then*

$$v = \underset{x}{\text{Max}} \left\{ \underset{y}{\text{Min}}\ M(x, y) \right\} = \underset{y}{\text{Min}}\ M(x_0, y).$$

Similarly, if player 2 possesses a pure strategy $y_0$ and player 1 possesses an arbitrary strategy, then

$$v = \underset{y}{\text{Min}} \left\{ \underset{x}{\text{Max}}\ M(x, y) \right\} = \underset{x}{\text{Max}}\ M(x, y_0).$$

The reader may easily prove this theorem.

**Theorem 13.27** *Let u be a real number.*

(i) *If $F_0$ is a mixed strategy for player 1, then*

$$u \le E(F_0, y) \quad \forall y \in Y \Rightarrow u \le v.$$

(ii) *If $G_0$ is a mixed strategy for player 2, then*

$$u \le E(x, G_0) \quad \forall x \in X \Rightarrow u \le v.$$

*Proof* We have $u \le E(F_0, y)\ \ \forall y \in Y$.

Now integrating both sides with respect to $G(y)$ over $Y$, we have

$$u \int_Y \mathrm{d}G(y) \le \int_Y E(F_0, y)\, \mathrm{d}G(y)$$

$$\text{or } u \le E(F_0, G) \quad \forall G \in D_Y$$

$$\text{Hence } u \le \inf_G E(F_0, G) \le \sup_F \left\{ \inf_G E(F, G) \right\} = v$$

$$\therefore u \le v.$$

The proof of (ii) is similar.

**Theorem 13.28**

(i) *In order for the strategy $F_0$ for player 1 to be optimal, it is necessary and sufficient that*

$$E(F_0, y) \ge v\ \forall y \in Y.$$

(ii) *In order for the strategy $G_0$ for player 2 be optimal, it is necessary and sufficient that*

$$E(x,\, G_0) \le v\,\forall x \in X.$$

*Proof of (i)* Let $F_0$ be an optimal strategy. This implies that

$$\inf_G E(F_0,\, G) = \sup_F \inf_G E(F,\, G) = v,$$

so that $E(F_0,\, G) \ge v$ for all $G \in D_Y$.

Let us choose $G$ to be a pure strategy $H(y - y_0)$. Then we obtain

$$E(F_0,\, y_0) \ge v \;\text{ for all } y_0 \in Y,$$

i.e. $E(F_0,\, y) \ge v$ for all $y \in Y$.

$$\textbf{\textit{Sufficient}}\text{: Let } E(F_0,\, y) \ge v \;\text{ for all } y \in Y. \tag{13.42}$$

Now, we have to show that $F_0$ is an optimal strategy for player 1.

Integrating both sides of (13.42) with respect to $G(y)$ over $Y$, we have

$$\int_Y E(F_0,\, y)\, dG(y) \ge v \int_Y dG(y)$$

or $E(F_0, G) \ge v \quad \forall G \in D_Y$.

$$\text{Hence, } \inf_G E(F_0,\, G) = v = \sup_F \left\{ \inf_G E(F,\, G) \right\},$$

i.e. the supremum of the function $\inf\limits_G E(F_0,\, G)$ is attained at $F_0$. This implies that $F_0$ is optimal.

## 13.10 Continuous Game

**Definition** An antagonistic game $\Gamma = \langle X,\, Y,\, M \rangle$ is called a game on the unit square if the sets of pure strategies for each of the players are the unit segments, i.e. $X = Y = [0,\, 1]$.

Hence, the mixed strategies for the players in a game on the unit square are probability distributions on the segment $[0,\, 1]$ such that $\int_0^1 \mathrm{d}F(x) = 1$ and $\int_0^1 \mathrm{d}G(y) = 1$.

$$\text{Here } D_X = D_Y = D.$$

**Definition** A game on the unit square is called continuous if the payoff function is continuous.

***Fundamental theorem for continuous game***

**Theorem 13.29** *For any continuous game on the unit square with payoff function M(x, y), where both players possess optimal strategies,*

$$\underset{F}{\text{Max}}\left\{\underset{G}{\text{Min}}\ E(F,\, G)\right\} \text{ and } \underset{G}{\text{Min}}\left\{\underset{F}{\text{Max}}\ E(F,\, G)\right\}$$

*exist and are equal.*

*The proof of the theorem is left to the reader.*

## 13.11 Separable Game

**Definition** A function $M(x, y)$ of two variables $x$, $y$ is called separable if there exist $m$ continuous functions $r_i\,(x)$, $i = 1, 2, \ldots, m$, $n$ continuous functions $s_j\,(y)$, $j = 1, 2, \ldots, n$ and $mn$ constants $a_{ij}$ $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$ such that

$$M(x,\, y) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\, r_i(x)\; s_j(y).$$

*Example 9* $M(x,\, y) = x\sin y + x\cos y + 2x^2$ is separable.

We can take

$$\begin{array}{lll} r_1(x) = x & s_1(y) = \sin y & a_{11} = 1,\ a_{12} = 1,\ a_{13} = 0 \\ r_2(x) = x^2 & s_2(y) = \cos y & a_{21} = 0,\ a_{22} = 0,\ a_{23} = 2 \\ & s_3(y) = 1 & \end{array}$$

We can also take

$$\begin{array}{lll} r_1(x) = x & s_1(y) = \sin y + \cos y & a_{11} = 1,\ a_{12} = 0 \\ r_2(x) = 2x^2 & s_2(y) = 1 & a_{21} = 0,\ a_{22} = 1 \end{array}$$

*Note*

(i) The representation $M(x,\ y) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\, r_i(x)\ s_j(y)$ of a separable function is not unique.

(ii) If $M(x, y)$ is separable, we may also write

$$M(x,\ y) = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ij}\, r_i(x)\right) s_j(y) = \sum_{j=1}^{n} t_j(x)\, s_j(y),$$

where $t_i(x)$ and $s_j(y)$ are continuous functions of $x$ and $y$ respectively.

Thus, a continuous game on the unit square whose payoff function is separable is called a separable game.

***Expectation function in a separable game***

A separable game is $\langle X,\ Y,\ M\rangle$, where $X = Y = [0, 1]$, and $M(x, y)$ is separable. Hence, a mixed strategy for player 1 is a probability distribution function $F(x)$, and a mixed strategy for player 2 is a probability distribution function $G(y)$, where $F \in D,\ \ G \in D$ and

$$D = \left\{F : 0 \le F \le 1,\ \int_0^1 \mathrm{d}F(x) = 1\right\}$$
$$\text{or } D = \left\{G : 0 \le G \le 1,\ \int_0^1 \mathrm{d}G(y) = 1\right\}.$$

If player 1 chooses a mixed strategy $F(x)$ and player 2, a mixed strategy $G(y)$ $(F,\ G \in D)$, then the expectation of player 1 is

$$\begin{aligned} E(F,\ G) &= \int_0^1\int_0^1 M(x,\ y)\,\mathrm{d}F(x)\,\mathrm{d}G(y) \\ &= \int_0^1\int_0^1 \left\{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\, r_i(x)\, s_j(y)\right\} \mathrm{d}F(x)\,\mathrm{d}G(y) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\left\{\int_0^1 r_i(x)\,\mathrm{d}F(x)\right\}\left\{\int_0^1 s_j(y)\,\mathrm{d}G(y)\right\} \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\, u_i\, v_j, \end{aligned}$$

where $u_i = \int_0^1 r_i(x)\,\mathrm{d}F(x), \quad i = 1, 2, \ldots, m$

and $v_j = \int_0^1 s_j(y)\,\mathrm{d}G(y), \quad j = 1, 2, \ldots, n.$

***Note* 1**: There may exist two different strategies $F_1(x)$, $F_2(x)$ such that

$$\int_0^1 r_i(x)\,\mathrm{d}F_1(x) = \int_0^1 r_i(x)\,\mathrm{d}F_2(x) \quad i = 1, 2, \ldots, m.$$

But

$$\begin{aligned} E(F_1, G) &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\,u_i v_j = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\left(\int_0^1 r_i(x)\,\mathrm{d}F_2(x)\right) v_j \\ &= E(F_2, G) \quad \forall\, G \in D. \end{aligned}$$

Thus, it is immaterial whether player 1 uses strategy $F_1(x)$ or $F_2(x)$. In this case we shall say that the two strategies $F_1$ and $F_2$ are equivalent (with respect to the game).

***P-space and Q-space***

Let $\alpha = (u_1, \ldots, u_m)'$, $\beta = (v_1, \ldots, v_n)'$; then $\alpha \in E^m$, $\beta \in E^n$.

Let $P = \left\{\alpha = (u_1, u_2, \ldots, u_m)';\ u_i = \int_0^1 r_i(x)\,\mathrm{d}F(x),\ i = 1, 2, \ldots, m;\ F \in D\right\}$.

Then obviously $P \subset E^m$. We also say that $\alpha$ and $F$ correspond. Note that a given point $\alpha$ of the $P$-space, in general, corresponds to many different distribution functions.

Similarly, let $Q = \left\{\beta = (v_1, v_2, \ldots, v_n)';\ v_j = \int_0^1 s_j(y)\,dG(y),\ j = 1, 2, \ldots, n;\ G \in D\right\}$.

Then obviously $Q \subset E^n$. We also say that $\beta$ and $G$ correspond. Note that a given point $\beta$ in the $Q$-space, in general, corresponds to many different distribution functions.

*Note* If $u$ corresponds to both $F(x)$ and $F_1(x)$, and $v$ corresponds to both $G(y)$ and $G_1(y)$, then

$$\begin{aligned} E(F, G) &= E(F_1, G) = E(F, G_1) = E(F_1, G_1) \\ &= \sum_{i,j} a_{ij}\,u_i v_j, \text{ and let us denote it by } E(\alpha, \beta) \end{aligned}$$

***Geometrical properties of P- and Q-spaces***

**Theorem 13.30** *Let $F_k(x)$ $(k = 1, \ldots, p)$ be a distribution function (i.e. strategies for player 1), and let $F(x) = \sum_{k=1}^{p} c_k F_k(x)$, where $c_k \geq 0$, $k = 1, 2, \ldots, p$ and*

$\sum_{k=1}^{p} c_k = 1$. *Then* $F \in D$. *Let* $\alpha_k$ *and* $F_k$ *correspond* $(k = 1, \ldots, p)$. *Then* $\alpha = \sum_{k=1}^{p} c_k \alpha_k$ *corresponds to* $F$ *so that* $\alpha \in P$.

A similar result holds for the $Q$-space.

*Proof* We have $M(x, y) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} r_i(x) s_j(y)$.

Let $\alpha^k = \left(u_1^{(k)}, \ldots, u_m^{(k)}\right)' \quad u = 1, \ldots, p$.

Since $\alpha_k$ corresponds to $F_k(x)$, $k = 1, \ldots, p$, we have

$$u_i = \int_0^1 r_i(x)\, \mathrm{d}F_k(x),\ i = 1, \ldots, m.$$

Let $\alpha = (u_1, \ldots, u_m)'$.

$$\begin{aligned}
\text{Since } \alpha = \sum_{k=1}^{p} c_k \alpha_k \text{, then } u_i &= \sum_{k=1}^{p} c_k u_i^{(k)}, \quad i = 1, 2, \ldots, m \\
&= \sum_{k=1}^{p} c_k \int_0^1 r_i(x)\, \mathrm{d}F_k(x) \\
&= \int_0^1 r_i(x)\, d\left(\sum_{k=1}^{p} c_k F_k(x)\right) = \int_0^1 r_i(x)\, \mathrm{d}F(x)
\end{aligned}$$

Since $F \in D$, $\alpha$ corresponds to $F$, and hence $\alpha \in P$.

The proof for the $Q$-space is similar.

***P′*- and *Q′*-spaces**

Let us define $P^* = \left\{\rho = (r_1, \cdots, r_m)' : r_i(t),\ 0 \le t \le 1\right\}$

$$\begin{aligned}
&\therefore P^* \in E^m \\
&Q^* = \left\{\sigma = (s_1, \ldots, s_n)' : s_j(t),\ 0 \le t \le 1\right\} \\
&\therefore Q^* \subset E^n.
\end{aligned}$$

***Properties of P′- and Q′-spaces***

***Property-1***: $P^* \subset P$ and $Q^* \subset Q$.

***Property-2***: Since $r_i(x)$ and $s_j(y)$ are continuous functions defined on a closed set, $P^*$ and $Q^*$ are bounded, closed and connected sets.

***Property-3***: The $P$-space is the convex hull of $P^*$. Similarly, the $Q$-space is the convex hull of $Q^*$.

*Example 10* Let the payoff function be $M(x, y) = \cos(2\pi x)\cos(2\pi y) + x + y$, $0 \le x,\ y \le 1$.

Here we can write $M(x, y) = \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}\, r_i(x)\, s_j(y)$,

where

$$\begin{array}{lll} r_1(x) = x, & s_1(y) = y, & a_{11} = 0,\ a_{12} = 0,\ a_{13} = 1 \\ r_2(x) = \cos 2\pi x, & s_2(y) = \cos 2\pi y, & a_{21} = 0,\ a_{22} = 1,\ a_{23} = 0 \\ r_3(x) = 1, & s_3(y) = 1, & a_{31} = 1,\ a_{32} = 0,\ a_{33} = 0 \end{array}$$

$$\text{Thus, } P^* = \left\{\rho = (r_1, r_2, r_3)' \colon r_1(t) = t,\ r_2(t) = \cos 2\pi t,\ r_3(t) = 1,\ 0 \le t \le 1\right\}$$
$$Q^* = \left\{\sigma = (s_1, s_2)' \colon s_1(t) = t,\ s_2(t) = \cos 2\pi t,\ s_3(t) = 1,\ 0 \le t \le 1\right\}.$$

***Image of*** $\alpha \in P$

**Definition** The image of $\alpha \in P$ is defined as the set of points $\beta \in Q$ such that $E(\alpha, \beta) = \underset{\eta \in Q}{\text{Min}}\ E(\alpha, \eta)$, and we denote this image of $\alpha$ by $Q(\alpha)$ and $Q(\alpha) \subset Q$. Note that $Q(\alpha)$ is in general a set of points $\in Q$.

***Image of*** $\beta \in Q$

**Definition** Similarly, the image of $\beta \in Q$ is defined as the set of points $\alpha \in P$ such that $E(\alpha, \beta) = \underset{\xi \in P}{\text{Max}}\ E(\xi, \beta)$, and we denote this image of $\beta$ by $P(\beta)$ and $P(\beta) \subset P$. Note that $P(\beta)$ in general is a set of points $\in P$.

***Fixed point***

**Definition** If $\beta \in Q$ is an image point of $\alpha \in P$ and $\alpha \in P$ is an image point of $\beta \in Q$, then $\alpha$ is called a fixed point of $P$-space and $\beta$ is a fixed point of $Q$-space.

**Theorem 13.31** *If $F_1$ is any distribution function and $\alpha \in P$ corresponds to $F_1$, then $F_1$ is an optimal strategy for player 1 iff $\alpha$ is a fixed point P. Similarly, G is an optimal strategy for player 2 iff the corresponding point $\beta \in Q$ is a fixed point Q.*

*Proof* Let $F_1$ be a mixed strategy for player 1 and let it correspond to $\alpha$. Now, $\alpha$ is a fixed point of $P$ means that for some point $\beta \in Q$, we have $\beta \in Q(\alpha)$ and $\alpha \in P(\beta)$. That is, $\beta$ is an image point of $\alpha$ and $\alpha$ is an image point of $\beta$, i.e.

$$E(\alpha, \beta) = \underset{\eta \in Q}{\text{Min}}\ E(\alpha, \eta)$$
$$\text{and } E(\alpha, \beta) = \underset{\xi \in P}{\text{Max}}\ E(\xi, \beta).$$

Let $G_1$ be a strategy (mixed) for player 2, and let it correspond to $\beta \in Q$. Then

$$E(F_1, G_1) \equiv E(\alpha, \beta) = \underset{\eta \in Q}{\text{Min}}\ E(\alpha, \eta) = \underset{G \in D}{\text{Min}}\ E(F_1, G)$$
$$\therefore E(F_1, G_1) \leq E(F_1, G) \quad \forall G \in D.$$
$$\text{Again } E(F_1, G_1) \equiv E(\alpha, \beta) = \underset{\xi \in P}{\text{Max}}\ E(\xi, \beta) = \underset{F \in D}{\text{Max}}\ E(F, G_1) \geq E(F, G_1) \quad \forall F \in D,$$
$$\text{i.e. } E(F, G_1) \leq E(F_1, G_1) \quad \forall F \in D,$$

Hence, we obtain

$$E(F, G_1) \leq E(F_1, G_1) \leq E(F_1, G) \quad \forall F, G \in D.$$

Thus, $F_1$ and $G_1$ are optimal strategies.

**Theorem 13.32** *If $\alpha$ is any fixed point of $P$ and $\beta$ is any fixed point of $Q$, then $\alpha$ is an image point of $\beta$, and $\beta$ is an image point of $\alpha$.*

*Proof* Since $\alpha$ is a fixed point of $P$, then $\exists$ a point $\beta_1$ of $Q$ such that $\beta_1$ is an image point of $\alpha$ and $\alpha$ is an image point $\beta_1$, i.e. $\alpha \in P(\beta_1)$ and $\beta_1 \in Q(\alpha)$, i.e. such that

$$E(\alpha, \beta_1) = \underset{\xi \in P}{\text{Max}}\ E(\xi, \beta_1) \tag{13.43}$$

$$\text{and } E(\alpha, \beta_1) = \underset{\eta \in Q}{\text{Min}}\ E(\alpha, \eta). \tag{13.44}$$

Similarly, since $\beta$ is a fixed point of $Q$, then there exists a point $\alpha_1 \in P$ such that $\alpha_1$ is an image point of $\beta$ and $\beta$ is an image point of $\alpha_1$,

$$\left.\begin{array}{l} \text{i.e. } \alpha_1 \in P(\beta) \text{ and } \beta \in Q(\alpha_1) \\ \text{i.e. } E(\alpha_1, \beta) = \underset{\xi \in P}{\text{Max}}\ E(\xi, \beta) \end{array}\right. \tag{13.45}$$

$$\text{and } E(\alpha_1, \beta) = \underset{\eta \in Q}{\text{Min}}\ E(\alpha_1, \eta). \tag{13.46}$$

Thus, we have

$$\begin{aligned}
E(\alpha, \beta) &\le \operatorname*{Max}_{\xi \in P} E(\xi, \beta) \quad \text{[by the definition of maxima]}\\
&= E(\alpha_1, \beta) \quad \text{[by (13.45)]}\\
&= \operatorname*{Min}_{\eta \in Q} E(\alpha_1, \eta) \quad \text{[by (13.46)]}\\
&\le E(\alpha_1, \beta_1) \quad \text{[by the definition of minima]}\\
&\le \operatorname*{Max}_{\xi} E(\xi, \beta_1) \quad \text{[by the definition of maxima]}\\
&= E(\alpha, \beta_1) \quad \text{[by (13.43)]}\\
&= \operatorname*{Min}_{\eta \in Q} E(\alpha, \eta) \quad \text{[by (13.44)]}\\
&\le E(\alpha, \beta) \quad \text{[by the definition of minima].}
\end{aligned}$$

Since the first and last terms of these inequalities are equal, we therefore conclude that all the quantities involved are equal, and thus in particular,

$$E(\alpha, \beta) = \operatorname*{Max}_{\xi \in P} E(\xi, \beta) = \operatorname*{Min}_{\eta \in Q} E(\alpha, \eta),$$

which means that $\alpha \in P(\beta)$ and $\beta \in Q(\alpha)$, i.e. $\alpha$ is an image point of $\beta$ and $\beta$ is an image point of $\alpha$.

***General procedure to solve a separable game***

Step-1 Plot the curves $P^*$ and $Q^*$ and determine their convex hulls to give the $P$-space and $Q$-space.

Step-2 Find $Q(\alpha)$ for every $\alpha \in P$ and $P(\beta)$ for every $\beta \in Q$; i.e. find all the image points of the $P$space and $Q$-space.

Step-3 Using the results of Step 2, find the fixed points of $P$ and $Q$.

Step-4 Express the fixed points as convex linear combinations of points in $P^*$ and $Q^*$ respectively, then find the distribution function in which these fixed points corresponding to the distribution function will be the optimal strategies

*Example 11*

Given that $M(x, y) = 3\cos 4x \ \cos 5y + 5\cos 4x \ \sin 5y + \sin 4x \cos 5y + \sin 4x \sin 5y$. Find $P^*$and $Q^*$

*Solution* We can write $M(x, y) = \sum_{i=1}^{2} \sum_{j=1}^{2} a_{ij}\, r_i(x)\, s_j(y)$,

$$\begin{aligned}
\text{where } r_1(x) &= \cos 4x \quad s_1(y) = \cos 5y \quad 0 \le x \le 1\\
r_2(x) &= \sin 4x \quad s_2(y) = \sin 5y \quad 0 \le y \le 1
\end{aligned}$$

and $a_{11} = 3,\ a_{12} = 5,\ a_{21} = 1,\ a_{22} = 1.$

Thus, $P^* = \left\{\rho = (r_1,\, r_2)' \quad r_1 = \cos 4t, \quad r_2 = \sin 4t, \quad 0 \le t \le 1\right\}$

and $Q^* = \left\{\sigma = (s_1,\, s_2)', \quad s_1 = \cos 5t, \quad s_2 = \sin 5t, \quad 0 \le t \le 1\right\}$.

***Canonical representation of the payoff function in a separable game***

For a separable game, let the payoff function $M(x, y)$ be represented by

$$M(x,\, y) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, r_i(x)\, s_j(y) + \sum_{i=1}^{n} b_i r_i(x) + \sum_{j=1}^{n} c_j s_j(y) + d,$$

where

(i) $r_1(x), r_2(x), \ldots r_n(x)$ and $s_1(y), \ldots, s_n(y)$ are continuous functions over [0, 1].
(ii) det $\left[a_{ij}\right] \neq 0$.

Any representation of a separable function in this form is called a canonical representation of the function $M(x, y)$.

We note that every separable function has a canonical form. For if $M(x, y)$ is separable, then it can always be written in the form $\sum_{j=1}^{n} r_j(x) s_j(y)$ so that det $\left[a_{ij}\right] \neq 0$.

*Example 12*

Examine whether the representation $M(x,\, y)$
$= xy - xe^y + 2x\cos y + 2e^x y + 3e^x e^y + e^x \cos y + 5\cos x e^y$
$- 3\cos x \cos y$ is canonical or not.

Let $r_1(x) = x \quad s_1(y) = y,\ r_2(x) = e^x \quad s_2(y) = e^y,\ r_3(x) = \cos x \quad s_3(y) = \cos y.$

Then $a_{11} = 1, \quad a_{12} = -1, \quad a_{13} = 2, \quad b_1 = b_2 = b_3 = 0$
$a_{21} = 2, \quad a_{22} = 3, \quad a_{23} = 1, \quad c_1 = c_2 = c_3 = 0$
$a_{31} = 0, \quad a_{32} = 5, \quad a_{33} = -3, \quad d = 0.$

Now, $\det\left[a_{ij}\right] = \begin{vmatrix} 1 & -1 & 2 \\ 2 & 3 & 1 \\ 0 & 5 & -3 \end{vmatrix} = 0.$

Now from $M(x,\, y) = xy - xe^y + 2x\cos y + 2e^x y + 3e^x e^y + e^x \cos y + 5\cos x e^y - 3\cos x \cos y$

we have

$$M(x,\ y) = (x - 2\cos x)\left(y + \frac{7}{5}\cos y\right) - (x - 2\cos x)\left(e^y - \frac{3}{5}\cos y\right)$$
$$+ 2(e^x + \cos x)\left(y + \frac{3}{5}\cos y\right) + 3(e^x + \cos x)\left(e^y - \frac{3}{5}\cos y\right),$$
$$\text{i.e. } M(x,\ y) = r_1(x)s_1(y) - r_1(x)s_2(y) + 2r_2(x)s_1(y) + 3r_2(x)s_2(y)$$

$$\text{where } r_1(x) = x - 2\cos x \quad s_1(y) = y + \frac{7}{5}\cos y$$
$$r_2(x) = e^x + \cos x \quad s_2(y) = e^y - \frac{3}{5}\cos y.$$

Hence,

$$a_{11} = 1, \quad a_{22} = -1, \quad b_1 = b_2 = 0 \quad a_{21} = 2, \quad a_{22} = 3 \quad c_1 = c_2 = 0 \quad d = 0.$$

Therefore, $\det\left[a_{ij}\right] = \begin{vmatrix} 1 & -1 \\ 2 & 3 \end{vmatrix} = 5 \neq 0.$

Thus, this representation is canonical.

***First critical point and second critical point***

Let a separable function $M(x,\ y)$ have the following canonical representation:

$$M(x,\ y) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\, r_i(x)\, s_j(y) + \sum_{i=1}^{n} b_i\, r_i(x) + \sum_{j=1}^{n} c_j\, s_j(x) + d,$$

where $r_i(x),\ s_j(y) \quad (i, j = 1, 2, \ldots, n)$ are continuous functions on [0, 1] and $\det\left[a_{ij}\right] \neq 0$.

In this case, let $F$, a mixed strategy for player 1, correspond to $\alpha = (u_1, \ldots, u_n)' \in P$, and let $G$, a mixed strategy for player 2, correspond to $\beta = (v_1, \ldots, v_n)' \in Q$ (note that both $P$, $Q \subset E^n$).

$$\text{Then } u_i = \int_0^1 r_i(x)\, \mathrm{d}F(x), \quad i = 1, 2, \ldots, n$$
$$v_j = \int_0^1 s_j(y)\, \mathrm{d}G(y), \quad j = 1, 2, \ldots, n$$

$$\text{and } E(F,\, G) \equiv E(\alpha,\, \beta) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}u_i v_j + \sum_{i=1}^{n} b_i u_i + \sum_{j=1}^{n} c_j v_j + d$$
$$= \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij}u_i + c_j\right)v_j + \sum_{i=1}^{n} b_i u_i + d.$$

Alternatively,

$$E(F,\, G) = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}v_i + b_i\right)u_i + \sum_{j=1}^{n} c_j v_j + d.$$

Since $\det\left[a_{ij}\right] \neq 0$, then the system of equations

$$\sum_{i=1}^{n} a_{ij}u_i + c_j = 0, \quad j = 1,\, 2,\, \ldots,\, n \tag{13.47}$$

has a unique solution, as does the system of equations

$$\sum_{j=1}^{n} a_{ij}v_j + b_i = 0, \quad i = 1,\, 2,\, \ldots,\, n \tag{13.48}$$

The unique solution of (13.47), say $\prod = (p_1,\, p_2,\, \ldots,\, p_n)'$, is called the first critical point of the game (with respect to the given representation), and the unique solution of (13.48), say $\chi = (q_1,\, q_2,\, \ldots,\, q_n)'$, is called the second critical point of the game (with respect to the given representation).

*Note* The point $\prod$ may not belong to the $P$-space, and the point $\chi$ may not belong to the $Q$-space. However, if $\prod \in P$ and $\chi \in Q$, then the corresponding game problem can be solved easily by using the following two theorems.

**Theorem 13.33** *Let the payoff function of a separable game have a given canonical form*

$$M(x, y) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}r_i(x)\, s_j(y) + \sum_{i=1}^{n} b_i r_i(x) + \sum_{j=1}^{n} c_j s_j(y) + d,\ \det\left[a_{ij}\right] \neq 0$$

*and also let* $\prod = (p_1,\, p_2,\, \ldots,\, p_n)'$, $\chi = (q_1,\, q_2,\, \ldots,\, q_n)'$ *be respectively the first and second critical points and suppose* $\prod \in P, \chi \in Q$. *Then* $\prod$ *and* $\chi$ *are the fixed points, and the value of the game is given by* $E(\prod,\, \chi) = \sum_{i=1}^{n} b_i p_i + d = \sum_{j=1}^{n} c_j q_j + d$.

*Proof* By the definition of critical points,

$$\sum_{i=1}^{n} a_{ij}p_i + c_j = 0, \quad j = 1, 2, \ldots, n$$
$$\text{and} \quad \sum_{j=1}^{n} a_{ij}q_j + b_i = 0, \quad i = 1, 2, \ldots, n.$$

Now, for any point $\beta \in Q$ $\left[\beta = (v_1, v_2, \ldots, v_n)'\right]$,

$$\begin{aligned} E\left(\prod, \beta\right) &= \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij}p_i + c_j\right)v_j + \sum_{i=1}^{n} b_ip_i + d \\ &= 0 + \sum_{i=1}^{n} b_ip_i + d, \text{which is independent of } \beta. \end{aligned}$$

Hence, $Q(\prod) = Q$. Similarly, $P(\chi) = P$.

Hence, $\prod \in P(\chi)$ and $\chi \in Q(\prod)$, so that $\prod$ and $\chi$ are fixed points, and the value of the game is

$$E\left(\prod, \chi\right) = \sum_{i=1}^{n} b_ip_i + d = \sum_{j=1}^{n} c_jq_j + d.$$

**Theorem 13.34** *Let the interior of Q contain a fixed point; then the first critical point belongs to P and is the only fixed point. Similarly, if the interior of P contains a fixed point, then the second critical point belongs to Q and is the only fixed point of Q.*

*Proof* Let $\beta = (v_1, \ldots, v_n)'$ be a fixed point of $Q$ which lies in the interior of $Q$. Let $\alpha = (u_1, \ldots, u_n)'$ be a fixed point of $P$ and let $\prod = (p_1, \ldots, p_n)'$ be the first critical point. We wish to show that $\alpha = \prod$. Suppose if possible, $\alpha \neq \prod$.

Since $\prod$ is the unique solution of $\sum_{i=1}^{n} a_{ij}p_i + c_j = 0, \quad j = 1, 2, \ldots, n,$ we have for some $k \leq n$,

$\sum_{i=1}^{n} a_{ik}u_i + c_k \neq 0$ and is equal to $g$.

Let $h$ be a real number of opposite sign to $g$, and let it be small enough to ensure that the point

$$\bar{\beta} = (v_1, \ldots, v_{k-1}, v_k + h, v_{k+1}, \ldots, v_n) \in Q.$$

$$\text{Now } E(\alpha, \beta) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij}u_i + c_j\right)v_j + \sum_{i=1}^{n} b_iu_i + d$$

$$\begin{aligned} E\left(\alpha, \bar{\beta}\right) &= E(\alpha, \beta) + h\left(\sum_{i=1}^{n} a_{ik}v_k + c_k\right) \\ &= E(\alpha, \beta) + hg \\ &< E(\alpha, \beta) \text{ as } hg < 0. \end{aligned}$$

This means that $\alpha \notin P(\beta)$, which contradicts the assumption that $\alpha$ and $\beta$ are both fixed points of $P$ and $Q$ respectively. Hence, $\alpha$ must coincide with $\prod$.

**Corollary 1** *Let the payoff of a separable game be given in a canonical form and let the first critical point $\notin P$. Then every fixed point of $Q$ is in the boundary of $Q$. Similarly, if the second critical point $\notin Q$, then every fixed point of $P$ is in the boundary of $P$.*

**Corollary 2** *Let the payoff function of a separable game be given in a canonical form and $\prod$ and $\chi$ be the first and second critical points respectively. Let $\prod$ belong to the interior of $P$ and $\chi$ belong to the interior of $Q$; then $\prod$ is the only fixed point of $P$ and $\chi$ is the only fixed point of $Q$.*

*Example 13* Solve the separable game whose payoff function is

$$M(x, y) = 3\cos 7x \cos 8y + 5\cos 7x \sin 8y + 2\sin 7x \cos 8y + \sin 7x \sin 8y.$$

**Solution** Let

$$r_1(x) = \cos 7x, \quad r_2(x) = \sin 7x$$
$$s_1(y) = \cos 8y, \quad s_2(y) = \sin 8y.$$

Then $M(x,y) = 3r_1(x)\,s_1(y) + 5r_1(x)s_2(y) + 2r_2(x)\,s_1(y) + r_2(x)\,s_2(y)$.

$$\text{Hence,}\; a_{11} = 3, \quad a_{12} = 5, \quad b_1 = b_2 = 0$$
$$a_{21} = 2, \quad a_{22} = 1, \quad c_1 = c_2 = 0$$
$$\therefore\ \det[a_{ij}] = \begin{vmatrix} 3 & 5 \\ 2 & 1 \end{vmatrix} = 3 - 10 = -7 \neq 0.$$

Hence, this representation is canonical.

***$P^*$ and $Q^*$ spaces***

$$\text{Here } P^* = \{\rho = (r_1, r_2)' : r_1 = \cos 7t,\ r_2 = \sin 7t, \quad 0 \le t \le 1\}$$
$$Q^* = \{\sigma = (s_1, s_2)' : s_1 = \cos 8t,\ s_2 = \sin 8t, \quad 0 \le t \le 1\}$$

The $P^*$ space is the circle with its interior, and the $Q^*$ space is the circle with its interior. See Fig. 13.4a, b.

***First and second critical points***

$\prod = (p_1, p_2)'$ is the unique solution of

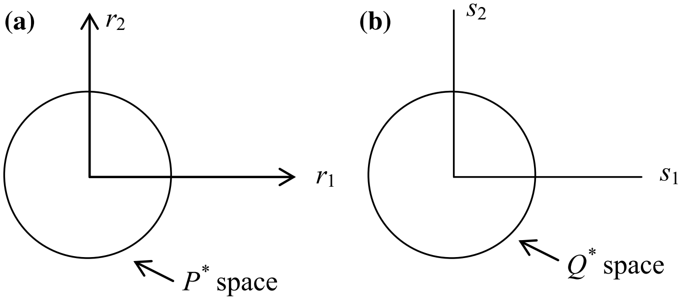



**Fig. 13.4** **a** Graphical representation of $P^*$ space. **b** Graphical representation of $Q^*$ space

$$\sum_{i=1}^{n} a_{ij} p_i + c_j = 0, \quad j = 1, 2,$$

$$\text{i.e. } 3p_1 + 2p_2 = 0$$
$$5p_1 + p_2 = 0$$
$$\Rightarrow\ p_1 = 0,\ p_2 = 0$$

$\therefore \prod = (0,\ 0)' \in$ interior of $P$-space.
Now $\chi = (q_1,\ q_2)'$ is the unique solution of

$$\sum_{j=1}^{q} a_{ij} q_j + c_i = 0, \quad i = 1, 2,$$

$$\text{i.e. } 2q_1 + 5q_2 = 0$$
$$2q_1 + q_2 = 0$$
$$\therefore q_1 = 0, \quad q_2 = 0$$

$\therefore \chi = (0, 0)' \in$ interior of $Q$.

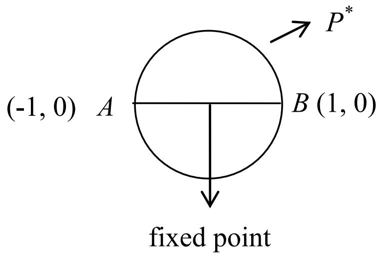
Hence, $\prod$ is the only fixed point of $P$ and $\chi$ is the only fixed point of $Q$.

***Value of the game***

The value of the game is

$$E\left(\prod, \chi\right) = \sum_{i=1}^{2} b_i p_i + d = \sum_{j=1}^{2} c_j q_j + d = 0.$$

**Fig. 13.5** Graphical representation of $P^*$ space




***Optimal strategies***

We write the fixed point of $P$, i.e. $\prod = (0,\,0)'$, as a convex combination of the points on $P^*$. See Fig. 13.5. We can obviously write

$$(0,0)' = \frac{1}{2}(-1,0)' + \frac{1}{2}(1,0)'.$$

Hence, an optimal strategy for player $-1$ is

$$F^*(x) = \frac{1}{2}F_A(x) + \frac{1}{2}F_B(x),$$

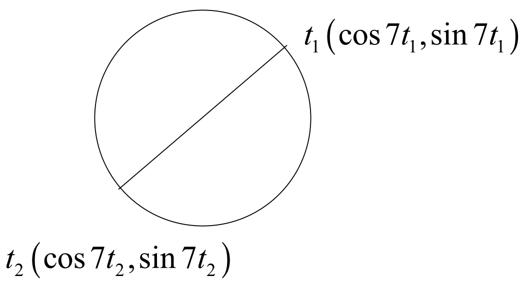
where $F_A(x)$ is a strategy corresponding to $A \equiv (-1,\,0)'$ and $F_B(x)$ is a strategy corresponding to $B = (1,\,0)'$.
Now to find $F_A(x)$, we note that it satisfies

$$-1 = \int_0^1 r_1(t)\ \mathrm{d}F_A(t) = \int_0^1 \cos 7t\,\mathrm{d}F_A(t)$$

$$0 = \int_0^1 r_2(t)\ \mathrm{d}F_A(t) = \int_0^1 \sin 7t\,\mathrm{d}F_A(t).$$

These expressions are satisfied by $F_A(x) = H\left(x - \frac{\pi}{7}\right)$.

Similarly, $F_B(x)$ satisfies

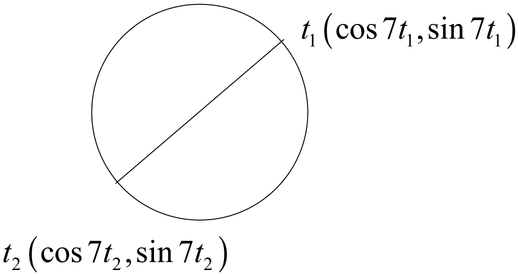


$$1 = \int_0^1 r_1(t)\ \mathrm{d}F_B(t) = \int_0^1 \cos 7t\, \mathrm{d}F_B(t)$$

$$0 = \int_0^1 r_2(t)\ \mathrm{d}F_B(t) = \int_0^1 \sin 7t\, \mathrm{d}F_B(t).$$

These are satisfied by $F_B(x) = H(x)$ or $H\left(x - \frac{2\pi}{7}\right)$.

Hence, an optimal strategy for player 1 is




$$\text{either } F^*(x) = \frac{1}{2}H\left(x - \frac{\pi}{7}\right) + \frac{1}{2}\,H(x)$$
$$\text{or } F^*(x) = \frac{1}{2}H\left(x - \frac{\pi}{7}\right) + \frac{1}{2}H\left(x - \frac{2\pi}{7}\right).$$

Hence, we can also write

$$(0,0)' = \frac{1}{2}(\cos 7t_1, \sin 7t_1)' + \frac{1}{2}(\cos 7t_2, \sin 7t_2)' \text{ where } t_2 - t_1 = \frac{\pi}{7} \quad 0 \le t_1, t_2 \le 1.$$

Hence, a most general optimal strategy for player 1 is

$$F^*(x) = \frac{1}{2}H(x - x_1) + \frac{1}{2}H(x - x_2) \text{ where } x_2 - x_1 = \frac{\pi}{7} \quad 0 \le x_1, x_2 \le 1.$$

Similarly, a most general optimal strategy for player 2 is

$$G^*(y) = \frac{1}{2}H(y-y_1) + \frac{1}{2}H(y-y_2),\ y_2 - y_1 = \frac{\pi}{8} \quad 0 \le y_1, x_2 \le 1.$$

$$\text{Also, } 0 = \int_0^1 \cos 7x \, \mathrm{d}F(x)$$

$$0 = \int_0^1 \cos 7x \, \mathrm{d}F(x)$$

are satisfied by choosing $F(x) = \begin{cases} \frac{7x}{2\pi}; & 0 \le x \le \frac{2\pi}{7} \\ 1 & \frac{2\pi}{7} \le x \le 1 \end{cases}.$

$$\text{In fact, } \int_0^1 \cos 7x \, dF(x) = \int_0^{\frac{2\pi}{7}} \cos 7x \, d\left(\frac{7x}{2\pi}\right) + \int_{\frac{2\pi}{7}}^1 \cos 7x \, d(1)$$

$$= \frac{7}{2\pi}\left[\frac{\sin 7x}{7}\right]_0^{\frac{2\pi}{7}} = 0$$

and

$$\int_0^1 \sin 7x \, \mathrm{d}F(x) = \int_0^{\frac{2\pi}{7}} \sin 7x \, \mathrm{d}\left(\frac{7x}{2\pi}\right) + \int_{\frac{2\pi}{7}}^1 \sin 7x \, \mathrm{d}(1)$$

$$= \frac{7}{2\pi}\left[-\frac{\cos 7x}{7}\right]_0^{\frac{2\pi}{7}} = -\frac{7}{2\pi}(1-1) = 0.$$

Similarly, it can easily be verified that

$$G(y) = \begin{cases} \dfrac{8y}{2\pi} & 0 \le y \le \dfrac{2\pi}{8} \\ 1 & \dfrac{2\pi}{8} \le y \le 1 \end{cases}$$

is an optimal strategy for player 2.

*Note* This is example of a game where there are an infinite number of optimal strategies.

## 13.12 Exercises

1. Solve the game whose payoff matrix is given below:

$$\begin{bmatrix} -2 & 0 & 0 & 5 & 3 \\ 3 & 2 & 1 & 2 & 2 \\ -4 & -3 & 0 & -2 & 6 \\ 5 & 3 & -4 & 2 & -6 \end{bmatrix}$$

2. Solve the following $4 \times 2$ game graphically:

$$\begin{array}{cc} & \text{Player } B \\ \text{Player} A & \begin{bmatrix} -2 & 0 \\ 3 & -1 \\ -3 & 2 \\ 5 & -4 \end{bmatrix} \end{array}$$

3. Determine the optimum minimax strategies for each player in the following games:

(i)
$$\begin{array}{cc} & \begin{array}{cccc} B_1 & B_2 & B_3 & B_4 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} & \begin{bmatrix} -5 & 2 & 0 & 7 \\ 5 & 6 & 4 & 8 \\ 4 & 0 & 2 & -3 \end{bmatrix} \end{array}$$

(ii)
$$\begin{array}{cc} & \begin{array}{ccc} B_1 & B_2 & B_3 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} & \begin{bmatrix} -2 & 15 & -2 \\ -5 & -6 & -4 \\ -5 & 20 & -8 \end{bmatrix} \end{array}$$

4. Use dominance to solve the following game:

$$\begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & 2 & 1 & 5 \\ 3 & 1 & 0 & -2 \\ 4 & 3 & 2 & 6 \end{bmatrix}$$

5. Solve the following two-person zero-sum game to find the value of the game:

$$\begin{array}{cc} & \text{Company } B \\ \text{Company } A & \begin{bmatrix} 2 & -2 & 4 & 1 \\ 6 & 1 & 12 & 3 \\ -3 & 2 & 0 & 6 \\ 2 & -3 & 7 & 7 \end{bmatrix} \end{array}$$

6. Use the notion of dominance to simplify the rectangular game with the following payoff, and then solve it graphically:

$$\begin{array}{cc} & \text{Player } K \\ & \begin{array}{cccc} I & II & III & IV \end{array} \\ \text{Player } L \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 18 & 4 & 6 & 4 \\ 6 & 2 & 13 & 7 \\ 11 & 5 & 17 & 3 \\ 7 & 6 & 12 & 2 \end{pmatrix} \end{array}$$

7. Convert the following game problem into a linear programming problem and solve:

$$\begin{array}{cc} & \text{Player } B \\ \text{Player } A & \begin{bmatrix} 5 & 7 & 2 \\ 10 & 4 & 9 \\ 6 & 2 & 0 \end{bmatrix} \end{array}$$

8. For the following two-person zero-sum game, find the strategy of each player and the value of the game:

$$\begin{array}{cc} & B \\ A & \begin{bmatrix} 8 & -3 & 7 \\ 6 & -4 & 5 \\ -2 & 2 & -3 \end{bmatrix} \end{array}$$

9. Solve the following $3 \times 3$ game:

$$\begin{bmatrix} 7 & 1 & 7 \\ 9 & -1 & 1 \\ 5 & 7 & 6 \end{bmatrix}$$

10. Consider the following $2 \times 2$ game:

$$\begin{pmatrix} 4 & 7 \\ 6 & 5 \end{pmatrix}$$

(i) Does it have a saddle point?
(ii) Is it correct to state that the value of game $G$ will satisfy $5 < G < 6$?
(iii) Determine the frequency of optimum strategies by the arithmetic method and find the value of the game.

11. Solve the following game:

$$\begin{array}{c c} & \begin{array}{cccc} I & II & III & IV \end{array} \\ \begin{array}{c} I \\ II \\ III \\ IV \end{array} & \begin{bmatrix} 3 & 2 & 4 & 0 \\ 3 & 4 & 2 & 4 \\ 4 & 2 & 4 & 0 \\ 0 & 4 & 0 & 8 \end{bmatrix} \end{array}$$

12. Solve the following game whose payoff matrix is given by the graphical method:

$$\begin{array}{c c} & \begin{array}{cccc} B_1 & B_2 & B_3 & B_4 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} & \begin{bmatrix} 4 & -2 & 3 & -1 \\ -1 & 2 & 0 & 1 \\ -2 & 1 & -2 & 0 \end{bmatrix} \end{array}$$

13. Obtain the optimal strategies for both persons and the value of the game for a two-person zero-sum game whose payoff matrix is as follows:

$$\begin{array}{c c} & \text{Player } B \\ \text{Player } A & \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 6 \\ 4 & 1 \\ 2 & 2 \\ -5 & 0 \end{bmatrix} \end{array}$$

14. Use the graphical method to solve the following game:

$$\begin{array}{c c} & \text{Player } A \\ \text{Player } B & \begin{bmatrix} 2 & 2 & 3 & -2 \\ 4 & 3 & 2 & 6 \end{bmatrix} \end{array}$$

# Chapter 14
# Project Management

## 14.1 Objectives

The objectives of this chapter are to:

- Discuss the importance of PERT and CPM techniques for project management.
- Distinguish between PERT and CPM network techniques.
- Discuss the different phases of any project and various activities to be done during these phases.
- Draw the network diagrams with single and three-time estimates of activities involved in a project.
- Determine the critical path and floats associated with non-critical activities and events along with the total project completion time.
- Determine the probability of completing a project within the scheduled time.
- Establish a time-cost trade-off for completion of a project.

## 14.2 Introduction

A project is a well-defined set of activities, jobs or tasks, all of which must be completed to finish the project. Construction of a highway or power plant, production and marketing of a new product and research & development work are the examples of projects. Such projects involve a large number of inter-related activities (or tasks) which must be completed in a specified time and in a specified sequence (or order) and require resources such as personnel, money, materials, facilities and/ or space. The main objective before starting any project is to schedule the required activities in an efficient manner so as to:

(i) Complete it on or before a specified time limit
(ii) Minimize the total time

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_14

(iii) Minimize the time for a prescribed cost
(iv) Minimize the cost for a specified time
(v) Minimize the total cost
(vi) Minimize the idle resources.

Therefore, before starting any project, it is essential to prepare a plan for scheduling and controlling the various activities involved in the project. The techniques of operations research (OR) used for planning, scheduling and controlling large & complex projects are often referred to as network analysis, network planning or network scheduling techniques. In all these techniques, a project is broken down into various activities which are arranged in a logical sequence in the form of a network. This approach helps managers to visualize a project as a number of tasks which can easily be defined in terms of duration, cost and starting time. The sequence of activities is also defined. There are two basic planning and control techniques that utilize a network to complete a predetermined project or schedule. These are PERT (program evaluation and review technique) and CPM (critical path method). A PERT network was developed during 1956–1958 by a research team of the US Navy's Polaris Nuclear Submarine Missile development project. Since 1958, this technique has been used to plan almost all types of projects. At the same time but independently, CPM was developed jointly by two companies: E. I. DuPont de Nemours and Company and Remington Rand Corporation. Other network techniques are PEP (performance evaluation programme), LCES (least cost estimating and scheduling), SCANS (scheduling and control by automated network system), etc.

## 14.3 Phases of Project Management

The work involved in a project can be divided into three phases corresponding to the management functions of planning, scheduling and control.

***Planning***

This phase involves setting the objectives of the project and the assumptions to be made. It also involves the listing of tasks or jobs that must be performed to complete a project under consideration. In this phase, the men, machines and materials required for the project in addition to the estimates of costs and duration of the various activities of the project are also determined.

***Scheduling***

This consists of organizing the activities according to the precedence order and to determine

(i) The starting and finishing times for each activity
(ii) The critical path for which the activities require special attention
(iii) The slack and float for the non-critical paths.

***Control***

This phase is exercised after the planning and scheduling and involves the following:

(i) Making periodical progress reports
(ii) Reviewing the progress
(iii) Analysing the status of the project
(iv) Management decisions regarding updating, crashing and resource allocation, etc.

## 14.4 Advantages of Network Analysis

The network analysis:

(i) Shows the inter-relationships of all jobs in the project
(ii) Gives a clear picture, other than a typical bar chart, of the relationship controlling the order of performance of various activities
(iii) Helps in communication of ideas. The pictorial approach helps to clarify the verbal instructions
(iv) Provides time schedules containing much more information than other methods such as bar charts
(v) Identifies jobs which are critical for a project completion date
(vi) Permits an accurate forecast of resource requirements
(vii) Provides a method of resource allocation to meet the limiting condition and to maintain or minimize the overall costs
(viii) Integrates all elements of a programme to whatever detail is desired by management
(ix) Relates time to costs, which allows a money value to be placed on proposed changes.

## 14.5 Basic Components

There are two basic components in a network. These are:

(i) Event/node
(ii) Activity.

### *Event/node*

An event/node is a particular instant in time showing the end or beginning of one or more activities. It is a point of accomplishment or decision. The starting and end points of an activity are thus described by two events, usually called the tail event and head event. An event is generally represented by a circle, rectangle, hexagon or some other geometric shape. These geometric shapes are numbered for distinguishing one activity from another. The occurrence of an event indicates that the work has been accomplished up to that point (See Fig. 14.1).

### *Merge and burst event*

It is necessary for an event to be the ending event of at least one activity, but it can also be the ending event of two or more activities. Then it is defined as a merge event.
If an event is the beginning event of two or more activities, it is defined as a burst event. See Fig. 14.2.

### *Activity*

An activity is a task or item of work to be done that consumes time, effort, money or other resources. Activities are represented by arrows.

Activities are identified by the numbers of their starting (tail) event and ending (head) event. Generally, an ordered pair $(i, j)$ represents an activity where events $i$ and $j$ represent the starting and ending of the activity respectively. It can also be denoted by $i - j$. Sometimes, activities are denoted by capital letters (Fig. 14.3).

The activities can further be classified into different categories:

(i) *Predecessor activity*: An activity which must be completed before one or more other activities start is known as a predecessor activity.
(ii) *Successor activity*: An activity which started immediately after the completion of one or more other activities is known as a successor activity.
(iii) *Dummy activity*: In connecting events by activities showing their interdependencies, very often a situation arises where a certain event $j$ cannot occur until another event $i$ has taken place, but the activity connecting $i$ and $j$ does not involve any time or expenditure of other resources. In such a case, the activity is called a dummy activity. It is depicted by a dotted line in the network diagram.

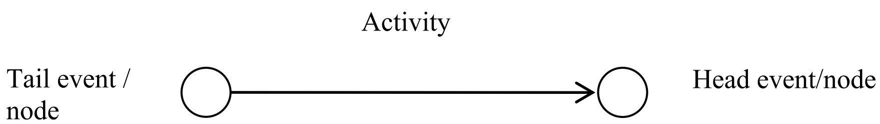



**Fig. 14.1** Head and trail events

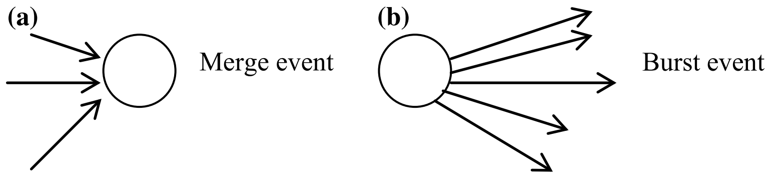



**Fig. 14.2** **a** Merge event. **b** Burst event

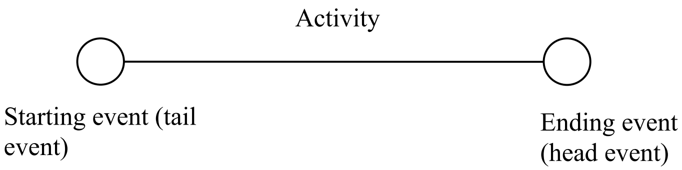



**Fig. 14.3** Activity

Let us consider an example of a car taken to a garage for cleaning. Both the inside and the outside of the car are to be cleaned before it is taken away from the garage. The events can be stated as follows:

Event 1: Starting of car from house
Event 2: Parking of car in garage
Event 3: Completion of outside cleaning
Event 4: Completion of inside cleaning
Event 5: Taking of car from garage
Event 6: Parking of car in house.

The network diagram for the problem is given in Fig. 14.4.

It is assumed that the inside cleaning and outside cleaning can be done concurrently by two garage assistants. Activities *B* and *C* represent these cleaning operations. What do activities *D* and *E* stand for? Their time consumptions are zero, but they express the condition that events 3 and 4 must occur before event 5 can take place. Activities *D* and *E* are called dummy activities.

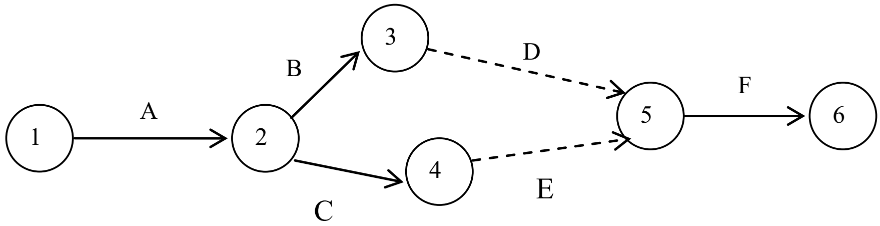



**Fig. 14.4** Project network

***Network***

A network is the graphical representation of logically and sequentially connected arrows and nodes representing the activities and events of a project. Networks are also called arrow diagrams.

***Path***

An unbroken chain of activity arrows connecting the initial event to some other event is called a path.

## 14.6 Common Errors

There are three common errors in a network construction: looping, dangling and redundancy.

***Looping (cycling)***

In a network diagram, a looping error is also known as a cycling error. The drawing of an endless loop in a network is known as an error of looping. A looping network is shown in Fig. 14.5.

***Dangling***

To disconnect an activity before the completion of all the activities in a network diagram is known as dangling. It should be avoided. In that case, a dummy activity is introduced in order to maintain the continuity of the system (see Fig. 14.6).

***Redundancy***

If a dummy activity is the only activity emanating from an event and it can be eliminated, this is known as redundancy (see Fig. 14.7).

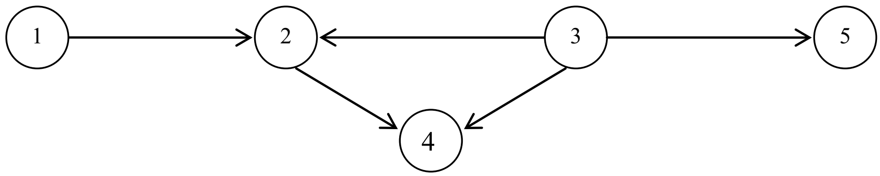



**Fig. 14.5** Looping in project network

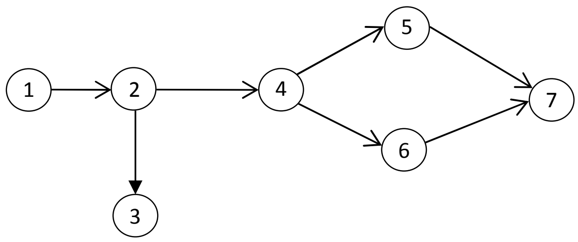



**Fig. 14.6** Dangling in project network

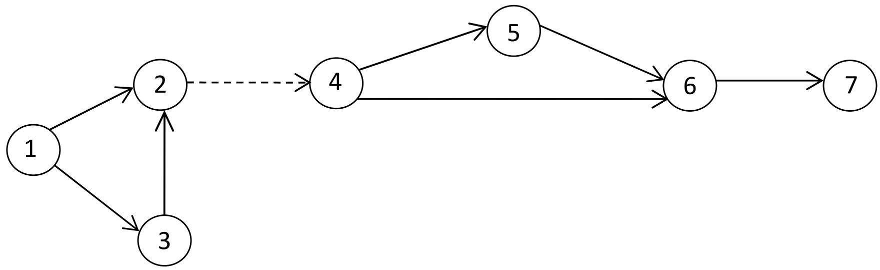



**Fig. 14.7** Redundancy in project network

## 14.7 Rules for Network Construction

For the construction of a network, certain rules are generally followed:

(i) Each activity is represented by one and only one arrow.
(ii) Crossing of arrows and curved arrows should be avoided; only straight arrows are to be used.
(iii) Each activity must be identified by its starting and ending nodes.
(iv) No event can occur until every activity preceding it has been completed.
(v) An event cannot occur twice; i.e. there must be no loops.
(vi) An activity succeeding an event cannot be started until that event has occurred.
(vii) Events are numbered to identify an activity uniquely. The number of the tail event (starting event) should be lower than that of the head (ending) event of an activity.
(viii) Between any pair of nodes (events), there should be one and only one activity. However, more than one activity may emanate from a node or terminate at a node.
(ix) Dummy activities should be introduced only if it is extremely necessary.
(x) The network has only one entry point, called the starting event, and one terminal point, called the end or terminal event.

## 14.8 Numbering of Events

After the network is drawn in a logical sequence, every event is assigned by a number. The number sequence must reflect the flow of the network. In numbering the events, Fulkerson's rules (for D. R. Fulkerson) are used. These rules are as follows:

(i) The event number should be unique.
(ii) Event numbering should be carried out on a sequential basis from left to right.
(iii) An initial event is one which has all outgoing arrows with no incoming arrow. In any network, there will be one such event. Number it as 1 (one).
(iv) Delete all arrows emerging from event 1 (one). This will create at least one more initial event.
(v) Number these initial events as 2, 3, …, etc.
(vi) Delete all emerging arrows from these numbered events, which will create new initial events.
(vii) Repeat steps (v) and (vi) until the last event is obtained which has no arrows emerging from it. Number the last event.

*Example 1* Construct a network of a project whose activities and their precedence relationships are given below. Then number the events (Fig. 14.8).

| Activity | *A* | *B* | *C* | *D* | *E* | *F* | *G* | *H* | *I* |
|---|---|---|---|---|---|---|---|---|---|
| Immediate predecessor | – | *A* | *A* | – | *D* | *B, C, E* | *F* | *D* | *G, H* |

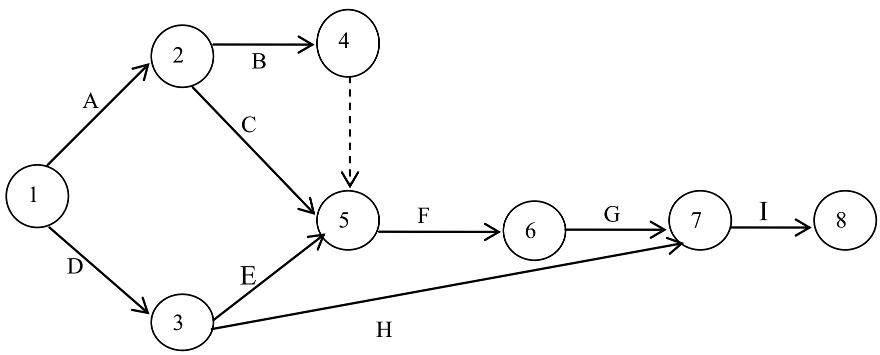



**Fig. 14.8** Construction of project network

## 14.9 Critical Path Analysis

Once the network of a project is constructed, the time analysis of the network becomes essential for planning various activities of the project. The main objective of the time analysis is to prepare a planning schedule of the project. The planning schedule should include the following factors:

(i) Total completion time for the project
(ii) Earliest time when each activity can start
(iii) Latest time when each activity can be started without delaying the total project
(iv) Float for each activity, i.e. the duration of time by which the completion of an activity can be delayed without delaying the total project completion
(v) Identification of critical activities and critical path.

***Notation***

The following notation is used in this analysis:

$E_i$ = Earliest occurrence time of event $i$, i.e. the earliest time at which the event $i$ can occur without affecting the total project duration
$L_i$ = Latest allowable occurrence time of event $i$. It is the latest allowable time at which an event can occur without affecting the total project duration
$t_{ij}$ = Duration of activity $(i, j)$
$\mathrm{ES}_{ij}$ = Earliest starting time of activity $(i, j)$
$\mathrm{LS}_{ij}$ = Latest starting time of activity $(i, j)$
$\mathrm{EF}_{ij}$ = Earliest finishing time of activity $(i, j)$
$\mathrm{LF}_{ij}$ = Latest finishing time of activity $(i, j)$.

The critical path calculations are done in the following two ways:

(a) Forward pass calculations method
(b) Backward pass calculations method.

***Forward pass calculations method***

In this method, calculation begins from the initial event, proceeds through the events in an increasing order of event numbers and ends at the final event of the network. At each node (event), the earliest starting and finishing times are calculated for each activity. The method may be summarized as follows:

Step 1 Set $E_1 = 0$, $i = 1$.
Step 2 Calculate the earliest starting time $\mathrm{ES}_{ij}$ for each activity that begins at event $i$, i.e. $\mathrm{ES}_{ij} = E_i$ for all activities $(i, j)$ that start at node $i$.
Step 3 Calculate the earliest finishing time $\mathrm{EF}_{ij}$ of each activity that begins at event $i$ by adding the earliest starting time of the activity with the duration of the activity. Thus, $\mathrm{EF}_{ij} = \mathrm{ES}_{ij} + t_{ij} = E_i + t_{ij}$.

Step 4 Go to the next event (node), say event $j(j > i)$, and compute the earliest occurrence time for event $j$. This is the maximum of the earliest finishing times of all activities ending at that event, i.e.
$E_j = \underset{i}{\text{Max}}\{\text{EF}_{ij}\} = \underset{i}{\text{Max}}\{E_i + t_{ij}\}$ for all immediate predecessor activities.

Step 5 If $j = n$ (final event number), then the earliest finishing time for the project is given by
$E_n = \underset{i}{\text{Max}}\{\text{EF}_{ij}\} = \underset{i}{\text{Max}}\{E_i + t_{ij}\}$ for all terminal activities.

***Backward pass calculations method***

In this method, calculations begin from the terminal event, proceed through the events in a decreasing order of event numbers and end at the initial event of the network. At each node (event), the latest starting and finishing times are calculated for each activity. The method may be summarized as follows:

Step 1 Set $L_n = E_n$, $j = n$.

Step 2 Calculate the latest finishing time $\text{LF}_{ij}$ for each activity that ends at event $j$, i.e. $\text{LF}_{ij} = L_j$ for all activities $(i, j)$ that end at node $j$.

Step 3 Calculate the latest starting time $\text{LS}_{ij}$ of each activity that ends at event $j$ by subtracting the duration of each activity from the latest finishing time of the activity. Thus,

$$\text{LS}_{ij} = \text{LF}_{ij} - t_{ij} = L_j - t_{ij}.$$

Step 4 Proceed backward to the node in the sequence that decreases $j$ by 1. Also compute the latest occurrence time of node $i$ $(i < j)$. This is the minimum of the latest starting times of all activities starting from that event, i.e.
$L_i = \underset{j}{\text{Min}}\{\text{LS}_{ij}\} = \underset{j}{\text{Min}}\{L_j - t_{ij}\}$ for all immediate successor activities.

Step 5 If $i = 1$ (initial node), then
$L_1 = \underset{j}{\text{Min}}\{\text{LS}_{ij}\} = \underset{j}{\text{Min}}\{L_j - t_{ij}\}$ for all initial activities.

***Determination of floats and slack times***

When the network is completely drawn and properly labelled, and the earliest and latest event times are computed, the next objective is to determine the floats of each activity and the slack time of each event.

The float of an activity is the amount of time by which it is possible to delay its completion time without affecting the total project completion time. There are three types of activity floats:

(i) Total float
(ii) Free float
(iii) Independent float.

### *Total float*

The total float of an activity represents the amount of time by which an activity can be delayed without a delay in the project completion time.

Mathematically, the total float of an activity (*i, j*) is the difference between the latest start time and earliest start time of that activity (or the difference between the earliest finish time and latest finish time). Hence, the total float for an activity (*i, j*) is denoted by $\mathrm{TF}_{ij}$ and is computed by the following formula:

$$\mathrm{TF}_{ij} = \mathrm{LS}_{ij} - \mathrm{ES}_{ij} \text{ or } \mathrm{TF}_{ij} = \mathrm{LF}_{ij} - \mathrm{EF}_{ij} \text{ or } \mathrm{TF}_{ij} = L_j - \left(E_i + t_{ij}\right) \text{ as } L_j = \mathrm{LF}_{ij}$$
$$\text{and } \mathrm{EF}_{ij} = E_i + t_{ij}.$$

### *Free float*

Sometimes we may need to know how much an activity's completion time may be delayed without causing any delay in its immediate successor activities. This amount of float is called a free float. Mathematically, the free float for an activity (*i, j*) is denoted by $\mathrm{FF}_{ij}$ and is computed by

$$\begin{aligned} & \mathrm{FF}_{ij} = E_j - E_i - t_{ij} \\ \text{as} \quad & \mathrm{TF}_{ij} = L_i - E_i - t_{ij} \text{ and } L_i \geq E_j \\ \therefore \quad & \mathrm{TF}_{ij} \geq E_j - E_i - t_{ij} \quad \text{i.e.} \quad \mathrm{TF}_{ij} \geq \mathrm{FF}_{ij}. \end{aligned}$$

Hence, for all activities, free float can take values from zero up to total float, but it does not exceed total float.

Again, free float is very useful for rescheduling activities with a minimum disruption of earlier plans.

### *Independent float*

In some cases, the delay in the completion of an activity affects neither its predecessor nor its successor activities. This amount of delay is called an independent float. Mathematically, independence of an activity (*i, j*) denoted by $\mathrm{IF}_{ij}$ is computed by the following formula:

$$\mathrm{IF}_{ij} = E_j - L_i - t_{ij}.$$

A negative independent float is always taken as zero.

### *Event slack or event float*

The slack of an event is the difference between its latest time and its earliest time. Hence, for an event *i*,

$$\text{slack} = L_i - E_i.$$

*Critical event*: An event is said to be critical if its slack is zero, i.e. $L_i = E_i$ for the $i$th event.
*Critical activity*: An activity is critical if its total float is zero, i.e. $LS_{ij} = ES_{ij}$ or $LF_{ij} = EF_{ij}$ for an activity $(i, j)$.
Otherwise, an activity is called non-critical.

***Critical path***

The critical path is the continuous chain or sequence of critical activities in a network diagram path from starting event to ending event. It is shown by a dark line or double lines to make a distinction from other, non-critical paths.

The length of the critical path is the sum of the individual times of all critical activities lying on it and defines the minimum time required to complete the project.

The critical path on a network diagram can be identified as follows:

(i) For all activities $(i, j)$ lying on the critical path, the $E$ values and $L$ values for tail and head events are equal, i.e. $E_j = L_j$ and $E_i = L_i$.
(ii) On the critical path, $E_j - E_i = L_j - L_i = t_{ij}$.

**Main features of the critical path**

The critical path has two main features:

(i) If the project has to be shortened, then some of the activities on that path must be shortened. The application of additional resources on other activities will not give the desired results unless that critical path is shortened first.
(ii) The variation in actual performance from the expected activity duration time will be completely reflected in one-to-one fashion in the anticipated completion of the whole project.

*Example 2* Determine the critical path and minimum time of completion of the project whose network diagram is shown in Fig. 14.9. Also find the different floats.

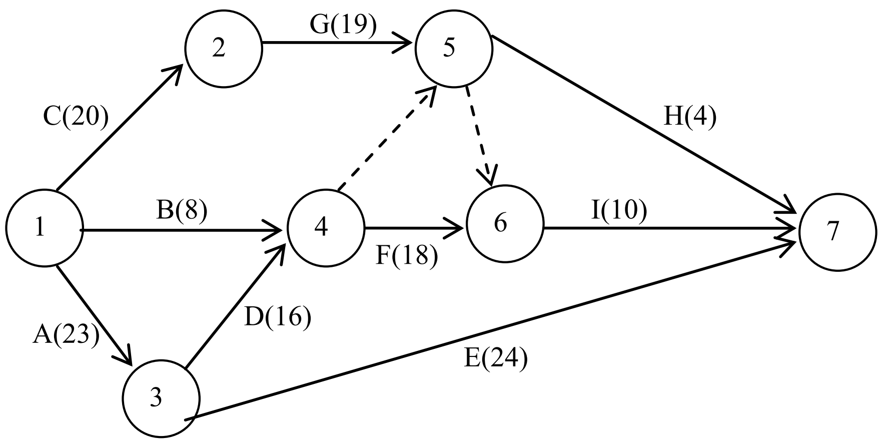



**Fig. 14.9** Project network

**Solution**
**Forward pass calculations**

At node 1: Set $E_1 = 0$
At node 2: $E_2 = E_1 + t_{12} = 0 + 20 = 20$
At node 3: $E_3 = E_1 + t_{13} = 0 + 23 = 23$
At node 4: $E_4 = \underset{i=1,3}{\text{Max}}\{E_i + t_{i4}\} = \text{Max}\{E_1 + t_{14}, E_3 + t_{34}\} = \text{Max}\{0 + 8, 23 + 16\} = 39$
At node 5: $E_5 = \underset{i=2,4}{\text{Max}}\{E_i + t_{i5}\} = \text{Max}\{E_2 + t_{25}, E_4 + t_{45}\} = \text{Max}\{20 + 19, 39 + 0\} = 39$
At node 6: $E_6 = \underset{i=4,5}{\text{Max}}\{E_i + t_{i6}\} = \text{Max}\{E_4 + t_{46}, E_5 + t_{56}\} = \text{Max}\{39 + 18, 39 + 0\} = 57$
At node 7: $E_7 = \underset{i=3,5,6}{\text{Max}}\{E_i + t_{i7}\} = \text{Max}\{E_3 + t_{37}, E_5 + t_{57}, E_6 + t_{67}\}$

$$= \text{Max}\{23 + 24, 39 + 4, 57 + 10\} = 67$$

**Backward pass calculations**

At node 7: Set $L_7 = E_7 = 67$
At node 6: $L_6 = L_7 - t_{67} = 67 - 10 = 57$
At node 5: $L_5 = \underset{j=6,7}{\text{Min}}\{L_j - t_{5j}\} = \text{Min}\{L_6 - t_{56}, t_7 - t_{57}\} = \text{Min}\{57 - 0, 67 - 4\} = 57$
At node 4: $L_4 = \underset{j=5,6}{\text{Min}}\{L_j - t_{4j}\} = \text{Min}\{L_5 - t_{45}, L_6 - t_{46}\} = \text{Min}\{57 - 0, 57 - 18\} = 39$
At node 3: $L_3 = \underset{j=4,7}{\text{Min}}\{L_j - t_{3j}\} = \text{Min}\{L_4 - t_{34}, L_7 - t_{37}\} = \text{Min}\{39 - 16, 67 - 24\} = 23$
At node 2: $L_2 = L_5 - t_{25} = 57 - 19 = 38$
At node 1: $L_1 = \underset{j=2,3,4}{\text{Min}}\{L_j - t_{1j}\} = \text{Min}\{L_2 - t_{12}, L_3 - t_{13}, L_4 - t_{14}\}$

$$= \text{Min}\{38 - 20, 23 - 23, 39 - 8\} = 0$$

To find the critical activities and different floats, we construct the table (cf. Table 14.1).

**Table 14.1** Computation of critical activity and different floats

| Activity | Duration of activity | Earliest time | | Latest time | | Float | | | Critical activity |
|---|---|---|---|---|---|---|---|---|---|
| | | Start ($E_i$) | Finish ($E_i + t_{ij}$) | Start ($L_j - t_{ij}$) | Finish ($L_j$) | Total $L_j - E_i - t_{ij}$ | Free $E_j - E_i - t_{ij}$ | Independent $E_j - L_i - t_{ij}$ | |
| (1, 2) | 20 | 0 | 20 | 18 | 38 | 18 | 0 | 0 | |
| (1, 3) | 23 | 0 | 23 | 0 | 23 | 0 | 0 | 0 | (1, 3) |
| (1, 4) | 8 | 0 | 8 | 31 | 39 | 31 | 31 | 31 | |
| (2, 5) | 19 | 20 | 39 | 38 | 57 | 18 | 0 | 0 | |
| (3, 4) | 16 | 23 | 39 | 23 | 39 | 0 | 0 | 0 | (3, 4) |
| (3, 7) | 24 | 23 | 47 | 43 | 67 | 20 | 20 | 20 | |

(continued)

**Table 14.1** (continued)

| Activity | Duration of activity | Earliest time | | Latest time | | Float | | | Critical activity |
|---|---|---|---|---|---|---|---|---|---|
| | | Start ($E_i$) | Finish ($E_i$+ $t_{ij}$) | Start ($L_j - t_{ij}$) | Finish ($L_j$) | Total $L_j - E_i - t_{ij}$ | Free $E_j - E_i - t_{ij}$ | Independent $E_j - L_i - t_{ij}$ | |
| (4, 5) | 0 | 39 | 39 | 57 | 57 | 18 | 0 | 0 | |
| (4, 6) | 18 | 39 | 57 | 39 | 57 | 0 | 0 | 0 | (4, 6) |
| (5, 6) | 0 | 39 | 39 | 57 | 57 | 18 | 18 | 0 | |
| (5, 7) | 4 | 39 | 43 | 63 | 67 | 24 | 24 | 46 | |
| (6, 7) | 10 | 57 | 67 | 57 | 67 | 0 | 0 | 0 | (6, 7) |

From the preceding table, it is clear that the critical activities (zero total float) are (1, 3), (3, 4), (4, 6) and (6, 7). Hence, the critical path is 1–3−4–6–7, and the duration of the project is 67 time units (as $E_7 = L_7 = 67$).

## 14.10 PERT Analysis

***Time estimates***

It is very difficult to estimate the time required for the execution of each activity or because of various uncertainties. Taking the uncertainties into account, three types of time estimates are generally obtained.

The PERT system is based on these three time estimates of the performance time of an activity:

(i) *Optimistic time* ($t_o$): This is the estimate of the shortest possible time in which an activity can be completed under ideal conditions.

(ii) *Pessimistic time* ($t_p$): This is the maximum time which is required to perform the activity under extremely bad conditions. However, such conditions do not include labour strikes or acts of nature (like floods, earthquakes, tornados, etc.).

(iii) *Most likely time* ($t_m$): This is the estimate of the normal time which an activity would require. This time estimate lies between the optimistic and pessimistic time estimates. Statistically, it is the modal value of the duration of the activity.

The range specified by the optimistic time ($t_o$) and pessimistic time ($t_p$) estimates supposedly must close every possible estimate of duration of the activity. The most likely time ($t_m$) estimate may not coincide with the midpoint $t_{mid} = (t_o + t_p)/2$ and may occur to its left or right, as shown in Fig. 14.10.

Keeping in mind the above-described time properties, it may be justified to assume that the duration of each activity may follow a beta ($\beta$)-distribution with its unimodal point occurring at $t_m$ and its end point at $t_o$ and $t_p$.

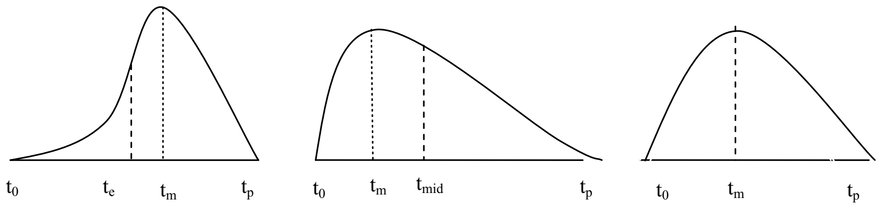



**Fig. 14.10** Time estimates of an activity

The expected or mean value of an activity duration can be approximated by a linear combination of the three time estimates or by the weighted average of three time estimates $t_o$, $t_p$ and $t_m$, i.e. $t_e = (t_o + 4t_m + t_p)/6$.

Again, to determine the activity duration variance in PERT, the unimodal property of the $\beta$-distribution is used. However, in PERT, the standard deviation is expressed as

$$\sigma = \frac{1}{6}\left(t_p - t_o\right) \text{ or variance } \sigma^2 = \left(\frac{t_p - t_o}{6}\right)^2.$$

Note that in PERT analysis, the $\beta$-distribution is assumed as it is unimodal, has non-negative end points and is approximately symmetric.

***Probability of meeting the schedule time***

After identifying the critical path and the occurrence time of all activities, there arises a question: What is the probability that a particular event will occur on or before the schedule date? This particular event may be any event in the network.

Let us recall that the expected time of an activity is the weighted average of the three time estimates $t_o$, $t_p$ and $t_m$,

$$\text{i.e. } t_e = \left(t_o + 4t_m + t_p\right)/6.$$

The probability that the activity $(i, j)$ will be completed in time $t_e$ is 0.5; i.e. the chance of completion of that activity is 50%. In the frequency distribution curve, for the activity $(i, j)$ the vertical line through $t_e$ will divide the area under the curve into two equal parts, as shown in Fig. 14.11.

For completing the activity in any other time $t_k$, the probability will be

$$p = \frac{\text{Area under } AEK}{\text{Area under } AEB}.$$

A project consists of a number of activities. All activities, as we know, are independent random variables, and hence the length of the project up to a certain event through a certain path is also a random variable. But the point of difference is that

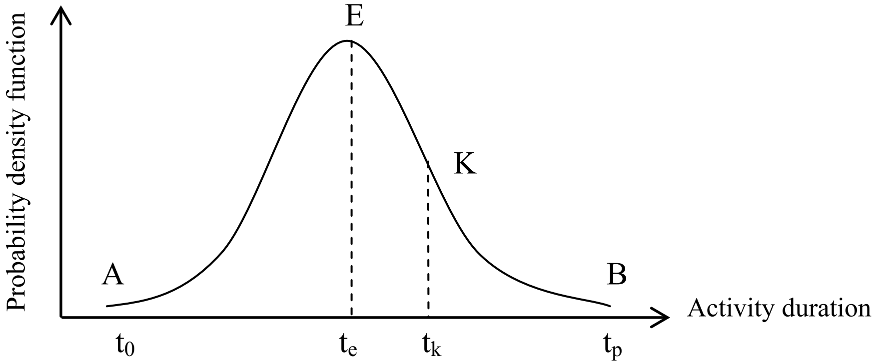



**Fig. 14.11** Time distribution curve of an activity

the expected project length $T_e$ does not have the same probability distribution as the expected activity time $t_e$. While a $\beta$-distribution curve approximately represents the activity time probability distribution, the project expected time $T_e$ follows approximately a standard normal distribution. This standard normal distribution curve has an area equal to unity and standard deviation 1 and is symmetrical about the mean, as shown in Fig. 14.12.

The probability of completing a project in scheduling time $T_s$ is given by

$$prob(T_s) = \frac{\text{Area under } ACS}{\text{Area under } ACB}.$$

Probability $prob(T_s)$ depends upon the location of $T_s$. Taking $T_e$ as a reference point, the distance $T_s - T_e$ can be expressed in terms of the standard deviation for a network, calculated as

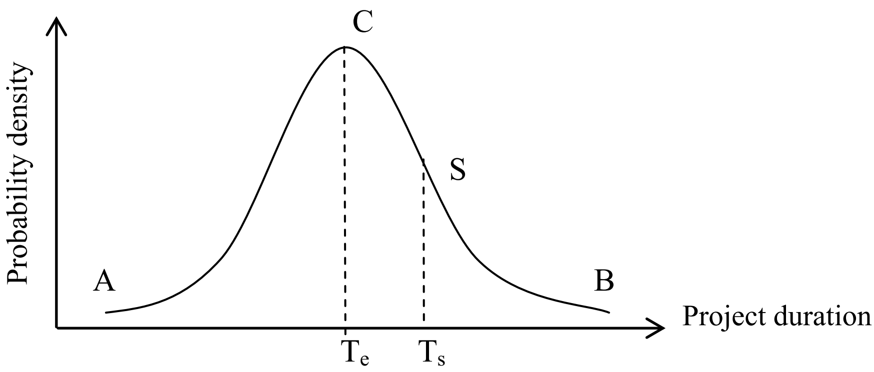



**Fig. 14.12** Time distribution of a project network

$$\text{Standard deviation for a network} = \sigma_e$$
$$= \sqrt{\text{sum of the variances along the critical path}},$$

i.e. $\delta$ for a network $= \sqrt{\sum \sigma_{ij}^2}$, where $\sigma_{ij}^2$ for an activity $(i, j) = \left(\frac{t_p - t_o}{6}\right)^2$.

Since the standard deviation for a standard normal curve is unity, the standard deviation $\sigma_e$ calculated above is used as a scale factor for calculation of the normal deviate,

$$\text{The normal deviate } D = \frac{T_s - T_e}{\sigma_e}.$$

Hence, the probability of completing the project by scheduling time ($T_s$) is given by $prob(Z \leq D)$, where $D = \frac{T_s - T_e}{\sigma_e}$ and $Z$ is the standard normal variate.

The values of the probabilities for a normal distribution curve corresponding to the different values of the normal deviate are available from the table of the standard normal curve.

*Example 3* A small project is composed of seven activities whose time estimates (in weeks) are listed in the Table 14.2.

(a) Draw the project network.
(b) Find the expected duration and variance of each activity.
(c) Calculate the earliest and latest occurrence time for each event and the expected project length.
(d) Calculate the variance and standard deviation of the project length.
(e) What is the probability that the project will be completed:
    (i) At least 4 weeks earlier than expected?
    (ii) Not more than 4 weeks later than expected?
(f) If the project due date is 19 weeks, what is the probability of meeting the due date?
(g) Find also the schedule time at which the project will be completed with a probability 0.90.

**Table 14.2** Activities and their different time estimates

| Activity | (1, 2) | (1, 3) | (1, 4) | (2, 5) | (3, 5) | (4, 6) | (5, 6) |
|---|---|---|---|---|---|---|---|
| $t_0$ | 1 | 1 | 2 | 1 | 2 | 2 | 3 |
| $t_m$ | 1 | 4 | 2 | 1 | 5 | 5 | 6 |
| $t_p$ | 7 | 7 | 8 | 1 | 14 | 8 | 15 |

**Solution**

(a) The project network diagram is shown in Fig. 14.13.
(b) The expected time and variance of each activity is computed and displayed in the Table 14.3.

**Table 14.3** Computation of expected time and variance of each activity

| Activity | $t_o$ | $t_m$ | $t_p$ | $t_e = \frac{t_o + 4t_m + t_p}{6}$ | $\sigma^2 = \left(\frac{t_p - t_o}{6}\right)^2$ |
|---|---|---|---|---|---|
| (1, 2) | 1 | 1 | 7 | 2 | 1 |
| (1, 3) | 1 | 4 | 7 | 4 | 1 |
| (1, 4) | 2 | 2 | 8 | 3 | 1 |
| (2, 5) | 1 | 1 | 1 | 1 | 0 |
| (3, 5) | 2 | 5 | 14 | 6 | 4 |
| (4, 6) | 2 | 5 | 8 | 5 | 1 |
| (5, 6) | 3 | 6 | 15 | 7 | 4 |

Let $E_i$ be the earliest occurrence time of event $i$.
Set $E_1 = 0$
$E_2 = E_1 + t_{12} = 0 + 2 = 2$
$E_3 = E_1 + t_{13} = 0 + 4 = 4$
$E_4 = E_1 + t_{14} = 0 + 3 = 3$

$$E_5 = \underset{i=2,3}{\text{Max}}\{E_i + t_{i5}\} = \text{Max}\{E_2 + t_{25}, E_3 + t_{35}\} = \text{Max}\{2+1, 4+6\} = 10$$

$$E_6 = \underset{i=4,5}{\text{Max}}\{E_i + t_{i6}\} = \text{Max}\{E_4 + t_{46}, E_5 + t_{56}\} = \text{Max}\{3+5, 10+7\} = 17$$

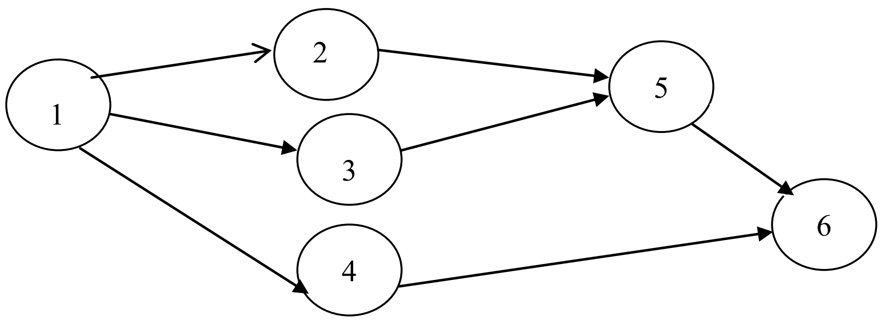



**Fig. 14.13** Project network diagram

(c) **Backward pass calculations**

Let $L_j$ be the latest occurrence time of event $j$.
Set $L_6 = E_6 = 17$
$L_5 = L_6 - t_{56} = 17 - 7 = 10$
$L_4 = L_6 - t_{46} = 17 - 5 = 12$
$L_3 = L_5 - t_{35} = 10 - 6 = 4$
$L_2 = L_5 - t_{25} = 10 - 1 = 9$

$$\begin{aligned} L_1 &= \underset{j=2,3,4}{\text{Min}}\left\{L_j - t_{1j}\right\} = \text{Min}\{L_2 - t_{12}, L_3 - t_{13}, L_4 - t_{14}\} \\ &= \text{Min}\{9 - 2, 4 - 4, 12 - 3\} = 0. \end{aligned}$$

From these calculations, it is seen that

$$E_1 = L_1, E_3 = L_3, E_5 = L_5, E_6 = L_6.$$

Hence, the critical events are 1, 3, 5, 6 and the critical path is 1–3–5–6. Also, the expected project length = $E_6 = L_6 = 17$ weeks.

(d) The variance of the project length is given by

$$\sigma_e^2 = 1 + 4 + 4 = 9 \quad \text{or}, \sigma_e = 3.$$

(e) The standard normal deviate is given by

$$D = \frac{\text{schedule time} - \text{expected time of completion}}{\sqrt{\text{variance or standard deviation}}}.$$

(i) Now, the probability that the project will be completed at least 4 weeks earlier than expected is given by

$$\begin{aligned} & prob(Z \le D) \text{ where } D = \frac{(17-4)-17}{3} = \frac{-4}{3} = -1.33 \text{ (approx.)} \\ & = prob(Z \le -1.33) \\ & = 0.5 - Prob(0 < Z \le 1.33) \\ & = 0.5 - \varphi(1.33) \\ & = 0.5 - 0.4082 \qquad \text{[From the table of the area under the standard normal curve]} \\ & = 0.0918. \end{aligned}$$

(ii) Again, the probability that the project will be completed not more than 4 weeks later than expected = the probability that the project will be completed within 17 + 4 weeks, i.e. within 21 weeks:

$$prob(Z \leq D), \text{where } D = \frac{21 - 17}{3} = \frac{4}{3} = -1.33 \text{ (approx.)},$$
$$= prob(Z \leq 1.33)$$
$$= 0.5 + \varphi(1.33) = 0.5 + 0.4082 = 0.9082.$$

(f) When the due date is 19 weeks, $D = \frac{19-17}{3} = \frac{2}{3} = 0.67$.

Then the probability of meeting the due date is given by

$$prob(Z \leq 0.67) = 0.5 + \varphi(0.67) = 0.5 + 0.2514 = 0.7514.$$

(g) Since the probability for the completion of the project is 0.90,

$prob(Z \leq D) = 0.90$, where $D = \frac{T_s - T_e}{\sigma}$ and $T_s$ is the schedule time.

As $prob(Z \leq 1.29) = 0.90$,

$\therefore D = 1.29$, which implies $\frac{T_s - 17}{3} = 1.29$, i.e. $T_s = 17 + 3 \times 1.29$, i.e. $T_s = 20.87$.

## 14.11 Difference Between PERT and CPM

The differences between PERT and CPM are listed as follows:

| PERT | CPM |
|---|---|
| 1. This technique was developed in connection with R&D (research and development) works; therefore, it had to cope with the uncertainty which is associated with R&D activities. In this case, the total project duration is regarded as a random variable. As a result, multiple time estimates are made to calculate the probability of completing the project within the schedule time. Therefore, it is a probabilistic model. | 1. This technique was developed in connection with a construction and maintenance project which consists of routine tasks or jobs whose resource requirements and duration are known with certainty. Therefore, it is basically a deterministic model. |
| 2. It is used for projects involving activities of a non-repetitive nature. | 2. It is used for projects involving activities of a repetitive nature. |
| 3. It is an event-oriented technique, because the results of analyses are expressed in terms of events or distinct points in time indicative of progress. | 3. It is an activity oriented technique, as the results of calculations are considered in terms of activities. |
| 4. It incorporates statistical analysis and thereby enables the determination of probabilities concerning the time by | 4. It does not incorporate statistical analysis in determining the estimates because the time is precise and known. |

(continued)

(continued)

| PERT | CPM |
|---|---|
| which each activity and the entire project would be completed. | |
| 5. It serves as a useful control device, as it assists management in controlling a project by calling attention through constant review to delays in activities which might lead to a delay in the project completion date. | 5. It is difficult to use this technique as a controlling device, because one must repeat the entire evaluation of the project each time changes are introduced into the network. |

## 14.12 Project Time-Cost Trade-Off

In this section, the costs of resources consumed by activities are taken into consideration. The project completion time can be reduced (crashing) by reducing the normal completion time of critical activities. The reduction in normal time of completion will increase the total budget of the project. However, the decision-maker will always look for a trade-off between the total cost of the project and the total time required to complete it.

**Project cost**

In order to include the cost aspect in project scheduling we have to find the cost duration relationships for various activities in the project. The total cost of any project comprises direct and indirect costs.

***Direct cost***

This cost is directly dependent upon the amount of resources in the execution of individual activities such as manpower loading, material consumed, etc. The direct cost increases if the activity duration is to be reduced.

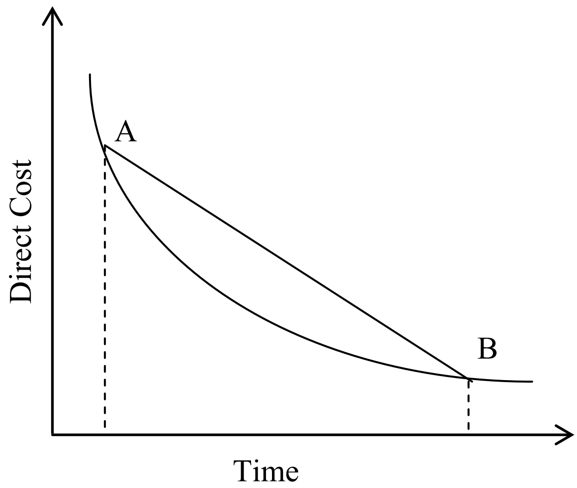
***Indirect cost***

This cost is associated with expenditures which cannot be allocated to individual activities of the project. This cost may include managerial services, loss of revenue, fixed overheads, etc. The indirect cost is computed on a per-day, per-week or per-month basis. This cost decreases if the activity duration is to be reduced.

The network diagram can be used to identify the activities whose duration should be shortened so that the completion time of the project can be shortened in the most economic manner. The process of reducing the activity duration by putting in extra effort is called crashing the activity.

The crash time ($T_c$) represents the minimum activity duration time that is possible; any attempts to further crash would only raise the activity cost without reducing the time. The activity cost corresponding to the crash time is called the crash cost ($C_c$), which is the minimum direct cost required to achieve the crash performance time.

**Fig. 14.14** Direct cost w.r.t. time




The normal cost ($C_n$) is equal to the absolute minimum of the direct cost required to perform an activity. The corresponding time duration taken by an activity is known as the normal time ($T_n$).

The direct cost curve (from the relationship of direct cost and time) is shown in Fig. 14.14. The point *B* denotes the normal time for completion of an activity, whereas point *A* denotes the crash time which indicates the least duration in which activity can be completed. The cost curve has a non-linear and asymptotic nature. But, for simplicity, it can be approximated by a straight line whose slope (in magnitude) is given by

$$\text{Cost slope} = \frac{\text{Crash cost} - \text{Normal cost}}{\text{Normal time} - \text{Crash time}} = \frac{C_c - C_n}{T_n - T_c}.$$

It is also called the crash cost slope or crash rate of increase in the cost of performing the activity per unit reduction in time and is also known as the time-cost trade-off. It varies from activity to activity. After assessing the direct and indirect project costs, the total project cost, which is the sum of the direct and indirect costs, can be determined.

**Time-cost optimization algorithm/time-cost trade-off procedure**

The following are the steps involved in the project crashing.

Step 1 Considering normal times of all activities, identify the critical activities and find the critical path.

Step 2 Calculate the cost slope for different activities and rank the activities in ascending order of cost slope.

Step 3 Crash the activities on the critical path as per ranking of cost slopes of different activities; i.e. an activity having a lower cost slope would be crashed first to the maximum extent possible. (For the crashing of a lower cost slope, i.e. for the reduction of activity duration time, the direct cost of the project would be increased very slowly.)

Step 4 Due to the reduction of the critical path duration by crashing in Step 3, other paths also become critical; i.e. we get parallel critical paths. In such cases, the project duration can be reduced by crashing of activities simultaneously in the parallel critical paths.

Step 5 Repeat the process until all the critical activities are fully crashed or no further crashing is possible.

In the case of indirect cost, the process of crashing is repeated until the total cost is minimum, beyond which it may increase. The minimum cost is called the optimum project cost and the corresponding time, the optimum project time.

*Example 4* The following table shows activities, their normal time, cost and crash time and cost for a project:

| Activity | Normal time (days) | Cost (Rs.) | Crash time (days) | Cost (Rs.) |
|---|---|---|---|---|
| (1, 2) | 6 | 1400 | 4 | 1900 |
| (1, 3) | 8 | 2000 | 5 | 2800 |
| (2, 3) | 4 | 1100 | 2 | 1500 |
| (2, 4) | 3 | 800 | 2 | 1400 |
| (3, 4) | Dummy | – | – | – |
| (3, 5) | 6 | 900 | 3 | 1600 |
| (4, 6) | 10 | 2500 | 6 | 3500 |
| (5, 6) | 3 | 500 | 2 | 800 |

The indirect cost for the project is Rs. 300 per day.

(i) Draw the network of the project.
(ii) What are the normal duration and associated cost of the project?
(iii) What will be the least project duration and the corresponding cost?
(iv) Find the optimum duration and minimum project cost.

**Solution**

(i) The network of the project is shown in Fig. 14.15.

(ii) Using the normal time duration of each activity, the earliest and latest occurrence times at various nodes are computed and displayed in Fig. 14.15 of the network.

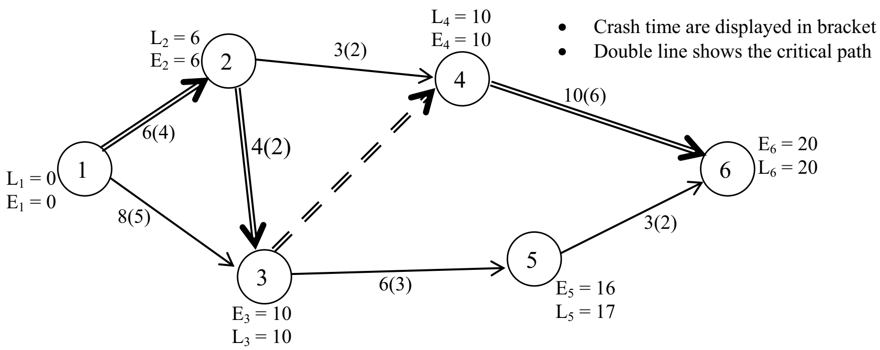



**Fig. 14.15** Project network with time duration 20 days

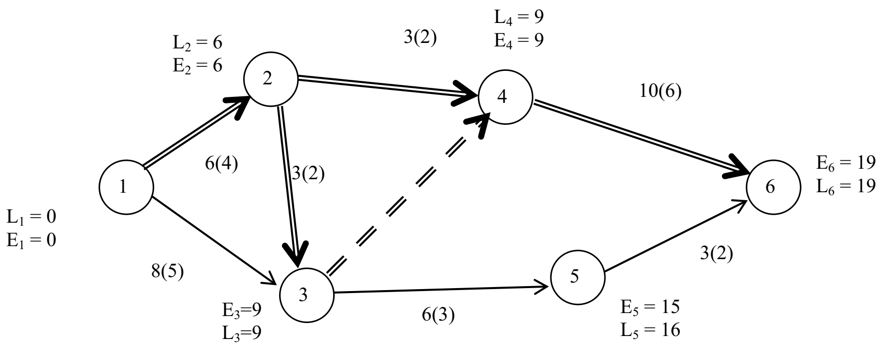



**Fig. 14.16** Project network with time duration 19 days

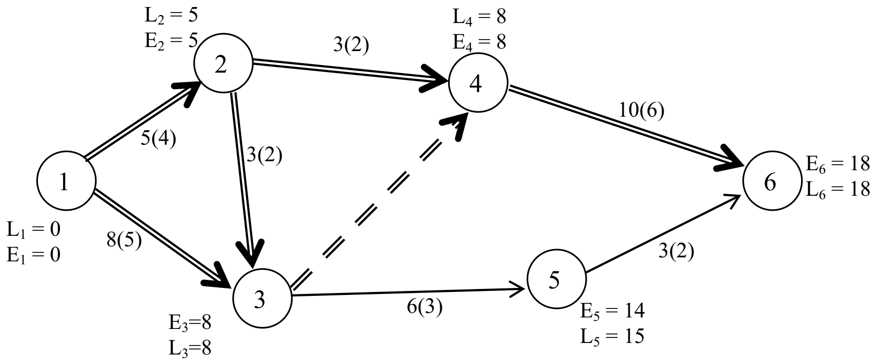



**Fig. 14.17** Project network with time duration 18 days

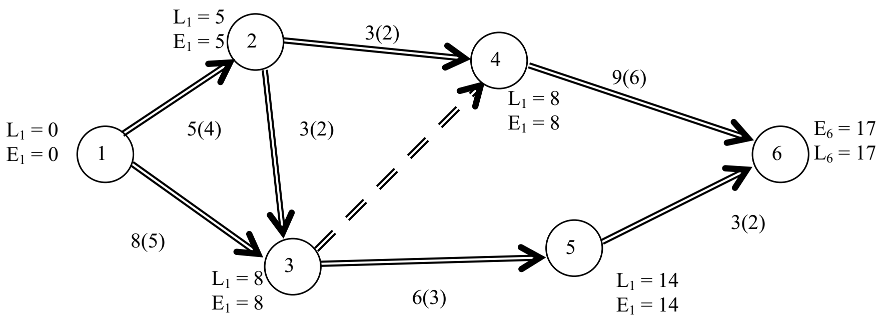



**Fig. 14.18** Project network with time duration 17 days

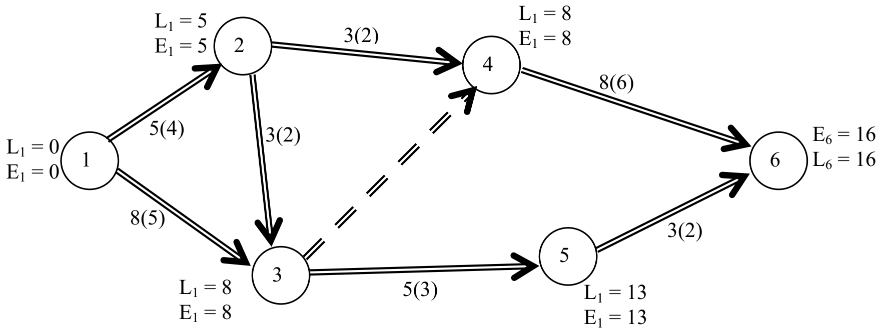



**Fig. 14.19** Project network with time duration 16 days

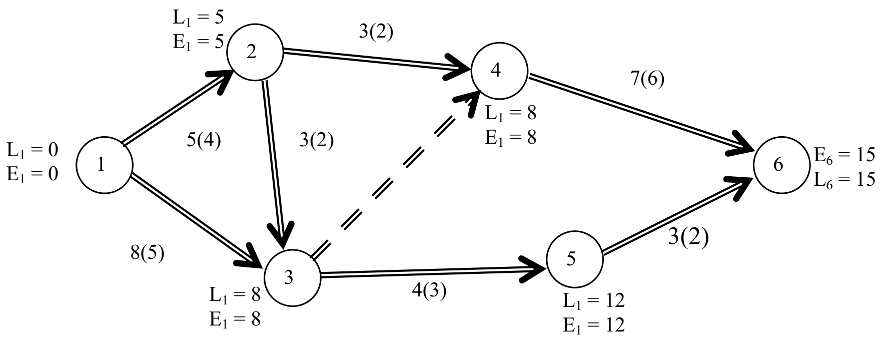



**Fig. 14.20** Project network with time duration 15 days

From the network, it is seen that the $L$ values and $E$ values at nodes 1, 2, 3, 4, 6 are same. This means that the critical path is 1–2–3–4–6 and the normal duration of the project is 20 days.

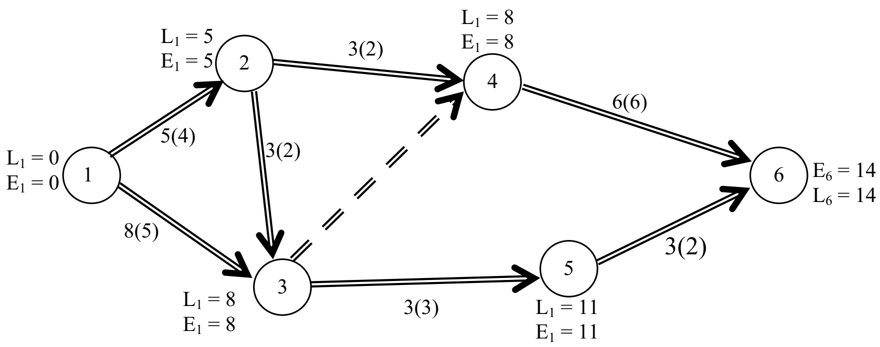



**Fig. 14.21** Project network with time duration 14 days

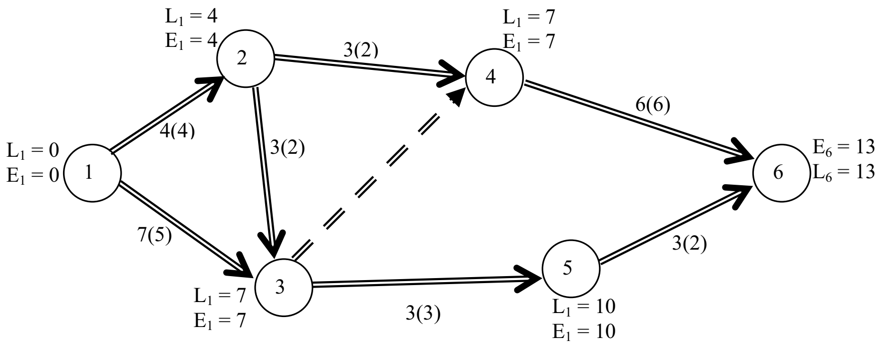



**Fig. 14.22** Project network with time duration 13 days

The associated cost of the project are as follows:

$$
\begin{aligned}
&= \text{Direct normal cost} + \text{indirect cost for 20 days}\\
&= \text{Rs. } [(1400 + 2000 + 1100 + 800 + 900 + 2500 + 500) + 20 \times 300]\\
&= \text{Rs. } [9200 + 6000] = \text{Rs. } 15{,}200.
\end{aligned}
$$

(iii) The cost slopes of different activities are computed by using the formula

$$\text{Cost slope} = \frac{\text{Crash cost} - \text{Normal cost}}{\text{Normal time} - \text{Crash time}}$$

and these are shown in the following table:

| Activity | (1, 2) | (1, 3) | (2, 3) | (2, 4) | (3, 5) | (4, 6) | (5, 6) |
|---|---|---|---|---|---|---|---|
| Slope | 250 | 267 | 200 | 600 | 233 | 250 | 300 |

**Table 14.4** Computation of total cost with duration due to crashing for Example 4

| Critical path(s) | See figure (before crashing) | Activities crashed (time) | Project length after crashing (days) | Normal direct cost (Rs.) | Crashing cost (Rs.) | Indirect cost (Rs. 300/day) | Total cost (Rs.) |
|---|---|---|---|---|---|---|---|
| | | | | (*A*) | (*B*) | (*C*) | (*A* + *B* + *C*) |
| 1–2–3–4–6 | Figure 14.15 | _ | 20 | 9200 | _ | 300 × 20 | 15,200 |
| 1–2–3–4–6 | Figure 14.15 | (2, 3) (1) | 19 | 9200 | 200 × 1 = 200 | 300 × 19 | 15,100 |
| 1–2–3–4–6, 1–2–4–6 | Figure 14.16 | (1, 2) (1) | 18 | 9200 | 200 + 250 × 1 = 450 | 300 × 18 | 15,050 |
| 1–2–3–4–6, 1–2–4–6 1–3–4–6 | Figure 14.17 | (4, 6) (1) | 17 | 9200 | 450 + 250 × 1 = 700 | 300 × 17 | 15,000 |
| 1–2–3–4–6, 1–2–4–6, 1–3–4–6, 1–3–5–6, 1–2–3–5–6 | Figure 14.18 | (3, 5) (1) (4, 6) (1) | 16 | 9200 | 700 + 233 × 1 + 250 × 1 = 1183 | 300 × 16 | 15,183 |
| 1–2–3–4–6, 1–2–4–6, 1–3–4–6, 1–3–5–6, 1–2–3–5–6 | Figure 14.19 | (3, 5)(1) (4, 6)(1) | 15 | 9200 | 1183 + 233 × 1 + 250 × 1 = 1666 | 300 × 15 | 15,366 |
| 1–2–3–4–6, 1–2–4–6, 1–3–4–6, 1–3–5–6, 1–2–3–5–6, | Figure 14.20 | (3, 5) (1) (4, 6) (1) | 14 | 9200 | 1666 + 233 × 1 + 250 × 1 = 2149 | 300 × 14 | 15,549 |
| 1–2–3–4–6, 1–2–4–6, 1–3–4–6, 1–3–5–6, 1–2–3–5–6 | Figure 14.21 | (1, 2) (1) (1, 3) (1) | 13 | 9200 | 2149 + 250 × 1 + 267 × 1 = 2666 | 300 × 13 | 15,766 |
| 1–2–3–4–6, 1–2–4–6, 1–3–4–6, 1–3–5–6, 1–2–3–5–6 | Figure 14.22 | (2,3) (1) (2,4) (1) (1,3) (1) | 12 | 9200 | 2666 + 200 × 1 + 600 × 1 + 267 × 1 = 3733 | 300 × 12 | 16,533 |

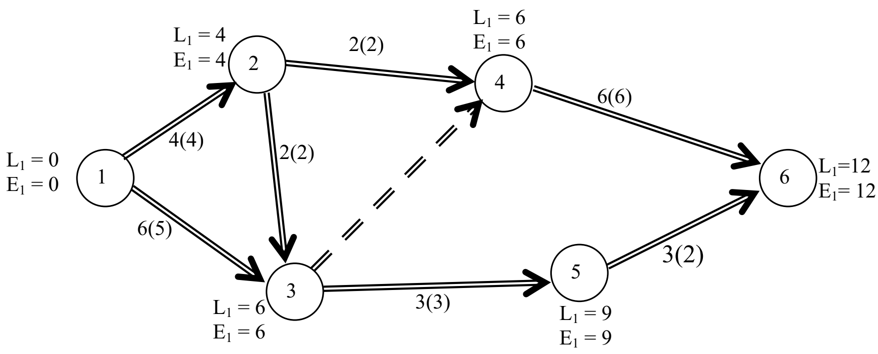



**Fig. 14.23** Project network with time duration 12 days

(iv) Now we construct the table for finding the optimal cost with duration and minimum project duration with cost (cf. Table 14.4).

From Fig. 14.23, it is seen that no further crashing is possible beyond 12 days. Hence, the least project duration is 12 days, and the corresponding cost of the project will be Rs. 16,533.00.

As the minimum cost occurs for a 17-day schedule, the optimum duration of the project is 17 days, and the minimum project cost is Rs. 15,000.00.

*Example 5* The following table gives data on normal time & cost and crashed time & cost for a project:

| Activity | Time (days) | | Cost (Rs.) | |
|---|---|---|---|---|
| | Normal | Crash | Normal | Crash |
| (1,2) | 9 | 4 | 1300 | 2400 |
| (1,3) | 15 | 13 | 1000 | 1380 |
| (2,3) | 7 | 4 | 700 | 1540 |
| (2,4) | 7 | 3 | 1200 | 1920 |
| (2,5) | 12 | 6 | 1700 | 2240 |
| (3,6) | 12 | 11 | 600 | 700 |
| (4,5) | 6 | 2 | 1000 | 1600 |
| (5,6) | 9 | 6 | 900 | 1200 |

The indirect cost for the project is Rs. 400 per day.

(i) Draw the network of the project.
(ii) What are the normal duration and associated cost of the project?
(iii) What are the optimum project duration and corresponding cost?
(iv) Find the optimum cost and corresponding duration.

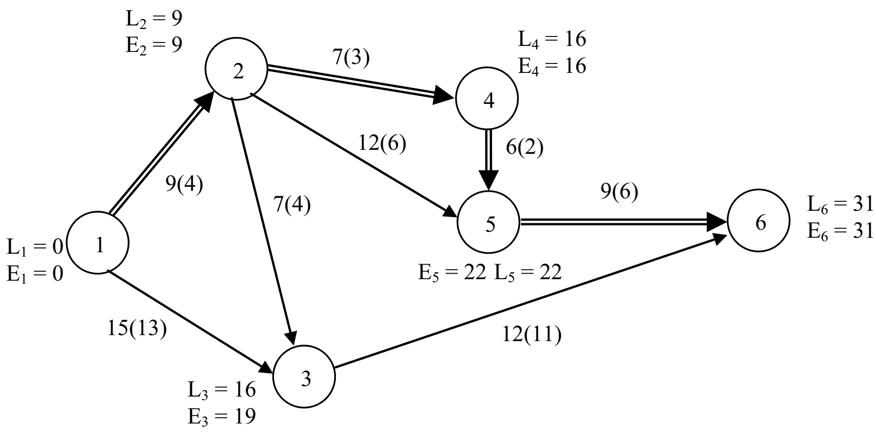
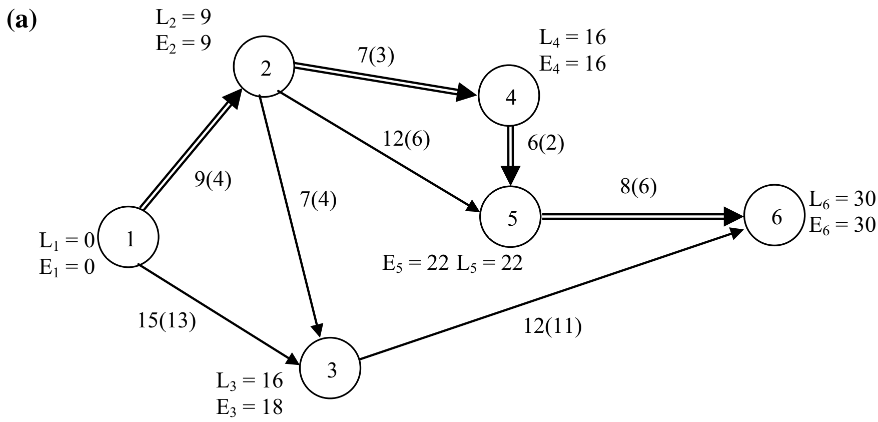



**Fig. 14.24** Project network with time duration 31 days



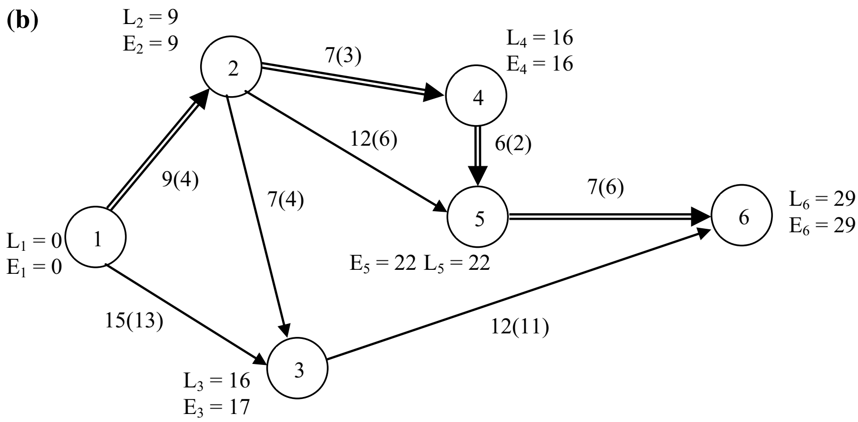




**Fig. 14.25** **a** Project network with time duration 30 days. **b** Project network with time duration 29 days

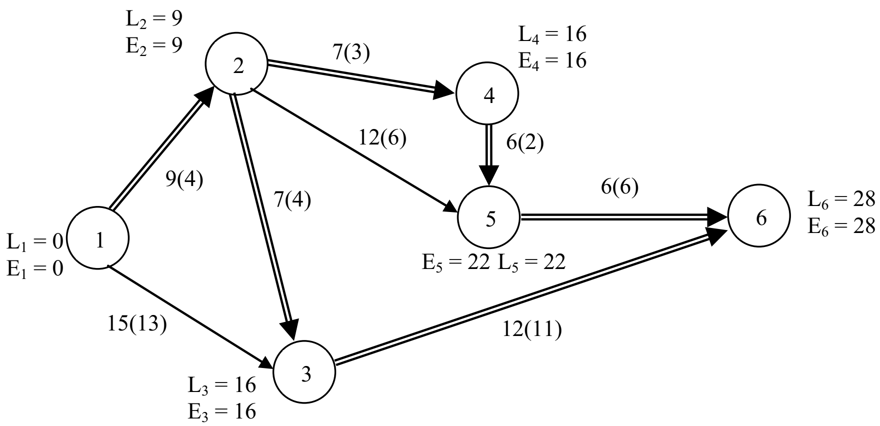



**Fig. 14.26** Project network with time duration 28 days

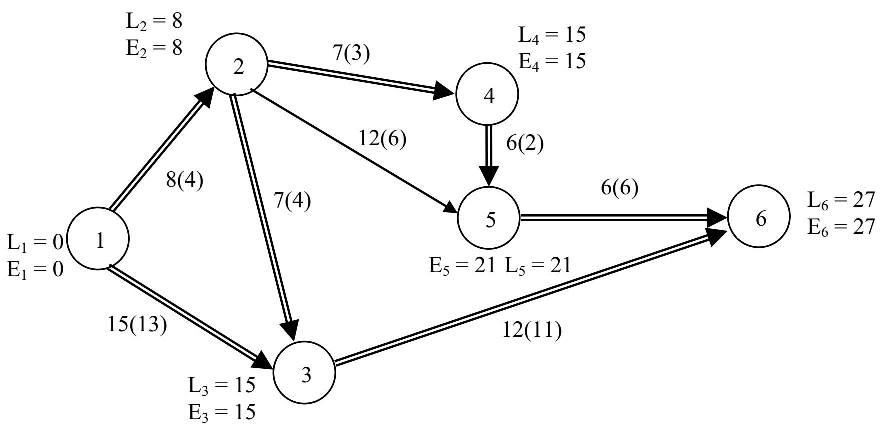



**Fig. 14.27** Project network with time duration 27 days

## Solution

(i) The network diagram of the project is shown in Fig. 14.24.

(ii) The network of the project has been drawn using the normal time duration for each activity. Then for each event, the earliest and latest occurrence times for each event have been computed and displayed in the network. The normal and crash time durations are also shown in the network.

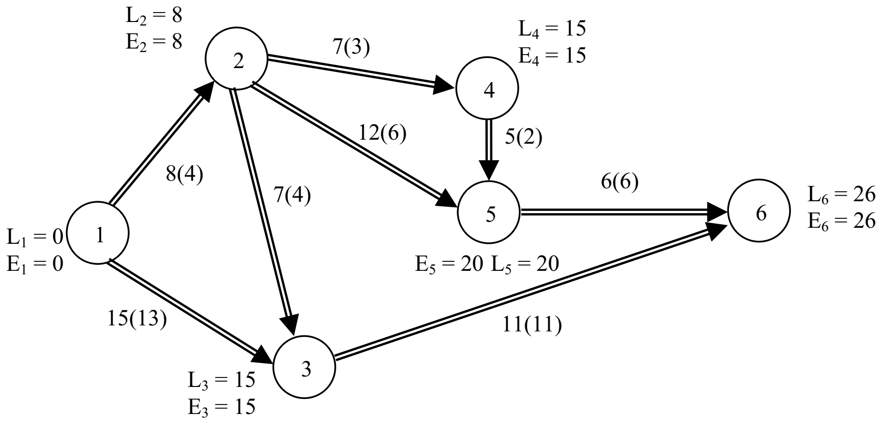



**Fig. 14.28** Project network with time duration 26 days

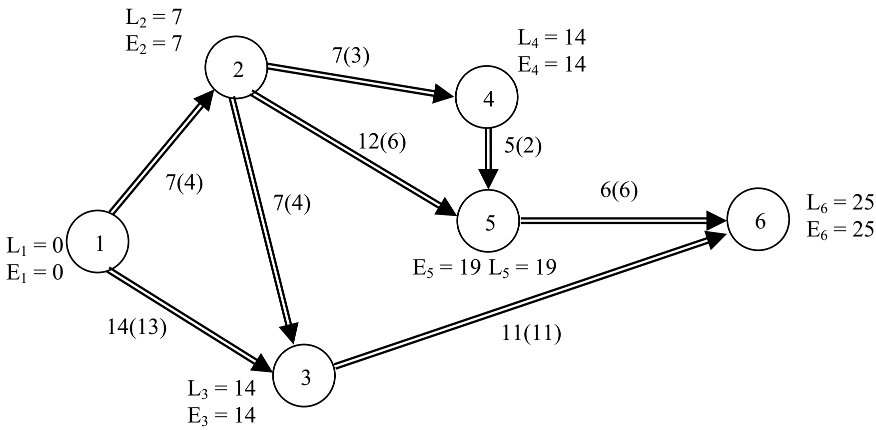



**Fig. 14.29** Project network with time duration 25 days

From the network, it is seen that

$$L_1 = E_1 = 0, L_2 = E_2 = 9, L_4 = E_4 = 16, L_5 = E_5 = 22, L_6 = E_6 = 31.$$

Hence, the critical path is 1–2–4–5–6.

As $E_6 = L_6 = 31$, the normal duration of the project is 31 days. The associated cost of the project

$$\begin{aligned}
&= \text{Direct normal cost} + \text{indirect cost for 31 days} \\
&= \text{Rs. } [(1300 + 1000 + 700 + 1200 + 1700 + 600 + 1000 + 900) + 31 \times 400] \\
&= \text{Rs. } 20{,}800.
\end{aligned}$$

(iii) The cost slopes of different activities are computed by using the formula

Cost slope = (Crash Cost − Normal cost)/(Normal time − Crash time), and these values are shown in the following table:

| Activity | (1,2) | (1,3) | (2,3) | (2,4) | (2,5) | (3,6) | (4,5) | (5,6) |
|---|---|---|---|---|---|---|---|---|
| Cost slope | 220 | 190 | 280 | 180 | 90 | 100 | 150 | 100 |

Now we construct the table for finding the optimal cost with duration and minimum duration with cost (cf. Table 14.5).

**Table 14.5** Computation of total cost with duration due to crashing for Example 5

| Critical path(s) | See figure (before crashing) | Activities crashed (time) | Project length after crashing (days) | Normal direct cost (Rs.) (*A*) | Crashing cost (Rs.) (*B*) | Indirect cost (Rs. 400/day) (*C*) | Total cost (Rs.) (*A* + *B* + *C*) |
|---|---|---|---|---|---|---|---|
| 1–2–4–5–6 | Figure 14.24 | | 31 | 8400 | | 400 × 31 | 20,800 |
| 1–2–4–5–6 | Figure 14.24 | (5,6) (1) | 30 | 8400 | 100 × 1 = 100 | 400 × 30 | 20,500 |
| 1–2–4–5–6 | Figure 14.25a | (5,6) (1) | 29 | 8400 | 100 + 100 × 1 = 200 | 40ion ion 0 × 29 | 20,200 |
| 1–2–4–5–6<br>1–2–3–6 | Figure 14.25b | (5,6) (1) | 28 | 8400 | 200 + 100 × 1 = 300 | 400 × 28 | 19,900 |
| 1–2–4–5–6<br>1–2–3–6 | Figure 14.26 | (1,2) (1) | 27 | 8400 | 300 + 220 × 1 = 520 | 400 × 27 | 19,720 |
| 1–2–4–5–6<br>1–2–3–6<br>1–3–6<br>1–2–5–6 | Figure 14.27 | (4,5) (1)<br>(3,6) (1) | 26 | 8400 | 520 + 150 × 1 + 100 × 1 = 770 | 400 × 26 | 19,570 |
| 1–2–4–5–6<br>1–2–3–6<br>1–3–6<br>1–2–5–6 | Figure 14.28 | (1,2) (1)<br>(1,3) (1) | 25 | 8400 | 770 + 220 × 1 + 190 × 1 = 1180 | 400 × 25 | 19,580 |
| Do | Figure 14.29 | (1,2) (1)<br>(1,3) (1) | 24 | 8400 | 1180 + 220 × 1 + 190 × 1 = 1590 | 400 × 24 | 19,590 |

From Fig. 14.30, it is seen that no further crashing is possible beyond 24 days. Hence, the optimum project duration, i.e. the least project duration, is 24 days and the corresponding cost of the project will be Rs. 19,590.00.

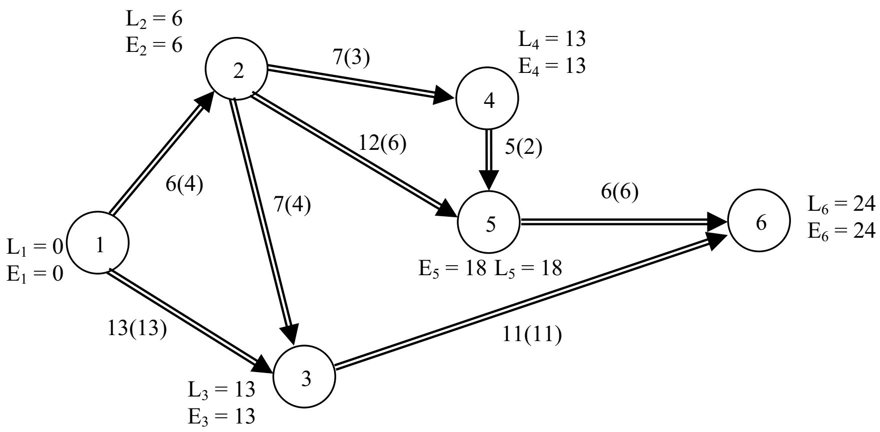



**Fig. 14.30** Project network with time duration 24 days

(iv) From Table 14.5, it is observed that the optimum cost is Rs. 19,570.00 and the corresponding project duration is 26 days.

## 14.13 Exercises

1. A project consists of a series of tasks labelled $A$, $B$, …, $H$, I with the following relationships ($W < X$, $Y$ means $X$ and $Y$ cannot start until $W$ is completed; $X$, $Y < W$ means $W$ cannot start until both $X$ and $Y$ are completed). With this notation, construct the network diagram having the following constraints:

$$A < D, E; B, D < F; C < G; G < H; F, G < I.$$

Find also the minimum time of completion of the project, when the time (in days) of completion of each task is as follows:

| Task | $A$ | $B$ | $C$ | $D$ | $E$ | $F$ | $G$ | $H$ | $I$ |
|---|---|---|---|---|---|---|---|---|---|
| Time | 23 | 8 | 20 | 16 | 24 | 18 | 19 | 4 | 10 |

2. The following are the details of the estimated times of activities of a certain project:

| Activity | *A* | *B* | *C* | *D* | *E* | *F* |
|---|---|---|---|---|---|---|
| Immediate predecessor | _ | *A* | *A* | *B, C* | _ | *E* |
| Estimated time (weeks) | 2 | 3 | 4 | 6 | 2 | 8 |

(a) Find the critical path and the expected time of the project.
(b) Calculate the earliest start time and earliest finish time for each activity.
(c) Calculate the float for each activity.

3. A project consists of eight activities with the following relevant information:

| Activity | Immediate predecessor | Estimated duration (days) | | |
|---|---|---|---|---|
| | | Optimistic | Most likely | Pessimistic |
| *A* | _ | 1 | 1 | 7 |
| *B* | _ | 1 | 4 | 7 |
| *C* | _ | 2 | 2 | 8 |
| *D* | *A* | 1 | 1 | 1 |
| *E* | *B* | 2 | 5 | 14 |
| *F* | *C* | 2 | 5 | 8 |
| *G* | *D, E* | 3 | 6 | 15 |
| *H* | *F, G* | 1 | 2 | 3 |

(ii) Draw the network and find the expected project completion time.
(iii) What duration will have 95% confidence for project completion?
(iv) If the average duration for activity F increases to 14 days, what will be its effect on the expected project completion time which will have 95% confidence?

4. Draw the network for the following project and compute the earliest and latest times for each event and also find the critical path:

| Activity | (1, 2) | (1, 3) | (2, 4) | (3, 4) | (4, 5) | (4,6) | (5, 7) | (6,7) | (7, 8) |
|---|---|---|---|---|---|---|---|---|---|
| Immediate predecessor | _ | _ | (1, 2) | (1, 3) | (2, 4) | (2,4), (3,4) | (4, 5) | (4, 6) | (6, 7), (5, 7) |
| Time (days) | 5 | 4 | 6 | 2 | 1 | 7 | 8 | 4 | 3 |

5. A small project consists of seven activities, the details of which are given below:

| Activity | Time estimates | | | Predecessor |
|---|---|---|---|---|
| | $t_o$ | $t_m$ | $t_p$ | |
| *A* | 3 | 6 | 9 | None |
| *B* | 2 | 5 | 8 | None |
| *C* | 2 | 4 | 6 | *A* |
| *D* | 2 | 3 | 10 | *B* |
| *E* | 1 | 3 | 11 | *B* |
| *F* | 4 | 6 | 8 | *C, D* |
| *G* | 1 | 5 | 15 | E |

Find the critical path. What is the probability that the project will be completed by 18 weeks?

6. The following table gives data on normal time-cost and crash time-cost for a project:

| Activity | Normal | | Crash | |
|---|---|---|---|---|
| | Time (days) | Cost (Rs.) | Time (days) | Cost (Rs.) |
| (1, 2) | 6 | 650 | 4 | 1000 |
| (1, 3) | 4 | 600 | 2 | 2000 |
| (2, 4) | 5 | 500 | 3 | 1500 |
| (2, 5) | 3 | 450 | 1 | 650 |
| (3, 4) | 6 | 900 | 4 | 2000 |
| (4, 6) | 8 | 800 | 4 | 3000 |
| (5, 6) | 4 | 400 | 2 | 1000 |
| (6, 7) | 3 | 450 | 2 | 800 |

The indirect cost per day is Rs. 100.

(i) Draw the network and identify the critical path.
(ii) What are the normal project duration and associated cost?
(iii) Crash the relevant activities systematically and determine the optimum project completion time and cost.

7. The following table gives the activities in a construction project and other relevant information:

| Activity | Immediate predecessor | Time (days) | | Direct cost (Rs.) | |
|---|---|---|---|---|---|
| | | Normal | Crash | Normal | Crash |
| *A* | – | 4 | 3 | 60 | 90 |
| *B* | – | 6 | 4 | 150 | 250 |
| *C* | – | 2 | 1 | 38 | 60 |
| *D* | *A* | 5 | 3 | 150 | 250 |
| *E* | *C* | 2 | 2 | 100 | 100 |
| *F* | *A* | 7 | 5 | 115 | 175 |
| *G* | *B, D, E* | 4 | 2 | 100 | 240 |

Indirect costs vary as follows:

| Days | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 |
|---|---|---|---|---|---|---|---|---|---|---|
| Cost (Rs.) | 600 | 500 | 400 | 250 | 175 | 100 | 75 | 50 | 35 | 25 |

(i) Draw the network of the project.
(ii) Determine the project duration which will result in minimum total project cost.

# Chapter 15
# Queueing Theory

## 15.1 Objective

The objective of this chapter is to discuss

- The different situations which generate queueing problems
- The various factors/components of a queueing system
- A variety of performance measures (operating characteristics) of a queueing system
- The probability distributions of arrivals and departures in a queueing system
- Poisson queueing models with finite and infinite capacity
- The single server infinite capacity queueing model with Poisson arrival and multiphase Erlang service
- The single server infinite capacity queueing model with Poisson arrival and general service
- A mixed queueing model with random arrival rate which follows the Poisson distribution and deterministic service rate
- Finding the optimum service level by minimizing the total cost of $M/M/1$ systems with finite and infinite capacity.

## 15.2 Basics of Queueing Theory

### *15.2.1 Introduction*

A group of customers waiting to receive a service including those receiving the service is known as a waiting line or queue. Queueing theory involves the mathematical study of queues or waiting lines. It analyses a production and servicing process system exhibiting random variability in market demand (arrival times) and

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_15

service times. This subject had its beginning in the research work of Danish engineer A.K. Erlang. In the year 1909, Erlang experimented with fluctuating demand in telephone traffic. After eight years he established a report addressing the delays in automatic dialing equipment. At the end of the Second World War, Erlang's earlier work was extended and applied to solve general problems and the business applications of waiting lines.

Queueing theory tries to answer questions involving the average number of customers in the queue, average number of customers in the system, mean waiting time in the queue, mean waiting time in the system, and so forth.

The objective of a queueing model is to determine how to provide the service to the customers so as to maintain an economic balance between the cost (time) of the service and the cost (time) associated with the wait for the service by manipulating certain variables such as number of servers, rate of service and order of service.

### *15.2.2 Reason Behind the Queue*

The formation of waiting lines (or queues) is a common phenomenon which occurs whenever the current demand for a service exceeds the current capacity to provide that service; i.e. a customer does not get the service immediately but must wait, or the service facilities stand idle or the total number of service facilities exceeds the number of customers requiring service. Waiting lines or queues are a common occurrence both in our daily life and in various business and industrial situations.

Most queueing problems involve the question of finding the level of services that a firm should provide. Table 15.1 shows examples of queueing services.

For example:

(i) Supermarkets must decide how many cash counters should be opened.
(ii) Gasoline stations must decide how many pumps should be opened and how many attendants should be on duty.
(iii) Manufacturing plants must determine the optimal number of machines to have running in each shift.
(iv) Banks must decide how many teller windows to keep open to serve customers during certain hours of the day.

The person waiting in a queue or receiving the service is called the customer, and the person by whom the customer is serviced is called a server. Queueing problems arise because the cost of service may not be cheap enough to ensure sufficient service facilities so that no customer has to wait. Customers arrive at a counter according to a certain probabilistic law (Poisson's input, Erlang input, etc.). On the other hand, the customers will be served by one or more servers following a certain principle (FCFS, i.e. first come, first served, random service, etc.). By providing additional service facilities, the customer's waiting time or the queue can be reduced; conversely, by decreasing the number of service facilities, a long queue

**Table 15.1** Examples of queueing services

| Situation | Customers/ arrivals | Service facility or servers | Service process |
|---|---|---|---|
| Flow of automobile traffic through a road network | Automobiles | Road network | Automatic signalling |
| Ships entering a port | Ships | Docks | Unloading the items |
| Telephone exchange | Telephone calls | Switchboard operators or electric switching operators | Computer |
| Number of runways at airports | Aeroplanes | Runways | Landing or flying the planes |
| Bank | Customer | Bank clerks, check drawn | Money deposited or withdrawn |
| Sale of railway tickets | Railway passengers | Booking clerks | Selling tickets |
| Maintenance and repair of machines | Machine breakdown | Mechanics | Repairing |
| Department store | Customer | Checking clerks and bag packers | Bill payment |
| Job interviewing | Applicant | Interviewee | Selection of the applicant |

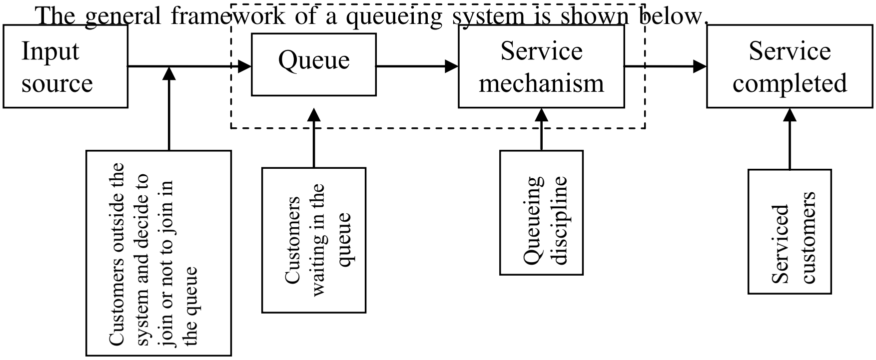
will be formed. The service times are random variables governed by a given probabilistic law. After being served, a customer leaves the queue. The objective of the queueing model is to determine how to provide the service to the customers so as to minimize the total cost of service and the waiting time of customers by manipulating certain factors (variables) such as number of servers, rate of service and order of service.

### *15.2.3 Characteristics of Queueing Systems*

Any queueing system can be characterized by the following set of characteristics:

- Input source
- Service mechanism
- Service discipline
- Capacity of the system.

The general framework of a queueing system is shown below.




**Input source**

The input source is a device or group of devices that provides customers at the service facility for the service. An input source is characterized by the following factors:

- Input size
- Arrival pattern
- Behaviour of the arrivals.

***Input size***

If the total number of customers requiring service is only a few, then the input size is said to be finite. But if the potential customers requiring service are sufficiently large in number, then the input source is considered to be infinite. Also, if the customers arrive at the service facility in batches of fixed size or variable size instead of one at a time, then the input is said to occur in bulk or batches. If the service is not available, they may be required to join the queue.

***Arrival pattern***

If the pattern by which customers arrive at the service system or time between successive arrivals (inter-arrival time) is uncertain, then the arrival pattern is measured by either mean arrival rate or inter-arrival time. These measures are characterized in the form of the probability distribution associated with these random processes. The most common stochastic queueing model assumes that the arrival rate follows a Poisson distribution or, equivalently, that the inter-arrival time follows an exponential distribution.

***Customer behaviour***

The customers generally behave in different ways as follows:

(a) *Balking*: A customer may leave the queue because the queue is too long and he/she has no time to wait or there is not sufficient waiting space.

(b) *Reneging*: A customer may become impatient after waiting in the queue for some time and leave the queue.
(c) *Priorities*: In certain applications some customers are served before others regardless of their order of arrival. These customers have priority over others.
(d) Jockeying: Customers may change position from one queue to another, hoping to receive the service more quickly.

**Service mechanism**

The service mechanism is concerned with the manner in which customers are served. It can be characterized by observing different factors:

(i) *Availability of service facility*: Service may be available only at certain times but not always.
(ii) *Capacity of service facility*: The capacity is measured in terms of customers who can be served simultaneously.

A queueing system may have one or more parallel service channels or a sequence of channels in series, as shown in Figs. 15.1, 15.2, 15.3 and 15.4.

(iii) *Service time or duration*:

Service time or duration may be constant or a random variable. In general, service time is not constant, since it is a common practice to use overtime or extra effort when an excessive queue condition is present.

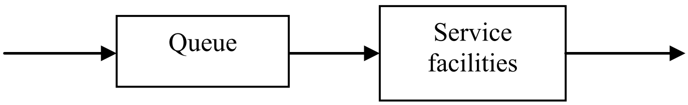
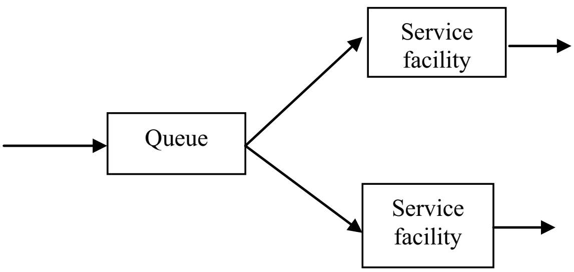



**Fig. 15.1** Single channel single phase queueing system


**Fig. 15.2** Single channel multiple phase queueing system

**Fig. 15.3** Multiple channel single phase queueing system



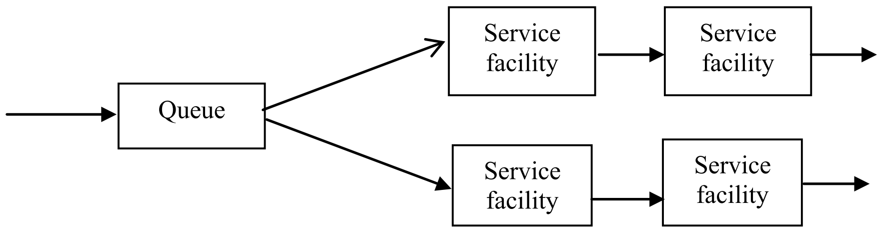



**Fig. 15.4** Multiple channel multiple phase queueing system

A variety of probability distributions can be used to characterize the service pattern. The service pattern is assumed to be independent with respect to the arrival of customers. In the most common stochastic queueing model, it is assumed that service time follows an exponential distribution or, equivalently, that service rate follows a Poisson distribution.

***Service discipline***

The service discipline refers to the order or manner in which customers in the queue will be served. The most common disciplines are as follows:

(i) ***First come, first served (FCFS)***: According to this discipline the customers are served in the order of their arrival. It is also known as FIFO (first input first output). This service discipline may be seen at a cinema ticket window or a railway ticket window, etc.
(ii) ***Last come, first served (LCFS)***: This discipline may be seen in a big god-own, where the units (items) which come last are taken out (served) first.
(iii) ***Serving in random order (SIRO)***: In this case, the arrivals are serviced randomly irrespective of their arrival in the system.
(iv) ***Service on some priority procedure***: Some customers are served before the others without considering their order of arrival; i.e. some customers are served on a priority basis. There are two types of priority services.

(a) *Preemptive priority*
In this case customers with the highest priority are allowed to enter the service immediately after entering the system, even if a customer with lower priority is already in service; i.e. lower priority customer service is interrupted to start service for a special customer. This interrupted service is resumed again after the highest priority customer is served.

(b) *Non-preemptive priority*
In this case the highest priority customer goes ahead of the queue, but service is started immediately on completion of the current customer service.

***Capacity of the system***

In certain cases a queueing system is unable to accommodate more than the required number of customers at a time. No further customers are allowed to enter until the space becomes available to accommodate new customers. This type of situation is referred to as a finite source queue.

### *15.2.4 Important Definitions*

**Queue length**: Queue length is defined by the number of persons (customers) waiting in the queue at any time.
**Average length of queue**: Average length of queue is defined by the number of customers in the queue per unit time.
**Waiting time**: The time up to which a unit has to wait in the queue before it is taken into the service.
**Servicing time**: The time taken for servicing of a unit.
**Busy period**: The busy period of a server is the time during which the server remains busy in servicing. Thus, it is the time between the starting of service of the first unit to the end of service of the last unit in the queue.
**Idle period**: When all the customers in the queue are served, the idle period of the server begins, and it continues up to the time of arrival of a customer. Thus, the idle period of a server is the time during which he remains free because there is no customer present in the system.
**Mean arrival rate**: The mean arrival rate in a queue is defined as the expected number of arrivals occurring in a time interval of unit length.
**Mean servicing rate**: The mean servicing rate for a particular servicing station is defined as the expected number of services completed in a time interval of unit length, given that the servicing is going on throughout the entire time unit.
**Traffic intensity**: For a simple queue, the traffic intensity is the ratio of the mean arrival and the mean servicing rate,

$$\text{i.e., Traffic intensity} = \frac{\text{Mean arrival rate}}{\text{Mean servicing rate}}$$

### *15.2.5 The State of the System*

The state of a system involves the study of a system's behaviour over time. The states of a system may be classified as follows:

(i) **Transient state**: A system is said to be in transient state when its operating characteristics are dependent on time. Thus, a queueing system is in transient state when the probability distributions of arrivals, waiting time and servicing time of the customers are dependent on time. This state occurs at the beginning of the operation of the system.

(ii) **Steady state**: A system is said to be in steady state when its operating characteristics become independent of time. Thus, a queueing system is in steady state when the probability distributions of arrivals, waiting time and servicing time of the customers are independent of time.

Let $p_n(t)$ denote the probability that there are $n$ units in the system at time $t$. Then if the probability $p_n(t)$ remains the same as time passes, the system acquires steady state. Mathematically, in steady state,

$$\lim_{t\to\infty} p_n(t) = p_n \text{ (independent of time)},$$

which implies $\lim_{t\to\infty} \frac{\mathrm{d}}{\mathrm{d}t} p_n(t) = \frac{\mathrm{d}p_n}{\mathrm{d}t} = 0.$

(iii) **Explosive state**: If the servicing rate is less than the arrival rate, the length of the queue will go on increasing with time and will tend to infinity as $t \to \propto$. The system is then said to be in explosive state.

### *15.2.6 Probability Distribution in a Queueing System*

In the queueing system, it is assumed that the customer's arrival is random and follows a Poisson distribution or, equivalently, that the inter-arrival times obey an exponential distribution. In most cases, service times are also assumed to be exponentially distributed. The basic assumptions (axioms) are as follows.

**Axiom 1** The probability that exactly one arrival will occur during the time interval $(t,\ t + \Delta t)$ is equal to $\lambda\Delta t + \Delta(t)$, i.e. $p_1(\Delta t)= \lambda\Delta t + O(\Delta t)$,
where $\lambda$ is a constant and independent of the total number of arrivals up to the time $t$, $\Delta t$ is a small time interval and $O(\Delta t)$ denotes a quantity which is of smaller order of magnitude than $\Delta t$ such that $\lim_{\Delta t\to o}\left\{\frac{O(\Delta t)}{\Delta t}\right\} = 0.$

**Axiom 2** The probability of more than one arrival during the time interval $(t,\ t + \Delta t)$ is negligible and is denoted by $O(\Delta t)$,

$$\text{i.e. } p_0(\Delta t) + p_1(\Delta t) + O(\Delta t) = 1.$$

**Axiom 3** The number of arrivals in non-overlapping intervals is statistically independent.

***Distribution of arrival (pure birth process)***

A process in which only arrivals are counted and no departure takes place is called a pure birth process. Stated in terms of queueing, a birth-death process usually arises by birth or arrival in the system and decreases by death or departure of serviced customers from the system.

***Property***

If the arrivals are completely random, then the probability distribution of the number of arrivals in a fixed time interval follows a Poisson distribution.

*Proof* Let there be $n$ units in the system at time $t$. In order to derive the arrival distribution in queues, we will consider the preceding three axioms.

Now we wish to determine the probability of $n$ arrivals at time $t$, denoted by $p_n(t)$. Clearly, $n$ will be an integer greater than or equal to zero.

**Case 1** When $n = 0$

If there is no unit in the system at time $t + \Delta t$, there will be no unit at time $t$ and no arrival during $\Delta t$.

Therefore, the probability of no unit in the system at time $t + \Delta t$ is given by

$$p_0(t + \Delta t) = p_0(t)(1 - \lambda\Delta t) + O(\Delta t). \tag{15.1}$$

**Case 2** When $n > 0$

In this case, there may be the following mutually exclusive ways of having $n$ units at time $t + \Delta t$:

(i) $n$ arrivals in the system at time $t$, and no arrival during the time $\Delta t$
(ii) $(n - 1)$ arrivals in the system at time $t$ and one arrival during the time $\Delta t$
(iii) $(n - 2)$ arrivals in the system at time $t$ and two arrivals during the time $\Delta t$, and so on.

Therefore, the probability of $n$ units in the system at time $t + \Delta t$ is given by

$$p_n(t + \Delta t) = p_n(t)(1 - \lambda\Delta t) + p_{n-1}(t)\lambda\Delta t + O(\Delta t). \tag{15.2}$$

Equations (15.1) and (15.2) can be written as

$$\frac{p_0(t+\Delta t)-p_0(t)}{\Delta t}=-\lambda p_0(t)+\frac{O(\Delta t)}{\Delta t},\quad n=0$$
$$\frac{p_n(t+\Delta t)-p_n(t)}{\Delta t}=-\lambda p_n(t)+\lambda p_{n-1}(t)+\frac{O(\Delta t)}{\Delta t},\quad n>0.$$

Taking the limit as $\Delta t \to 0$ on both sides of these equations, we have the following system of differential difference equations:

$$p_0'(t)=-\lambda p_0(t),\quad n=0 \tag{15.3}$$

$$p_n'(t)=-\lambda p_n(t)+\lambda p_{n-1}(t),\quad n>0. \tag{15.4}$$

**Solution of differential difference equations**

From (15.3), $\frac{p_0'(t)}{p_0(t)}=-\lambda$.
Integrating, we have $\log p_0(t)=-\lambda t+A$, where $A$ is an arbitrary constant.
When $t = 0$, $p_0(0)=1$, which implies $A = 0$.
Hence,

$$p_0(t)=\mathrm{e}^{-\lambda t}. \tag{15.5}$$

Now substituting $n = 1$ in (15.4), we have

$$p_1'(t)+\lambda p_1(t)=\lambda \mathrm{e}^{-\lambda t}\quad \left[\because\ p_0(t)=\mathrm{e}^{-\lambda t}\right].$$

This is a first-order linear differential equation.
Solving and using $p_1(t) = 0$ when $t = 0$, we have

$$p_1(t)=\lambda t\mathrm{e}^{-\lambda t}. \tag{15.6}$$

Now, substituting $n = 2$ in (15.4) and using the result of $p_1(t)$, we have

$$p_2'(t)+\lambda p_2(t)=\lambda\left(\lambda t\mathrm{e}^{-\lambda t}\right).$$

Solving this equation and using $p_2(0)=0$ when $t = 0$, we have

$$p_2(t)=\mathrm{e}^{-\lambda t}\frac{(\lambda t)^2}{2!}.$$

Similarly, for $n = 3$,

$$p_3(t) = \mathrm{e}^{-\lambda t} \frac{(\lambda t)^3}{3!}.$$

Let us assume that for $n = m$,

$$p_m(t) = \mathrm{e}^{-\lambda t} \frac{(\lambda t)^m}{\underline{|m}}.$$

Now, we shall prove that this result is true for $n = m + 1$ also.
For this purpose, substituting $n = m + 1$ in (15.4) and using the result $p_m(t) = \mathrm{e}^{-\lambda t} \frac{(\lambda t)^m}{m!}$, we have

$$p'_{m+1}(t) + \lambda p_{m+1}(t) = \lambda \frac{(\lambda t)^m \mathrm{e}^{-\lambda t}}{m!}.$$

On solving and using $p_{m+1}(0) = 0$ when $t$ =0, we have

$$p_{m+1}(t) = \frac{(\lambda t)^{m+1}}{(m+1)!} \mathrm{e}^{\lambda t}.$$

Hence, in general,

$$p_n(t) = \frac{(\lambda t)^n}{n!} \mathrm{e}^{\lambda t}, \quad n = 0, 1, 2, \ldots$$

which is the probability mass function of the Poisson distribution.
Therefore, the probability distribution of the number of arrivals in a fixed time interval follows a Poisson distribution.

***Distribution of inter-arrival times***

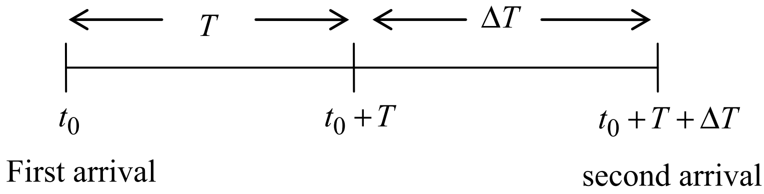
The inter-arrival time is defined as the time interval between two successive arrivals.

*Property of inter-arrival times*

If the arrival process in a queueing system follows a Poisson distribution, then the associated random variable defined as inter-arrival time follows an exponential distribution and vice versa.

*Proof* If $n$ is the number of arrivals at time $t$, then the probability of $n$ arrivals at time $t$ is given by

$$p_n(t) = \frac{(\lambda t)^n \mathrm{e}^{-\lambda t}}{n!}. \tag{15.7}$$




Let $t_0$ be the instant of an initial arrival.
Let $T$ be the random variable of inter-arrival time.
If there is no arrival during the intervals $(t_0,\, t_0 + T)$ and $(t_0 + T,\, t_0 + T + \Delta T)$, then $t_0 + T + \Delta T$ will be the instant of consecutive arrival.
From (15.7),

$$\begin{aligned} p_0(T) &= \text{probability of no arrival at time } T \\ &= \frac{(\lambda T)^0 \mathrm{e}^{-\lambda T}}{0!} = \mathrm{e}^{-\lambda T}, \end{aligned}$$

and

$$\begin{aligned} p_0(T + \Delta T) &= \text{probability of no arrival at time } \; T + \Delta T \\ &= \frac{\{\lambda(T + \Delta T)\}^0 \mathrm{e}^{-\lambda(T + \Delta T)}}{0!} = \mathrm{e}^{\lambda(T + \Delta T)} = \mathrm{e}^{-\lambda T} \cdot \mathrm{e}^{-\lambda \Delta T} \\ &= p_0(T)[1 - \lambda \Delta T + O(\Delta T)], \end{aligned}$$

where $O(\Delta T)$ contains the terms of second and higher powers of $\Delta T$.

$$\begin{aligned} \therefore \;\; & p_0(T + \Delta T) - p_0(T) = \lambda p_0(T)\Delta t + O(\Delta T) p_0(T) \\ \text{or} \;\; & \frac{p_0(T + \Delta T) - p_0(T)}{\Delta} = -\lambda p_0(T) + p_0(T)\frac{O(\Delta T)}{\Delta T}. \end{aligned}$$

Taking the limit as $\Delta T \to 0$, we have

$$\begin{aligned} & \frac{\mathrm{d}}{\mathrm{d}T}[p_0(T)] = -\lambda p_0(T). \\ \because & \frac{\mathrm{d}}{\mathrm{d}T}[p_0(T)] = -\lambda \mathrm{e}^{-\lambda T} \quad \left[\because p_0(T) = \mathrm{e}^{-\lambda T}\right]. \end{aligned} \tag{15.8}$$

According to probability theory, $F'(x) = f(x)$, where $F(x)$ is the distribution function and $f(x)$ is the probability density function. Hence, by a similar argument, we may write $\frac{\mathrm{d}}{\mathrm{d}T}[p_0(T)]$ as $f(T)$, where $p_0(T)$ is the probability distribution function for no arrival at time $T$ and $f(T)$ is the corresponding probability density function of $T$.

Again, since the probability density function is always non-negative, we disregard the negative sign from the right-hand side of (15.8).

Hence, $f(T) = \lambda e^{-\lambda T}$, which is the probability density function of the exponential distribution. That is, the random variable $T$ follows an exponential distribution.

The converse of the property may be proved in a similar way.

***Distribution of departure (pure death process)***

Sometimes a situation may arise where no additional customer joins the system while service is continued for those who are in the queue. At time $t$ let there be $N\,(\geq 1)$ customers in the system. It is clear that service will be provided at a rate $\mu$. The number of customers in the system at time $t \geq 0$ is equal to $N$ minus the total departure up to the time $t$. The distribution of the departure can be obtained with the help of the following basic axioms.

***Basic axioms***

(i) The probability of one departure during the time interval $(t, t + \Delta t)$ is equal to $\mu\Delta t + O(\Delta t)$.
(ii) The probability of more than one departure during the time interval $(t, t + \Delta t)$ is negligible and is denoted by $O(\Delta t)$.
(iii) The numbers of departures in non-overlapping intervals are statistically independent.

## *15.2.7 Classification of Queueing Models*

Generally, a queueing model may be specified in a symbolic form as $(a/b/c)$:$(d/e/f)$, where:

$a$ = Arrival distribution, i.e. type of arrival process
$b$ = Service time distribution, i.e. type of service process
$c$ = Number of servers
$d$ = Capacity of the system
$e$ = Service discipline
$f$ = Number of calling source capacity of input service.

Generally, the last symbol $f$ is not used. The standard notation for $a$ and $b$ is taken as $M$, which is called the Poisson (or Markovian) arrival or departure distribution or equivalently the exponential inter-arrival service time distribution.

$E_k$ = Erlangian or gamma inter-arrival for service time distribution with parameter $k$
$G$ = General service time distribution or departure distribution.

Thus, $(M/E_k/1)$:($\infty$/FIFO/$\infty$) defines a queueing system in which arrival follows the Poisson distribution; service time distributions are Erlangian; there is a single server; the capacity of the system is infinite, i.e. the system can hold an infinite number of customers; the service discipline is first input first output; and finally the source generating the arriving customers has an infinite capacity.

Generally queues with arrivals and departures start under a transient condition and gradually reach steady state after a sufficiently long time has elapsed, provided that the parameters of the system permit reaching steady state.

Now we define some parameters for a steady state queueing system:

$p_n$ = (steady state) Probability of $n$ customers in the system
$L_s$ = Expected number of customers in the system
$L_q$ = Expected number of customers in the queue
$W_s$ = Expected waiting time in the system
$W_q$ = Expected waiting time in the queue

## 15.3 Poisson Queueing Models

### *15.3.1 Introduction*

In this section, we shall discuss single as well as multiserver queueing models with Poisson arrival and departure processes. According to the capacity of the queueing system, we can classify these models as follows:

(i) Single server queueing model with infinite capacity of the system,

$$\text{i.e., } (M/M/1){:}(\infty/\text{FCFS}/\infty)$$

(ii) Single server queueing model with finite capacity of the system,

$$\text{i.e., } (M/M/1){:}(N/\text{FCFS}/\infty)$$

(iii) Multiserver queueing model with infinite capacity of the system,

$$\text{i.e., } (M/M/c){:}(\infty/\text{FCFS}/\infty)$$

(iv) Multiserver queueing model with finite capacity of the system,

$$\text{i.e.}(M/M/c){:}(\text{N}/\text{FCFS}/\infty).$$

We shall also discuss the repair problem of machine breakdowns and the power supply problem.

### 15.3.2 Model (M/M/1):(∞/FCFS/∞)

**Assumptions** In this model, we shall discuss a single server queueing system under the following assumptions:

(i) The arrival rate follows a Poisson distribution.
(ii) The servicing rate follows a Poisson distribution.
(iii) The queueing system has only a single service channel.
(iv) The service discipline is first come, first served (FCFS).
(v) The capacity of the system is infinite.

Let $\lambda$ and $\mu$ be the mean arrival rate and mean service rate of units respectively, and let $n$ be the number of customers in the system.

Hence, the probability of one arrival during time $\Delta t$ is $\lambda\Delta t + O(\Delta t)$ and the probability of one departure or service during $\Delta t$ is $\mu\Delta t + O(\Delta t)$.

If there is no customer in the system at time $t + \Delta t$, there may be the following mutually exclusive cases:

(i) No customer in the system at time $t$, no arrival in time $\Delta t$
(ii) One customer in the system at time $t$, no arrival in time $\Delta t$, one service in time $\Delta t$
(iii) No customer in the system at time $t$, one arrival in time $\Delta t$, one service in time $\Delta t$,

and so on.

Therefore, the probability of no customers in the system at time $t + \Delta t$ is given by

$$\begin{aligned} p_0(t+\Delta t) &= p_0(t)[1-\lambda\Delta t + O(\Delta t)] + p_1(t)[1-\lambda\Delta t + O(\Delta t)][\mu\Delta t + O(\Delta t)] + O(\Delta t) \\ &= p_0(t) - \lambda p_0(t)\Delta t + \mu p_1(t)\Delta t + O(\Delta t) \end{aligned}$$

$$\text{or,}\quad \frac{p_0(t+\Delta t) - p_0(t)}{\Delta t} = -\lambda p_0(t) + \mu p_1(t) + \frac{O(\Delta t)}{\Delta t}.$$

Taking the limit as $\Delta t \to 0$ on both sides in the preceding expression, we have

$$p_0'(t) = -\lambda p_0(t) + \mu p_1(t). \tag{15.9}$$

To have $n$ (>0) customers in the system at time $t + \Delta t$, there may be the following mutually exclusive cases:

(i) $n$ customers in the system at time $t$, no arrival in time $\Delta t$, no service in time $\Delta t$
(ii) $(n - 1)$ customers in the system at time $t$, one arrival in time $\Delta t$, no service in time $\Delta t$
(iii) $(n + 1)$ customers in the system at time $t$, no arrival in time $\Delta t$, one service in time $\Delta t$, and so on.

Therefore, the probability of $n$ customers in the system at time $(t + \Delta t)$ is given by

$$\begin{aligned} p_n(t+\Delta t) &= p_n(t)\{1-\lambda\Delta t+O(\Delta t)\}\{1-\mu\Delta t+O(\Delta t)\}+p_{n-1}(t) \\ &\quad [\lambda\Delta t+O(\Delta t)][1-\mu\Delta t+O(\Delta t)] \\ &\quad +p_{n+1}(t)[1-\lambda\Delta t+O(\Delta t)][\mu\Delta t+O(\Delta t)]+O(\Delta t) \text{ for } n>0 \\ &= p_n(t)+[-(\lambda+\mu)p_n(t)+\lambda p_{n-1}(t)+\mu p_{n+1}(t)]\Delta t+O(\Delta t). \end{aligned}$$

$$\therefore\; p_n(t+\Delta t)-p_n(t) = [\lambda p_{n-1}(t)-(\lambda+\mu)p_n(t)+\mu p_{n+1}(t)]\Delta t+O(\Delta t)$$

$$\text{or}\quad \frac{p_n(t+\Delta t)-p_n(t)}{\Delta t} = \lambda p_{n-1}(t)-(\lambda+\mu)p_n(t)+\mu p_{n+1}(t)+\frac{O(\Delta t)}{\Delta t} \text{ for } n>0.$$

Now taking the limit as $\Delta t \to 0$ of both sides in this equation, we have

$$p_n'(t) = \lambda p_{n-1}(t)-(\lambda+\mu)p_n(t)+\mu p_{n+1}(t), \quad n>0. \tag{15.10}$$

Under the steady state condition of the system, i.e. $\lim_{t\to\infty} p_n(t) = p_n$ and $\lim_{t\to\infty} p_n'(t) = 0$, Eqs. (15.9) and (15.10) reduce to

$$-\lambda p_0+\mu p_1 = 0 \tag{15.11}$$

$$\lambda p_{n-1}-(\lambda+\mu)p_n+\mu p_{n+1} = 0 \quad \text{for } n>0, \tag{15.12}$$

which are the steady state difference equations of the system.

***Solution of the Model***

From (15.11), $-\lambda p_0+\mu p_1 = 0$,
which implies

$$p_1 = \frac{\lambda}{\mu}p_0 = \rho\, p_0 \quad \left[\because \frac{\lambda}{\mu} = \rho\right].$$

Now setting $n = 1$ in (15.12), we have

$$\lambda p_0-(\lambda+\mu)p_1+\mu p_2 = 0$$

$$\text{or}\quad p_2 = -\frac{\lambda}{\mu}p_0+\left(\frac{\lambda}{\mu}+1\right)p_1$$

$$\text{or}\quad p_2 = -\rho p_0+(\rho+1)\rho p_0 = \rho^2 p_0.$$

Again setting $n = 2$ in (15.12), we have

$$\lambda p_1 - (\lambda + \mu)p_2 + \mu p_3 = 0$$

$$\text{or} \;\; p_3 = -\frac{\lambda}{\mu}p_1 + \left(\frac{\lambda}{\mu} + 1\right)p_2$$
$$= -\rho^2 p_0 + \rho^3 p_0 + \rho p_0 = \rho^3 p_0.$$

Hence,

$$p_3 = \rho^3 p_0. \tag{15.13}$$

Proceeding in this way, we will have

$$p_n = \rho^n p_0.$$

The value of $p_0$ is obtained from the condition $\sum_{n=0}^{\infty} p_n = 1$,
i.e. $p_0(1 + \rho + \rho^2 + \ldots) = 1$

$$\text{or} \quad p_0 = (1 - \rho), \;\; \text{since} \;\; \rho = \lambda/\mu < 1. \tag{15.14}$$

Therefore, (15.13) becomes

$$p_n = (1 - \rho)p^n, \quad n = 0, 1, 2, \ldots \tag{15.15}$$

This gives the probability that there are $n$ units in the system at any time. Equations (15.14) and (15.15) rather give the required probability distribution of the queue length.

We note that the probability distribution in (15.15) depends only on the traffic intensity ratio $\rho = \lambda/\mu$. Obviously in order for $p_0$ to exist, it is required that $(\rho =)\lambda/\mu < 1$, i.e. $\lambda < \mu$; otherwise, the series will diverge. This is the stability condition for the *M/M/*1 model.

***Characteristics of the Model***

(i) *Probability of queue size greater than or equal to N*, i.e. prob *(queue size* $\geq$ *N)*

$$\begin{aligned}\text{prob (queue size} \geq N) &= p_N + p_{N+1} + p_{N+2} + \cdots \\ &= p_0 \rho^N \left(1 + \rho + \rho^2 + \cdots\right) \\ &= p_0 \frac{\rho^N}{1 - \rho} = \rho^N = \left(\frac{\lambda}{\mu}\right)^N [\because p_0 = 1 - \rho].\end{aligned}$$

$$\therefore \quad \text{prob (queue size} \geq N) = \rho^N = \left(\frac{\lambda}{\mu}\right)^N.$$

(ii) *Expected line length* $L_s$

$L_s$ = expected number of customers in the system, i.e. expected line length or average number of customers in the system

$$\begin{aligned}
&= \sum_{n=0}^{\infty} np_n = \sum_{n=0}^{\infty} n\rho^n p_0 = p_0 \sum_{n=0}^{\infty} n\rho^n \\
&= p_0\rho[1+2\rho+3\rho^2+4\rho^3+\cdots] \\
&= p_0\rho(1-\rho)^{-2} = (1-\rho)\rho\frac{1}{(1-\rho)^2} \quad [\because\ p_0 = 1-\rho]. \\
&= \frac{\rho}{1-\rho} = \frac{\lambda}{\mu-\lambda} \quad \left[\because\ \rho = \frac{\lambda}{\mu}\right]
\end{aligned}$$

(iii) *Expected queue length* $L_q$

$L_q$= expected number of customers in queue (i.e. expected queue length) $= \sum_{n=1}^{\infty}(n-1)p_n$ [Since there are $(n-1)$ customers in the queue excluding the one in service]

$$\begin{aligned}
&= \sum_{n=1}^{\infty} np_n - \sum_{n=1}^{\infty} p_n = L_s - \left(\sum_{n=0}^{\infty} p_n - p_0\right) \\
&= L_s - (1-p_0) = L_s - \rho = \frac{\rho}{1-\rho} - \rho \quad \left[\because L_s = \frac{\rho}{1-\rho}\right] \\
&= \frac{\lambda^2}{\mu(\mu-\lambda)} \quad \left[\because\ \rho = \frac{\lambda}{\mu}\right]
\end{aligned}$$

Note: $L_s = L_q + \rho = L_q + \frac{\lambda}{\mu}$ as $L_q = L_s - \rho$.

(iv) *Variance of queue length*

From *the definition of variance*, $\mathrm{Var}(n) = E\{n^2\} - \{E(n)\}^2$

$$\begin{aligned}
&= \sum_{n=1}^{\infty} n^2 p_n - \left\{\sum_{n=1}^{\infty} np_n\right\}^2 \\
&= \sum_{n=1}^{\infty} n^2\rho^n(1-\rho) - \left(\frac{\rho}{1-\rho}\right)^2 \\
&\left[\because L_s = \sum_{n=1}^{\infty} np_n = \frac{\rho}{1-\rho} \text{ and } p_n = \rho^n(1-\rho)\right.
\end{aligned}$$

$$
\begin{aligned}
&= \rho(1-\rho) + 2^2\rho^2(1-\rho) + 3^2\rho^3(1-\rho) + 4^2\rho^4(1-\rho) + \cdots - \frac{\rho^2}{(1-\rho)^2} \\
&= \rho(1-\rho)\left[1 + 2^2\rho + 3^2\rho^2 + 4^2\rho^3 + \cdots\right] - \frac{\rho^2}{(1-\rho)^2} \\
&= \rho(1-\rho)S - \frac{\rho^2}{(1-\rho)^2}, \quad \text{where } S = 1 + 2^2\rho + 3^2\rho^2 + 4^2\rho^3 + \cdots
\end{aligned}
$$

Now, we have to calculate the simplified expression of $S$.

$$S = 1 + 2^2\rho + 3^2\rho^2 + 4^2\rho^3 + \cdots$$

Integrating both sides with respect to $\rho$ from 0 to $\rho$, we have

$$
\begin{aligned}
\int_{\rho=0}^{\rho} S \mathrm{d}\rho &= \int_{\rho=0}^{\rho} \left(1 + 2^2\rho + 3^2\rho^2 + 4^2\rho^3 + \cdots\right)\mathrm{d}\rho \\
&= \rho + 2\rho^2 + 3\rho^3 + 4\rho^4 + 5\rho^5 + \cdots \\
&= \rho(1-\rho)^{-2} \\
\therefore \int_{\rho=0}^{\rho} S \mathrm{d}\rho &= \rho(1-\rho)^{-2} = \frac{\rho}{(1-\rho)^2}.
\end{aligned}
$$

Differentiating both sides with respect to $\rho$, we have

$$
\begin{aligned}
S &= \frac{1}{(1-\rho)^2} + \rho\frac{-2}{(1-\rho)^3}(-1) = \frac{1}{(1-\rho)^2} + \frac{2\rho}{(1-\rho)^3} = \frac{1-\rho+2\rho}{(1-\rho)^3} \\
&= \frac{1+\rho}{(1-\rho)^3}.
\end{aligned}
$$

$$
\begin{aligned}
\therefore \mathrm{Var}(n) &= \rho(1-\rho)S - \frac{\rho^2}{(1-\rho)^2} \\
&= \rho(1-\rho)\frac{1+\rho}{(1-\rho)^3} - \frac{\rho^2}{(1-\rho)^2}. \\
&= \frac{\rho(1+\rho)}{(1-\rho)^2} - \frac{\rho^2}{(1-\rho)^2} = \frac{\rho}{(1-\rho)^2}
\end{aligned}
$$

(v) *Probability density function of waiting time excluding the service time distribution*

In a steady state, each customer has the same waiting time distribution. Let this distribution be a continuous distribution with probability density function $\psi(w)$, and we denote by $\psi(w)\,dw$ the probability that a customer begins to be served in the interval $(w, w + dw)$, where $w$ is measured from the time of his arrival. We suppose that a customer arrives at time $w = 0$ and his service begins in the interval $(w, w + dw)$.


There may be two possibilities:

(i) There is a finite probability $p_0$ (the probability that the system is empty) that the waiting time is zero.
(ii) Let there be $n$ customers in the system, $(n - 1)$ waiting, one in the service, when customer A arrives. Therefore, before the service of $A$ begins, $(n - 1)$ customers must leave in the time interval $(0, w)$ and with the $n$th customer in $(w, w + dw)$.

As the server's mean rate of service is $\mu$ in unit time or $\mu w$ in time $w$, and the service distribution is Poisson, we have

prob $[(n - 1)$ customers waiting are served in time $w]$

$$= \frac{(\mu w)^{n-1} e^{-\mu w}}{(n-1)!}$$

and prob [one customer is served in time $dw$] = $\mu\, dw$.

$$\begin{aligned}
\therefore\ \psi_n(w)dw &= \text{probability that a new arrival is taken into service} \\
&\quad \text{after a time lying between } w \text{ and } w + dw \\
&= \text{prob}[(n-1)\text{customers waiting are served in time } w] \times \text{prob} \\
&\quad [\text{one customer is served in time } dw] \\
&= \frac{(\mu w)^{n-1} e^{-\mu w}}{(n-1)!} \mu\, dw.
\end{aligned}$$

Let $W$ be the waiting time of the customer who has to wait such that

$$w \leq W \leq w + dw.$$

Since the queue length can vary between 1 and $\infty$, the probability density of the waiting time is given by

$$\begin{aligned}
\psi(w)\mathrm{d}w &= p(w \le W \le w + \mathrm{d}w) \\
&= \sum_{n=1}^{\infty} [\text{probability that there are } n \text{ customers in the system}] \times \psi_n(w)\mathrm{d}w \\
&= \sum_{n=1}^{\infty} \rho^n (1-\rho) \frac{(\mu w)^{n-1} \mathrm{e}^{-\mu w}}{(n-1)!} \mu \mathrm{d}w \quad [\because\ p_n = \rho^n(1-\rho) \\
&= (1-\rho)\mathrm{e}^{-\mu w} \mu \mathrm{d}w \sum_{n=1}^{\infty} \rho^n \frac{(\mu w)^{n-1}}{(n-1)!} = \rho(1-\rho)\mathrm{e}^{-\mu w}\mu \mathrm{d}w \sum_{n=1}^{\infty} \frac{(\rho\mu w)^{n-1}}{(n-1)!} \\
&= \frac{\lambda}{\mu}\left(1-\frac{\lambda}{\mu}\right)\mathrm{e}^{-\mu w}\mu \mathrm{d}w \sum_{n=1}^{\infty} \frac{\left(\frac{\lambda}{\mu}\mu w\right)^{n-1}}{(n-1)!} = \lambda\left(1-\frac{\lambda}{\mu}\right)\mathrm{e}^{-\mu w}\mathrm{d}w \sum_{n=1}^{\infty} \frac{(\lambda w)^{n-1}}{(n-1)!} \\
&= \lambda\left(1-\frac{\lambda}{\mu}\right) e^{-\mu w}\mathrm{d}w \left[1 + \frac{\lambda w}{1!} + \frac{(\lambda w)^2}{2!} + \cdots\right] = \lambda\left(1-\frac{\lambda}{\mu}\right)\mathrm{e}^{-\mu w} \cdot \mathrm{e}^{\lambda w}\mathrm{d}w \\
&= \lambda\left(1-\frac{\lambda}{\mu}\right)\mathrm{e}^{-(\mu-\lambda)w}\mathrm{d}w \quad \text{where } w > 0.
\end{aligned}$$

Hence,

$$\begin{aligned}
\int_0^{\infty} \psi(w)\mathrm{d}w &= \int_0^{\infty} \lambda\left(1-\frac{\lambda}{\mu}\right)\mathrm{e}^{-(\mu-\lambda)w}\mathrm{d}w = \lim_{B\to\infty}\int_0^{B} \lambda\left(1-\frac{\lambda}{\mu}\right)\mathrm{e}^{-(\mu-\lambda)w}\mathrm{d}w \\
&= \lambda\left(1-\frac{\lambda}{\mu}\right)\lim_{B\to\infty}\left[\frac{\mathrm{e}^{-(\mu-\lambda)w}}{-(\mu-\lambda)}\right]_0^B \\
&= \lambda\left(1-\frac{\lambda}{\mu}\right)\left[0 + \frac{1}{(\mu-\lambda)}\right] = \lambda\frac{\mu-\lambda}{\mu}\frac{1}{\mu-\lambda} = \frac{\lambda}{\mu} \ne 1.
\end{aligned}$$

This happens because the case for which $w = 0$ is not included.
Thus,

$$\text{prob}[w > 0] = \int_0^{\infty} \psi(w)\mathrm{d}w = \rho$$

and prob[$w = 0$] = prob[no customer in the system] = $p_0 = 1 - \rho$.

Since the sum of these probabilities of waiting time is 1, the complete distribution of the waiting time is:

(a) Continuous for $w \le W \le w + \mathrm{d}w$ with probability density function $\psi(w)$ given by

$$\psi(w)\mathrm{d}w = \lambda\left(1 - \frac{\lambda}{\mu}\right)\mathrm{e}^{-(\mu-\lambda)w}\mathrm{d}w$$

(b) Discrete for $w = 0$ with $\mathrm{prob}(w = 0) = 1 - \frac{\lambda}{\mu}$.

Note that the probability that the waiting time exceeds $w_1$ is

$$\int_{w_1}^{\infty} \psi(w)\mathrm{d}w = \left[-\frac{\lambda}{\mu}\mathrm{e}^{-(\mu-\lambda)w}\right]_{w_1}^{\infty} = \frac{\lambda}{\mu}\mathrm{e}^{-(\mu-\lambda)w_1},$$

which does not include the service time.

(vi) *Busy period distribution, i.e. the conditional density function for waiting time, given that a person has to wait*

Here we find the probability density function for the distribution of total time (waiting plus service) that an arrival spends in the system.

Let $\psi(w/w > 0)$ = probability density function for waiting time such that a person has to wait

$$\begin{aligned}
&= \frac{\psi(w)}{\mathrm{prob}(w > 0)} \\
&= \frac{\psi(w)}{\rho} \qquad \left[\begin{array}{l} \because \mathrm{prob}(w > 0) + \mathrm{prob}(w = 0) = 1 \\ \therefore \mathrm{prob}(w > 0) = 1 - \mathrm{prob}(w = 0) = 1 - (1 - \rho) = \rho \end{array}\right. \\
&= \frac{\lambda\left(1 - \frac{\lambda}{\mu}\right)\mathrm{e}^{-(\mu-\lambda)w}}{\frac{\lambda}{\mu}} = (\mu - \lambda)\mathrm{e}^{-(\mu-\lambda)w}
\end{aligned} \tag{15.16}$$

Now $\int_0^\infty \psi(w/w > 0) = \int_0^\infty (\mu - \lambda)\mathrm{e}^{-(\mu-\lambda)w}\mathrm{d}w = 1.$
Thus, (15.16) gives the required probability density function for the busy period.

(vii) *Mean or expected waiting time in the queue, i.e.* $W_\mathrm{q}$ (*average waiting time of an arrival in the queue*)

We have

$$\begin{aligned}
W_\mathrm{q} &= \int_0^\infty w\psi(w)\mathrm{d}w \\
&= \int_0^\infty w\lambda\left(1 - \frac{\lambda}{\mu}\right)\mathrm{e}^{-(\mu-\lambda)w}\mathrm{d}w
\end{aligned}$$

$$= w\lambda\left(1-\frac{\lambda}{\mu}\right)\frac{e^{-(\mu-\lambda)w}}{-(\mu-\lambda)}\Bigg|_0^\infty + \lambda \underset{B\to\infty}{Lt}\int_0^B\left(1-\frac{\mu}{\lambda}\right)\frac{e^{-(\mu-\lambda)w}}{\mu-\lambda}\,dw$$

[Integrating by parts taking was the first function]

$$= 0 - 0 + \lambda\left(1-\frac{\lambda}{\mu}\right)\underset{B\to\infty}{Lt}\left.\frac{e^{-(\mu-\lambda)w}}{-(\mu-\lambda)^2}\right]_0^B$$

$$= -\lambda\left(\frac{\mu-\lambda}{\mu}\right)\left.\frac{1}{(\mu-\lambda)^2}e^{-(\mu-\lambda)w}\right]_0^\infty$$

$$= 0 + \frac{\lambda(\mu-\lambda)}{\mu(\mu-\lambda)^2}\cdot 1 = \frac{\lambda}{\mu(\mu-\lambda)} = \frac{1}{\lambda}L_q, \quad \text{where } L_q = \frac{\lambda^2}{\mu(\mu-\lambda)}.$$

(viii) *Expected waiting time in the system $W_s$, i.e. average time that an arrival spends in the system*

$$\begin{aligned} W_s &= \text{expected waiting time in the system} \\ &= \text{expected waiting time in queue plus expected service time} \\ &= W_q + \frac{1}{\mu} \quad \left[\text{Since the expected service time} = \text{mean service time} = \frac{1}{\mu}\right] \\ &= \frac{\lambda}{\mu(\mu-\lambda)} + \frac{1}{\mu} \quad \left[\because\ W_q = \frac{\lambda}{\mu(\mu-\lambda)}\right] \\ &= \frac{1}{\mu-\lambda}. \end{aligned}$$

As $W_s = \dfrac{1}{\mu-\lambda}$ and $L_s = \dfrac{\lambda}{\mu-\lambda}$, $L_s = \lambda W_s$.

(ix) *Expected length of non-empty queue, or average length of non-empty queue, i.e. L/L > 0.*

We have

$$\begin{aligned} (L/L>0) &= \frac{L_q}{\text{prob[an arrival has to wait, } L>0]} = \frac{L_q}{\text{prob}[(n-1)>0]} = \frac{L_q}{\sum_{n=2}^\infty p_n} \\ &= \frac{L_q}{\sum_{n=0}^\infty p_n - (p_0+p_1)} = \frac{L_q}{1-p_0-\rho p_0} = \frac{L_q}{1-p_0(1+\rho)} \\ &= \frac{1}{1-\rho} = \frac{\mu}{\mu-\lambda}. \end{aligned}$$

**Note**: From the preceding characteristics, we can write $L_s = L_q + \frac{\lambda}{\mu}$, $L_s = \lambda W_s$, $L_q = \lambda W_q$ and $W_s = W_q + \frac{1}{\mu}$. These formulas are known as Little's formula.

*Example 1* Arrivals at a telephone booth are considered to have a Poisson distribution with an average time of 10 min between one arrival and the next. The length of a phone call is assumed to be distributed exponentially with mean 3 min.

(a) What is the probability that a person arriving at the booth will have to wait?
(b) What is the average length of queues that form from time to time?
(c) The telephone company will install a second booth when convinced that an arrival would expect to have to wait at least 3 min for the phone. By how much time should the flow of arrivals be increased to justify a second booth?
(d) Find the average number of units in the system.
(e) What is the probability that an arrival has to wait more than 10 min before the phone is free?
(f) Estimate the fraction of a day that the phone will be in use (or busy).

**Solution**
This is an $(M/M/1)$:$(\infty/\text{FCFS}/\infty)$ problem.

In this problem, the mean arrival rate is $\lambda = \frac{1}{10}$ and the mean service time is $\mu = \frac{1}{3}$.

(a) Now the probability that a person arriving at the telephone booth will have to wait

$$= \text{prob}[w > 0] = 1 - \text{prob}(w = 0) = 1 - p_0 = 1 - (1 - \rho) = \rho = \frac{\lambda}{\mu} = 0.3.$$

(b) The average length of the queues that form from time to time

$$= L/L > 0 = \frac{\mu}{\mu - \lambda} = 1.43 \text{ persons}$$

(c) The installation of a second booth will be justified if the arrival rate is greater than the waiting time. Then the length of the queue will go on increasing.

In that case, let $\lambda'$ be the mean arrival rate.

We know that the average waiting time of an interval in the queue is $W_q = \frac{\lambda'}{\mu(\mu - \lambda')}$.

Here,

$$W_q = 3, \quad \mu = \frac{1}{3} \quad \therefore \; 3 = \frac{\lambda'}{\frac{1}{3}\left(\frac{1}{3} - \lambda'\right)} \text{or,} \quad \lambda' = \frac{1}{3} - \lambda' \text{ or } \lambda' = \frac{1}{6}.$$

Hence, the increase in mean arrival rate $= \lambda' - \lambda = \frac{1}{6} - \frac{1}{10} = \frac{1}{15} = 0.067$ arrival per minute.

Again, the increase in the flow of arrivals $= \frac{1}{15} \times 60 = 4$ per hour.

Therefore, the second booth is justified if the increase in arrival rate is 0.067 arrival per minute (or 4 persons per hour).

(d) The average number of units in the system is $L_s = \frac{\lambda}{\mu - \lambda} = 0.43$ persons.

(e) prob [waiting time of an arrival in the queue > 10]

$$= \int_{10}^{\infty} \psi(w)\mathrm{d}w = \int_{10}^{\infty} \frac{\lambda}{\mu}(\mu - \lambda)\mathrm{e}^{-(\mu-\lambda)w}\mathrm{d}w = \lim_{B\to\infty} \frac{\lambda}{\mu}(\mu - \lambda)\left[\frac{\mathrm{e}^{-(\mu-\lambda)w}}{-(\mu - \lambda)}\right]_{10}^{B}$$

$$= -\frac{\lambda}{\mu}\left[0 - \mathrm{e}^{-(\mu-\lambda)10}\right] = 0.03 \text{ (approx)}$$

(f) The fraction of a day that the phone will be busy = traffic intensity $= \rho = \frac{\lambda}{\mu} = 0.3$.

*Example 2* In a railway yard, goods trains arrive at a rate 30 trains per day. Assume that the inter-arrival time follows an exponential distribution, and the service time (the time taken to hump a train) distribution is also exponential with an average of 36 min. Calculate the following:

(i) The average number of trains in the system
(ii) The probability that the number of trains in the system exceeds 10
(iii) Expected waiting time in the queue
(iv) Average number of trains in the queue
(v) If the input of trains increases to an average 33 per day, what will be the change in (i) and (ii)?

**Solution**

This is an $(M/M/1)$:($\infty$/FCFS/$\infty$) problem.

In this problem, the mean arrival rate $\lambda = 30$ trains/day $= \frac{30}{60\times 24}$ trains/minute $= \frac{1}{48}$ trains/min, and the mean service time $\mu = \frac{1}{36}$ trains/min.

Hence, the traffic intensity $\rho = \frac{\lambda}{\mu} = \frac{\frac{1}{48}}{\frac{1}{36}} = \frac{36}{48} = \frac{3}{4} = 0.75$.

(i) The average number of trains in the system is given by $L_s = \frac{\rho}{1-\rho} = \frac{0.75}{1-0.75} = \frac{0.75}{0.25} = 3$ trains.

(ii) The probability that the number of trains in the system exceeds $10 = \text{prob(queue size } \geq 10) = \rho^{10} = (0.75)^{10} = 0.056$.

(iii) The expected waiting time in the queue is given by $W_q = \frac{\lambda}{\mu(\mu-\lambda)} = \frac{1/48}{\frac{1}{36}\left(\frac{1}{36} - \frac{1}{48}\right)}$ min = 108 min*or*1 h 48 min.

(iv) The average number of trains in the queue is given by

$L_{\mathrm{q}} = \frac{\lambda^2}{\mu(\mu-\lambda)} = \frac{\left(\frac{1}{48}\right)^2}{\frac{1}{36}\left(\frac{1}{36}-\frac{1}{48}\right)} = \frac{9}{4} = 2.25$or nearly 2 trains.

(v) If the input of trains increases to 33 trains per day, then $\lambda = \frac{33}{60\times 24}$ trains/min $= \frac{11}{480}$ trains/min. Hence, $\rho = \frac{\lambda}{\mu} = \frac{\frac{11}{480}}{\frac{1}{36}} = \frac{33}{40}$.

In this case, $L_{\mathrm{s}} = \frac{\rho}{1-\rho} = \frac{\frac{33}{40}}{1-\frac{33}{40}} = \frac{33}{7} = 5$ trains (approx.) and prob(queue size $\geq 10) = \rho^{10} = \left(\frac{33}{40}\right)^{10} = 0.2$ (approx.).

### *15.3.3 Model (M/M/1):(N/FCFS/∞)*

**Assumptions** In this model, we shall discuss a single server finite queueing system under the following assumptions:

(i) Customers arrive according to a Poisson fashion.
(ii) The service rate follows a Poisson distribution.
(iii) The queueing system has only a single service channel.
(iv) The service discipline is FCFS (first come, first served).
(v) The maximum number of customers in the system is limited to $N$.

If a new customer arrives when $N$ persons are already in the system, then the customer is restricted from entering, and is said to be lost from the system. This is often referred to as blocking.

***Model Formulation***

Let $n$ be the customers in the system. Let $\lambda$ and $\mu$ be the mean arrival and service rate of units respectively.

In this model,

$$\lambda_n = \begin{cases} \lambda \text{ when } n < N \text{ i.e., } (n = 0, 1, 2, \ldots .N-1) \\ 0 \text{ when } n \geq N \end{cases}$$

and $\mu_n = \mu$ (independent of $n$).

Therefore, the probability of one arrival during $\Delta t$ is $\lambda_n \Delta t + O(\Delta t)$ for $n \geq N$, and the probability of one departure or service during $\Delta t$ is $\mu \Delta t + O(\Delta t)$.

To have 0 (zero) customer in the system at time $t + \Delta t$, there may be the following mutually exclusive cases:

(i) No customer in the system at time $t$, no arrival in time $\Delta t$
(ii) One customer in the system at time $t$, no arrival in time $\Delta t$, one service in time $\Delta t$, and so on.

Therefore, the probability of no customer in the system at time $t + \Delta t$ is given by

$$\begin{aligned} p_0(t+\Delta t) &= p_0(t)(1-\lambda\Delta t) + p_1(t)(1-\lambda\Delta t)\mu\Delta t + O(\Delta t) \\ &= p_0(t) - \lambda p_0(t)\Delta t + \mu p_1(t)\Delta t + O(\Delta t) \end{aligned}$$

$$\text{or} \quad \frac{p_0(t+\Delta t) - p_0(t)}{\Delta t} = -\lambda p_0(t) + \mu p_1(t) + \frac{O(\Delta t)}{\Delta t} \quad \text{for } n = 0. \tag{15.17}$$

To have $n$ (>0) customers in the system at time $t + \Delta t$, there may be the following mutually exclusive cases:

(i) $n$ customers in the system at time $t$, no arrival in time $\Delta t$, no service in time $\Delta t$

(ii) $(n - 1)$ customers in the system at time $t$, one arrival in time $\Delta t$, no service in time $\Delta t$

(iii) $(n + 1)$ customers in the system at time $t$, no arrival in time $\Delta t$, one service in time $\Delta t$, and so on.

Therefore, the probability of $n$ customers in the system at time $(t + \Delta t)$ is given by

$$\begin{aligned} p_n(t+\Delta t) &= p_{n-1}(t)\lambda_{n-1}\Delta t(1-\mu\Delta t) + p_n(t)(1-\lambda_n\Delta t)(1-\mu\Delta t) \\ &\quad + p_{n+1}(t)(1-\lambda_{n+1}\Delta t)\mu\Delta t + O(\Delta t) \end{aligned}$$

$$\text{or} \quad \frac{p_n(t+\Delta t) - p_n(t)}{\Delta t} = \lambda_{n-1}p_{n-1}(t) - (\lambda_n + \mu)p_n(t) + \mu p_{n+1}(t) + \frac{O(\Delta t)}{\Delta t}. \tag{15.18}$$

When $n < N$, then $\lambda_{n-1} = \lambda, \lambda_n = \lambda$.
Hence, from (15.18), we have

$$\frac{p_n(t+\Delta t) - p_n(t)}{\Delta t} = \lambda p_{n-1}(t) - (\lambda + \mu)p_n(t) + \mu p_{n+1}(t) + \frac{O(\Delta t)}{\Delta t}. \tag{15.19}$$

When $n = N$, $p_{n+1} = 0, \lambda_{n-1} = \lambda,\ \lambda_n = 0$.
From (15.18), we have

$$\frac{p_N(t+\Delta t) - p_N(t)}{\Delta t} = \lambda p_{N-1}(t) - \mu p_N(t) + \frac{O(\Delta t)}{\Delta t}. \tag{15.20}$$

Taking the limit as $\Delta t \to 0$ in Eqs. (15.17), (15.19) and (15.20), we have

$$\begin{aligned} p_0'(t) &= -\lambda p_0(t) + \mu p_1(t) \text{ for } n = 0 \\ p_n'(t) &= \lambda p_{n-1}(t) - (\lambda + \mu)p_n(t) + \mu p_{n+1}(t) \quad \text{for } 0 < n < N \\ p_n'(t) &= \lambda p_{N-1}(t) - \mu p_N(t) \text{ for } n = N \end{aligned}$$

Under a steady state condition of the system, these equations reduce to

$$-\lambda p_0 + \mu p_1 = 0 \quad \text{for} \;\; n = 0 \tag{15.21}$$

$$\lambda p_{n-1} - (\lambda + \mu)p_n + \mu p_{n+1} = 0 \quad \text{for} \;\; 1 \le n < N \tag{15.22}$$

$$\lambda p_{N-1} - \mu p_N = 0 \quad \text{for} \;\; n = N. \tag{15.23}$$

***Solution of the Model***

From (15.21), we have $p_1 = \frac{\lambda}{\mu} p_0 = \rho p_0$ where $\rho = \frac{\lambda}{\mu}$ .
Setting $n$ = 1 in (15.22), we have

$$\begin{aligned} &\lambda p_0 - (\lambda + \mu)p_1 + \mu p_2 = 0 \\ \text{or} \;\; &p_2 = -\frac{\lambda}{\mu} p_0 + \left(1 + \frac{\lambda}{\mu}\right) p_1 = \rho p_1 = \rho^2 p_0. \end{aligned}$$

Setting $n$ = 2 in (15.22), we have

$$\begin{aligned} &\lambda p_1 - (\lambda + \mu)p_2 + \mu p_3 = 0 \\ \therefore \;\; &p_3 = -\frac{\lambda}{\mu} p_1 + \left(1 + \frac{\lambda}{\mu}\right) p_2 = \rho p_2 = \rho^3 p_0. \end{aligned}$$

Proceeding in this way, we have

$$\begin{aligned} p_n &= \left(\frac{\lambda}{\mu}\right)^n p_0 = \rho^n p_0 \quad \text{for} \;\; 1 \le n < N \\ \therefore \;\; p_{N-1} &= \left(\frac{\lambda}{\mu}\right)^{N-1} p_0 = \rho^{N-1} p_0. \end{aligned} \tag{15.24}$$

From (15.23), we have $p_N = \frac{\lambda}{\mu} p_{N-1} = \rho^N p_0$

$$\therefore \;\; p_n = \rho^n p_0 \;\text{for}\; 0 \le n \le N. \tag{15.25}$$

Since the capacity of the system is $N$,

$$\sum_{n=0}^{N} p_n = 1$$

$$\text{or } \; p_0 \frac{1-\rho^{N+1}}{1-\rho} = 1 \; \text{ for } \; \rho \neq 1 \text{ and } (N+1)p_0 = 1 \; \text{ for } \; \rho = 1$$

$$\text{or} \quad p_0 = \begin{cases} \frac{1-\rho}{1-\rho^{N+1}} & \text{for } \; \rho \neq 1 \\ \frac{1}{N+1} & \text{for } \; \rho = 1 \end{cases}.$$

From (15.9), we have

$$p_n = \rho^n p_0 = \begin{cases} \frac{(1-\rho)\rho^n}{1-\rho^{N+1}} & \text{for } \; 0 \leq n \leq N \text{ and } \; \rho \neq 1 \\ \frac{1}{N+1} & \text{for } \; 0 \leq n \leq N \text{ and } \; \rho = 1 \end{cases}$$

**Note**

(i) This solution holds for any value of $\rho$, as the number of customers allowed in the system is controlled by the queue length ($N$ – 1), not by the rates of arrival and departure $\lambda$ and $\mu$. So in this model we do not require the condition $\rho < 1$, and hence the system is stable for $\rho < 1$ as well as $\rho \geq 1$.

(ii) $\lambda_{\text{eff}}$ is the effective arrival rate in the system, i.e. the mean rate of customers entering the system. It equals the arrival rate $\lambda$ when all arriving customers can join the system; otherwise, $\lambda_{\text{eff}} < \lambda$. Since the probability that a customer does not join the system is $p_N$, the probability that the customers join the system is $(1 - p_N)$. Hence, $\lambda_{\text{eff}} = \lambda(1 - p_N)$.

(iii) In this model, customers arrive at the rate $\lambda$, but not all arrivals can join the system. So the equations $L_s = \lambda W_s$ and $L_q = \lambda W_q$ are modified by redefining $\lambda$ to include only those customers who actually join the system.

Hence, for this model, $L_s = \lambda_{\text{eff}} W_s$ and $L_q = \lambda_{\text{eff}} W_q$, where $\lambda_{\text{eff}}$ = effective arrival rate for those who join the system.

***Characteristics of the Model***

(i) ***Expected line length, i.e. average number of customers in the system***

The expected line length, i.e. the average number of customers in the system, is given by

$$L_s = E(n) = \sum_{n=0}^{N} np_n = \sum_{n=0}^{N} n \frac{1-\rho}{1-\rho^{N+1}} \rho^n = \frac{1-\rho}{1-\rho^{N+1}} \sum_{n=0}^{N} n\rho^n \quad \text{when } \; \rho \neq 1$$

$$= \frac{1-\rho}{1-\rho^{N+1}} \left(\rho + 2\rho^2 + 3\rho^3 + \cdots + N\rho^N\right).$$

[Let $S = \rho + 2\rho^2 + 3\rho^3 + \cdots + N\rho^N$.

$\therefore \rho S = \rho^2 + 2\rho^3 + \cdots + N\rho^{N+1}$.
Subtracting the second equation from the first, we have

$$
\begin{aligned}
& S(1-\rho) = \rho + \rho^2 + \rho^3 + \cdots n \text{ times } - N\rho^{N+1} \\
\text{or} \quad & S(1-\rho) = \frac{\rho(1-\rho^N)}{1-\rho} - N\rho^{N+1} \\
\text{or} \quad & S = \left[\frac{\rho(1-\rho^N)}{1-\rho} - N\rho^{N+1}\right]\frac{1}{1-\rho} \\
& = \frac{1-\rho}{1-\rho^{N+1}}\left[\frac{\rho(1-\rho^N)}{1-\rho} - N\rho^{N+1}\right]/(1-\rho) \\
& = \frac{1}{1-\rho^{N+1}}\left[\frac{\rho - \rho^{N+1} - N\rho^{N+1} + N\rho^{N+2}}{1-\rho}\right] \\
& = \frac{\rho[1-(N+1)\rho^N + N\rho^{N+1}]}{(1-\rho)(1-\rho^{N+1})} \quad \text{when } \rho \neq 1.
\end{aligned}
$$

For

$$\rho = 1,\ L_s = \sum_{n=0}^{N} np_n = \sum_{n=0}^{N} \frac{n}{N+1} = \frac{1}{N+1}(1+2+\cdots+N) = \frac{N}{2}$$

$$\left[\text{For } \rho = 1, p_n = \frac{1}{N+1} \text{ and } \sum_{n=0}^{N} p_n = 1\right].$$

(ii) ***Average number of customers in the queue***

The average number of customers in the queue

$$
\begin{aligned}
& = L_q = \sum_{n=1}^{N}(n-1)p_n = \sum_{n=1}^{N} np_n - \sum_{n=1}^{N} p_n \\
& = \sum_{n=0}^{N} np_n - \sum_{n=0}^{N} p_n + p_0 = L_s - 1 + p_0 \left[\because\ L_s = \sum_{n=0}^{N} np_n\right] \\
& = \frac{\rho[1-(N+1)\rho^N + N\rho^{N+1}]}{(1-\rho)(1-\rho^{N+1})} - \left(1 - \frac{1-\rho}{1-\rho^{N+1}}\right) \\
& = \rho^2\left[1 - N\rho^{N-1} + (N-1)\rho^N\right]/(1-\rho)\left(1-\rho^{N-1}\right).
\end{aligned}
$$

(iii) ***Waiting time in the queue***

The waiting time in the queue is

$$W_q = \frac{L_q}{\lambda_{\text{eff}}} = \frac{L_q}{\lambda(1-p_N)} [\because \ \lambda_{\text{eff}} = \lambda(1-p_N)]$$

$$\frac{L_q}{\lambda\left[1 - \frac{(1-\rho)\rho^N}{1-\rho^{N+1}}\right]} = \frac{\rho^2[1 - N\rho^{N-1} + (N-1)\rho^N]}{\lambda(1-\rho^N)(1-\rho)}.$$

(iv) ***Waiting time in the system***

The waiting time in the system is given by

$$W_s = W_q + \frac{1}{\mu}\left(\text{or} \frac{L_s}{\lambda_{\text{eff}}}\right).$$

**Special case**:

If $\rho < 1$ and $N \to \infty$, then the steady state solution is $p_n = (1-\rho)\rho^n$, and the resulting system reduces to the result $(M/M/1)$:$(\infty/\text{FCFS}/\infty)$.

*Example 3* In a car wash service facility, information gathering indicates that cars arrive for service according to a Poisson distribution with mean 5 per hour. The time for washing and cleaning for each car varies but is found to follow an exponential distribution with mean 10 min per car. The facility cannot handle more than one car at a time and has a total of 5 parking spaces. If the parking spot is full, newly arriving cars balk to seek services elsewhere.

(a) How many customers is the manager of the facility losing due to the limited parking space?
(b) What is the expected waiting time until a car is washed?

**Solution**
In this system, there may be 5 cars in the 5 parking spots and one can be serviced; i.e. the capacity of this system is $N = 5 + 1 = 6$.
Hence, this problem is of the model ($M/M/1$):(6/FCFS/∞).
Here $\lambda = 5$ per hour, $\mu = \frac{1}{10} \times 60 = 6$ per hour $\therefore \ \rho = \frac{\lambda}{\mu} = \frac{5}{6}$.
Now $p_N = p_6 = \frac{1-\rho}{1-\rho^{6+1}}\rho^6 = 0.0774$ [Since $p_N = \frac{1-\rho}{1-\rho^{N+1}}\rho^N$].

(a) Therefore, the rate at which the cars balk
$= \lambda - \lambda_{\text{eff}} = \lambda - \lambda(1-p_N) = \lambda p_6 = 5 \times .0774 = 0.387$ car/h.

Assuming 8 working hours per day, a manager will lose $0.387 \times 8 = 3.096 = 3$ cars a day.

(b) The expected waiting time until a car is washed is given by

$$W_s = \frac{L_s}{\lambda_{\text{eff}}}$$

Here $L_s = \frac{\rho[1-(N+1)\rho^N + N\rho^{N+1}]}{(1-\rho)(1-\rho^{N+1})} = 2.29$ h,
and $\lambda_{\text{eff}} = \lambda(1 - p_N) = \lambda(1 - p_6) = 4.613$.
Hence, $W_s = \frac{L_s}{\lambda_{eff}} = 0.496$ h.

### 15.3.4 Model (M/M/C):(∞/FCFS/∞)

**Assumptions** In this model, we shall discuss a multiserver queueing system under the following assumptions:

(i) The arrival rate follows a Poisson distribution.
(ii) This queueing system deals with queues which are being served by parallel service channels in which the service time of each server has an independently and identically distributed with exponential service time distribution.
(iii) This queueing system has $c(>1)$ service channels.
(iv) The service discipline is first come first served.
(v) The capacity of the system is infinite.
(vi) The length of the waiting line will depend on the number of occupied channels.

***Model Formulation***
Let $n$ be the number of customers in the system.
Therefore, according to assumption (i), the mean arrival rate is given by $\lambda_n = \lambda$ for all $n$.

For the mean service rate, if there are more than $c$ customers in the system $(n > c)$, all the $c$ servers will remain busy and each is providing service at a mean rate $\mu$, while $(n - c)$ customers will be waiting in the queue. Thus, the mean service rate is $c\mu$.

If $n = c$, then all the service channels will be busy, each providing at a mean service rate $\mu$, and thus the mean rate of service is $c\mu$.

If there are fewer than $c$ customers in the system $(n < c)$, only $n$ of the $c$ servers will be busy, and in this case $(c - n)$ servers will remain idle; there will be no customer waiting in the queue. Thus, the mean service rate is $n\mu$.

Hence, for this model,

$$\lambda_n = \lambda,\ n = 0, 1, 2, \ldots$$
$$\mu_n = \begin{cases} n\mu, & 0 \le n < c \\ c\mu & n \ge c \end{cases},$$

i.e. the probability of one arrival during $\Delta t$ is $\lambda.\ \Delta t + O(\Delta t)$, and the probability of one departure during $\Delta t$ is

$$n\mu\Delta t + O(\Delta t) \quad \text{if} \quad 0 < n < c \quad \text{and } c\mu\Delta t + O(\Delta t) \quad \text{if } n \geq c.$$

To have zero customer in the system during time $t + \Delta t$, there may be the following mutually exclusive cases:

(i) No customer in the system at time $t$, no arrival in time $\Delta t$
(ii) One customer in the system at time $t$, no arrival in time $\Delta t$, one service in time $\Delta t$, and so on.

Therefore, the probability of no customer in the system at time $(t + \Delta t)$ is given by

$$\begin{aligned} & p_0(t+\Delta t) = p_0(t)(1-\lambda\Delta t) + p_1(t)(1-\lambda\Delta t)\mu\Delta t + \mathrm{O}(\Delta t) \\ \text{or}\quad & p_0(t+\Delta t) - p_0(t) = -\lambda p_0(t)\Delta t + p_1(t)\mu\Delta t + \mathrm{O}(\Delta t) \\ \text{or}\quad & \frac{p_0(t+\Delta t)-p_0(t)}{\Delta t} = -\lambda p_0(t) + \mu p_1(t) + \frac{O(\Delta t)}{\Delta t}. \end{aligned} \tag{15.26}$$

To have $n$ customers in the system at time $t + \Delta t$, there may be the following mutually exclusive cases:

(i) $n$ customers in the system at time $t$, no arrival in time $\Delta t$, no service in time $\Delta t$
(ii) $(n - 1)$ customers in the system at time $t$, one arrival in time $\Delta t$, no service in time $\Delta t$
(iii) $(n + 1)$ customers in the system at time $t$, no arrival in time $\Delta t$, one service in time $\Delta t$, and so on.

Therefore, the probability of $n$ customers in the system at time $(t + \Delta t)$ is given by

$$\begin{aligned} p_n(t+\Delta t) &= p_n(t)(1-\lambda\Delta t)(1-\mu_n\Delta t) + p_{n-1}(t)\lambda\Delta t(1-\mu_{n-1}\Delta t) \\ &\quad + p_{n+1}(t)(1-\lambda\Delta t)\mu_{n+1}\Delta\Delta t + O(\Delta t) \\ \text{or}\quad \frac{p_n(t+\Delta t)-p_n(t)}{\Delta t} &= \lambda p_{n-1}(t) - (\lambda+\mu_n)p_n(t) + \mu_{n+1}p_{n+1}(t) + \frac{O(\Delta t)}{\Delta t}. \end{aligned} \tag{15.27}$$

(a) When $0 < n < c$, then $\mu_n = n\mu$, $\mu_{n+1} = (n+1)\mu$.

Hence, from (15.27), we have

$$\frac{p_n(t+\Delta t)-p_n(t)}{\Delta t} = \lambda p_{n-1}(t) - (\lambda+n\mu)p_n(t) + (n+1)\mu p_{n+1}(t) + \frac{O(\Delta t)}{\Delta t}. \tag{15.28}$$

(b) When $n \geq c$, then $\mu_n = c\mu$, $\mu_{n+1} = c\mu$.

Hence, from (15.27), we have

$$\frac{p_n(t+\Delta t) - p_n(t)}{\Delta t} = \lambda p_{n-1}(t) - (\lambda + c\mu)p_n(t) + c\mu p_{n+1}(t) + \frac{O(\Delta t)}{\Delta t}. \qquad (15.29)$$

Now taking the limit as $\Delta t \to 0$, the differential difference equations for this system obtained from (15.26), (15.28) and (15.29) are as follows:

$$\begin{aligned} p_0'(t) &= -\lambda p_0(t) + \mu p_1(t) \quad \text{for } n = 0 \\ p_n'(t) &= \lambda p_{n-1}(t) - (\lambda + n\mu)p_n(t) + (n+1)\mu p_{n+1}(t), \text{ for } 0 < n < c \\ p_n'(t) &= \lambda p_{n-1}(t) - (\lambda + c\mu)p_n(t) + c\mu p_{n+1}(t), \text{ for } n \geq c. \end{aligned}$$

Under the steady state condition of the system, i.e. $\lim_{t\to\infty} p_n(t) = p_n$ and $\lim_{t\to\infty} p_n'(t) = 0$, the preceding three equations reduce to

$$-\lambda p_0 + \mu p_1 = 0 \text{ for } n = 0 \qquad (15.30)$$

$$\lambda p_{n-1} - (\lambda + n\mu)p_n + (n+1)\mu p_{n+1} = 0 \quad \text{for} \quad 0 < n < c \qquad (15.31)$$

$$\lambda p_{n-1} - (\lambda + c\mu)p_n + c\mu p_{n+1} = 0 \quad \text{for} \quad n \geq c, \qquad (15.32)$$

which are the steady state differential difference equations of the system.

**Solution**
From (15.30),

$$-\lambda p_0 + \mu p_1 = 0 \text{ or, } \mu p_1 = \lambda p_0 \text{ or, } p_1 = \frac{\lambda}{\mu} p_0$$

Now setting $n = 1$ in (15.31), we have

$$\begin{aligned} &\lambda p_0 - (\lambda + \mu)p_1 + 2\mu p_2 = 0 \\ &\text{or } p_2 = \frac{\lambda + \mu}{2\mu} p_1 - \frac{\lambda}{2\mu} p_0 \\ &\qquad = \frac{\lambda}{2\mu} p_1 = \frac{1}{(2)!} \left(\frac{\lambda}{\mu}\right)^2 p_0 \left[\because p_1 = \frac{\lambda}{\mu} p_0\right]. \\ &\therefore\ p_2 = \frac{1}{(2)!} \left(\frac{\lambda}{\mu}\right)^2 p_0. \end{aligned}$$

Again setting $n = 2$ in (15.31), we have

$$\lambda p_1 - (\lambda + 2\mu)p_2 + 3\mu p_3 = 0$$
$$\text{or,}\quad p_3 = \frac{\lambda + 2\mu}{3\mu}p_2 - \frac{\lambda}{3\mu}p_1$$
$$= \frac{1}{(3)!}\left(\frac{\lambda}{\mu}\right)^3 p_0.$$

Hence, we have $p_n = \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n p_0$ for $1 \le n < c$.

$$\therefore\ p_{c-1} = \frac{1}{(c-1)!}\left(\frac{\lambda}{\mu}\right)^{c-1} p_0 = \frac{1}{c-1}\frac{\lambda}{\mu}p_{c-2}. \tag{15.33}$$

Now setting $n = c - 1$ in (15.31), we have

$$\lambda p_{c-2} - [\lambda + (c-1)\mu]p_{c-1} + c\mu p_c = 0$$
$$\text{or}\quad p_c = \frac{\lambda + (c-1)\mu}{c\mu}p_{c-1} - \frac{\lambda}{c\mu}p_{c-2}$$
$$= \frac{\lambda}{c\mu}p_{c-1} + \frac{c-1}{c}p_{c-1} - \frac{c-1}{c}p_{c-1}\ \text{[by (15.8)]}$$
$$= \frac{1}{(c)!}\left(\frac{\lambda}{\mu}\right)^c p_0 = \frac{(c\rho)^c}{(c)!}p_0 \quad \text{where}\ \ \rho = \frac{\lambda}{c\mu}.$$

Setting $n = c$ in (15.32), we have

$$\lambda p_{c-1} - (\lambda + c\mu)p_c + c\mu p_{c+1} = 0$$
$$\text{or}\quad p_{c+1} = \frac{\lambda + c\mu}{c\mu}p_c - \frac{\lambda}{c\mu}p_{c-1}$$
$$= \frac{\lambda}{c\mu}p_c + p_c - p_c \quad \left[\because\ p_c = \frac{\lambda}{c\mu}p_{c-1}\right]$$
$$= \frac{1}{c(c)!}\left(\frac{\lambda}{\mu}\right)^{c+1} p_0.$$

Similarly, $p_{c+2} = \frac{1}{c^2(c)!}\left(\frac{\lambda}{\mu}\right)^{c+2} p_0$.

In general,

$$p_n = \frac{1}{c^{n-c}(c)!}\left(\frac{\lambda}{\mu}\right)^n p_0 \ \ \text{for}\ \ n \ge c$$
$$p_n = \frac{c^c}{c^n(c)!}\left(\frac{\lambda}{\mu}\right)^n p_0 \ \ \text{for}\ \ n \ge c.$$

Therefore,

$$\begin{aligned} p_n &= \frac{1}{(n)!}\left(\frac{\lambda}{\mu}\right)^n p_0 \quad \text{for} \quad 1 \le n < c \\ &= \frac{c^c}{c^n (c)!}\left(\frac{\lambda}{\mu}\right)^n p_0 \quad \text{for} \quad n \ge c. \end{aligned}$$

***Determination of*** $p_0$:

We have $\sum\limits_{n=0}^{\infty} p_n = 1$

$$\begin{aligned} &\text{or} \quad p_0 + \sum_{n=1}^{c-1} p_n + \sum_{n=c}^{\infty} p_n = 1 \\ &\text{or} \quad p_0 + \sum_{n=1}^{c-1} \tfrac{1}{(n)!}\left(\tfrac{\lambda}{\mu}\right)^n p_0 + \sum_{n=c}^{\infty} \tfrac{c^c}{c^{n-a}(c)!}\left(\tfrac{\lambda}{\mu}\right)^n p_0 = 1 \\ &\text{or} \quad p_0\left[\sum_{n=0}^{c-1} \tfrac{1}{(n)!}\left(\tfrac{\lambda}{\mu}\right)^n + \tfrac{c^c}{(c)!}\sum_{n=c}^{\infty}\left(\tfrac{\lambda}{c\mu}\right)^n\right] = 1 \qquad (15.34) \\ &\text{or} \quad p_0\left[\sum_{n=0}^{c-1} \tfrac{1}{(n)!}(c\rho)^n + \tfrac{c^c}{(c)!}\sum_{n=c}^{\infty} \rho^n\right] = 1 \quad \text{where} \quad \rho = \tfrac{\lambda}{c\mu} \\ &\text{or} \quad p_0 = \frac{1}{\left[\sum_{n=0}^{c-1} \frac{(c\rho)^n}{(n)!} + \frac{(c\rho)^c}{(c)!}\frac{1}{1-\rho}\right]}. \end{aligned}$$

Hence,

$$p_n = \begin{cases} \frac{1}{(n)!}\left(\frac{\lambda}{\mu}\right)^n p_0 = \frac{(c\rho)^n}{(n)!} p_0 & \text{for} \quad 1 \le n < c \quad \text{where} \quad \rho = \frac{\lambda}{c\mu} \\ \frac{c^c}{c^n (c)!}\left(\frac{\lambda}{\mu}\right)^n p_0 = \frac{c^c}{(c)!}(\rho)^n p_0 & \text{for} \quad n \ge c \quad \text{where} \quad \rho = \frac{\lambda}{c\mu} \end{cases} \qquad (15.35)$$

where $p_0$ is given by (15.34).

**Observation**:

The preceding result is valid only if $\rho = \frac{\lambda}{c\mu} < 1$; i.e. the mean arrival rate must be less than the mean maximum potential service rate of the system.

**Special case**:

When $c$ = 1, Eqs. (15.34) and (15.35) reduce to $p_n = \rho^n p_0$ and $p_0 = 1 - \rho$, which are the steady state solutions of the queueing model $(M/M/1)$:$(\infty/\text{FCFS}/\infty)$.

**Characteristics of the Model**

(i) *Expected queue length (average number of customers in the queue)*

If $n > c$, a queue of $n$ customers would consist of $c$ customers being served together with a genuine queue of $(n-c)$ waiting customers. Hence,

$$\begin{aligned}
L_{\mathrm{q}} &= \sum_{n=c+1}^{\infty} (n-c)p_n \\
&= \sum_{n=c+1}^{\infty} (n-c)\frac{c^c\rho^n}{(c)!}p_0 \quad \left[\because\ p_n = \frac{c^c}{(c)!}\rho^n p_0 \ \text{ where } \ \rho = \frac{\lambda}{c\mu}\right] \\
&= \frac{c^c\rho^c}{(c)!}p_0 \sum_{n=c+1}^{\infty} (n-c)\rho^{n-c} \\
&= \frac{(c\rho)^c}{(c)!}p_0\left[\rho + 2\rho^2 + 3\rho^3 + \cdots\right] \\
&= p_c\frac{\rho}{(1-\rho)^2} \\
\therefore\ L_{\mathrm{q}} &= p_c\frac{\rho}{(1-\rho)^2}.
\end{aligned}$$

**Particular case**: When $c = 1$; i.e. for a system with one channel only,

$$L_{\mathrm{q}} = \frac{\rho}{(1-\rho)^2}p_1 = \frac{\rho}{(1-\rho)^2}\rho(1-\rho) = \frac{\rho^2}{1-\rho}.$$

(ii) *Average number of customers in the system*

We know that $L_{\mathrm{s}} = L_{\mathrm{q}} + \frac{\lambda}{\mu}$.

$$\therefore\ L_{\mathrm{s}} = \frac{\rho p_c}{(1-\rho)^2} + \frac{\lambda}{\mu} = \frac{\rho p_c}{(1-\rho)^2} + c\rho \quad \text{where } \ \rho = \frac{\lambda}{c\mu}$$

(iii) *Average waiting time*

The average waiting time in the system is

$$\begin{aligned}
W_{\mathrm{s}} &= \frac{\text{Average number of customers in the system}}{\text{rate of arrival}} \\
&= \frac{L_{\mathrm{s}}}{\lambda} = \frac{\rho p_c/(1-\rho)^2 + c\rho}{\lambda} = \frac{\rho p_c}{\lambda(1-\rho)^2} + \frac{c\rho}{\lambda}.
\end{aligned}$$

The average waiting time in the queue is

$$W_q = \frac{\text{Average number of customers in the queue}}{\text{rate of arrival}}$$
$$= \frac{L_q}{\lambda} = \frac{\rho p_c}{\lambda(1-\rho)^2}.$$

(iv) *Expected length of non-empty queue, i.e. to find L/L > 0*

$$(L/L > 0) = \frac{L_q}{\text{prob[an arrival has to wait]}}$$
$$= \frac{L_q}{\text{prob}(n > c)} = \frac{\rho p_c/(1-\rho)^2}{\sum_{n=c+1}^{\infty} p_n} = \frac{\rho p_c/(1-\rho)^2}{\sum_{n=c+1}^{\infty} \frac{c^c}{\underline{|c}} \rho^n p_0}$$
$$= \frac{\rho p_c/(1-\rho)^2}{\frac{(c\rho)^2}{\underline{|c}} p_0 \sum_{n=c+1}^{\infty} \rho^{n-c}} = \frac{\rho p_c/(1-\rho)^2}{p_c(\rho+\rho^2+\rho^3+\cdots)} = \frac{\rho/(1-\rho)^2}{\rho(1+\rho+\rho^2+\cdots)}$$
$$= \frac{\rho/(1-\rho)^2}{\frac{\rho}{1-\rho}} = \frac{1}{1-\rho}.$$

(v) *Probability that a customer has to wait*

Probability that a customer has to wait

$$= \text{prob}(n \geq c) = \sum_{n=c}^{\infty} p_n = \sum_{n=c}^{\infty} \frac{c^c}{\underline{|c}} \rho^n p_0 = \frac{c^c}{\underline{|c}} p_0 \sum_{n=c}^{\infty} \rho^n$$
$$= \frac{c^c}{\underline{|c}} p_0 \left(\rho^c + \rho^{c+1} + \cdots\right)$$
$$= \frac{c^c}{\underline{|c}} p_0 \rho^c (1+\rho+\rho^2+\cdots) = \frac{(\rho c)^c}{\underline{|c}} p_0 . \frac{1}{1-\rho} = p_c . \frac{1}{1-\rho} = \frac{p_c}{1-\rho}$$

since $p_c = \frac{(\rho c)^c}{\underline{|c}} p_0$.

(vi) *Probability that on arrival a customer will not have to wait*

$= 1 - \text{prob}(n \geq c) = 1 - \frac{p_c}{1-\rho}$.

Alternatively, we can find this probability using the formula prob $(n < c) = \sum_{n=0}^{c-1} p_n$.

The probability that all channels will be occupied

$$=\text{prob}(n \geq c) = \frac{p_c}{1-\rho}.$$

*Example 4* A telephone exchange has two long-distance operators. The telephone company finds that, during the peak load, long-distance calls arrive in a Poisson fashion at an average rate of 15 per hour. The length of service on these calls is approximately exponentially distributed with mean length 5 min.

(a) What is the probability that a subscriber will have to wait for this long-distance call during the peak hours of the day?
(b) If the subscriber waits and is serviced in turn, what is the expected waiting time?

**Solution**
This problem is of the model (*M*/*M*/*c*):(∞/FCFS/∞).

Here $c = 2$ and $\lambda = 15$ calls/h $= \frac{15}{60} = \frac{1}{4}$calls/min, $\mu = \frac{1}{5}$ and $\rho = \frac{\lambda}{c\mu} = \frac{\frac{1}{4}}{2.\frac{1}{5}} = \frac{5}{8}$ and $c\rho = 2 \cdot \frac{5}{8} = \frac{5}{4}$.

Now, $p_0 = \frac{1}{\sum_{n=0}^{c-1} \frac{(c\rho)^n}{\lfloor n} + \frac{(c\rho)^c}{\lfloor c(1-\rho)}} = \frac{1}{\sum_{n=0}^{1} \frac{\left(\frac{5}{4}\right)^n}{\lfloor n} + \frac{\left(\frac{5}{4}\right)^n}{\lfloor 2\left(1-\frac{5}{8}\right)}} = \frac{1}{1+\frac{\frac{5}{4}}{\lfloor 1} + \frac{\frac{25}{16}}{2.\frac{3}{8}}} = \frac{3}{13}.$

(a) The probability that a subscriber will have to wait for his long-distance call

$$= \text{prob}(n \geq c) = \frac{p_c}{1-\rho} = \frac{\frac{(c\rho)^c}{\lfloor c} p_0}{1-\rho} = \frac{(c\rho)^c p_0}{\lfloor c(1-\rho)} = \frac{\left(\frac{5}{4}\right)^2 \cdot \frac{3}{13}}{\lfloor 2\left(-\frac{5}{8}\right)} = \frac{25}{52} = 0.48.$$

(b) The expected waiting time

$$= W_q = \frac{L_q}{\lambda} = \frac{\rho p_c}{\lambda(1-\rho)^2} = \frac{\rho}{\lambda(1-\rho)^2} \cdot \frac{(c\rho)^2}{\lfloor c} p_0 = \frac{\frac{5}{8}}{\frac{1}{4}\left(1-\frac{5}{8}\right)^2} \cdot \frac{\left(\frac{5}{4}\right)^2}{\lfloor 2} \cdot \frac{3}{13}$$

$$= \frac{\frac{5}{8}}{\frac{1}{4} \cdot \left(\frac{3}{8}\right)^2} \cdot \frac{\frac{25}{16}}{2} \cdot \frac{3}{13} = \frac{125}{39} = 3.2 \text{ min}.$$

*Example 5* A supermarket has two girls ringing up sales at the counters. If the service time for each customer is exponential with mean 4 min, and if the people arrive in a Poisson fashion at the counter at the rate 10 per hour, then:
(a) What is the probability of having to wait for service?
(b) What is the expected percentage of idle time for each girl?
(c) If a customer has to wait, what is the expected length of his waiting time?

**Solution**
This problem is of the model (*M/M/c*):(∞/FCFS/∞).
Here $c = 2$ $\lambda = \frac{10}{60} = \frac{1}{6}$ people per minute, $\mu = \frac{1}{4}$ people per minute.

$$\therefore\ \rho = \frac{\lambda}{c\mu} = \frac{\frac{1}{6}}{2.\frac{1}{4}} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{2}{6} = \frac{1}{3}$$
$$\therefore\ c\rho = 2\cdot\frac{1}{3} = \frac{2}{3}$$

Now, $p_0 = \frac{1}{\sum_{n=0}^{c-1}\frac{(c\rho)^2}{\lfloor n} + \frac{(c\rho)^2}{\lfloor c(1-\rho)}} = \frac{1}{\sum_{n=0}^{1}\frac{\left(\frac{2}{3}\right)^n}{\lfloor n} + \frac{\left(\frac{2}{3}\right)^2}{\lfloor 2\left(1-\frac{1}{3}\right)}} = \frac{1}{1+\frac{\frac{2}{3}}{\lfloor 1} + \frac{\left(\frac{2}{3}\right)^2}{2\cdot\frac{2}{3}}} = \frac{1}{2}.$

(a) The probability of having to wait for service

$= \text{prob}(n \geq 2) = \sum_{n=2}^{\infty} p_n = \frac{p_c}{1-\rho} = \frac{\frac{(c\rho)^c}{\lfloor c}p_0}{1-\rho} = \frac{\frac{\left(\frac{2}{3}\right)^2}{\lfloor 2}.\frac{1}{2}}{1-\frac{1}{3}} = 0.167.$

(b) The expected number of girls who are idle

$= 2\cdot p_0 + 1\cdot p_1 = 2\cdot\frac{1}{2} + 1\cdot\frac{1}{3} = 1 + \frac{1}{3} = \frac{4}{3}\left[p_1 = \frac{\lambda}{\mu}p_0 = \frac{\frac{1}{6}}{\frac{1}{4}}\cdot\frac{1}{2} = \frac{1}{3}\right].$

Now the probability of any girl being idle

$= \frac{\text{expected number of idle girls}}{\text{total number of girls}} = \frac{\frac{4}{3}}{2} = \frac{2}{3} = 0.67.$

Hence, the expected percentage of idle time for each girl = 0.67 × 100% = 67%.

(c) The expected length of the customer waiting time on the condition that the customer has to wait

$$= W/W > 0 = \frac{1}{\text{Mean service rate} - \text{mean arrival rate for } n \geq c} = \frac{1}{c\mu - \lambda}$$
$$= \frac{1}{2\mu - \lambda} = 3 \text{ min.}$$

## 15.3.5 *Model (*M/M/C*):(*N*/FCFS/∞)*

**Assumptions** In this model, we shall discuss a multiserver queueing system under the following assumptions:

(i) Customers arrive in a Poisson fashion.
(ii) The service rate follows a Poisson distribution.
(iii) This queueing system has $c(>1)$ service channels.
(iv) The maximum number of customers in the system is limited to $N$, where $N > c$; the capacity of the system is finite.
(v) The service discipline is first come first served.

### *Model Formulation*

Let $n$ be the number of customers in the system and $\lambda$, $\mu$ be the mean arrival rate and mean service rate of units respectively.

In this model,

$$\lambda_n = \begin{cases} \lambda & \text{when } 0 \le n < N \\ 0 & \text{when } n \ge N \end{cases}$$

and

$$\mu_n = \begin{cases} n\mu & \text{when } 0 \le n < c \\ c\mu & \text{when } c \le n \le N \end{cases}$$

Hence, the probability of one arrival during time $\Delta t$ is $\lambda_n \Delta t + O(\Delta t)$, and the probability of one departure or service during time $\Delta t$ is $\mu_n \Delta t + O(\Delta t)$.

To have zero customer in the system at time $t + \Delta t$, there may be the following mutually exclusive cases:

(i) No customer in the system at time $t$, no arrival during time $\Delta t$
(ii) One customer in the system at time $t$, no arrival during time $\Delta t$, one service during time $\Delta t$,

and so on.
Therefore, the probability of no customer in the system at time $t + \Delta t$ is given by

$$p_0(t+\Delta t) = p_0(t)(1 - \lambda_0 \Delta t) + p_1(t)(1 - \lambda_1 \Delta t)\mu_1 \Delta t + O(\Delta t)$$

$$\text{or } \; p_0(t+\Delta t) - p_0(t) = -\lambda p_0(t)\Delta t + \mu p_1(t)\Delta t + O(\Delta t) \left[\begin{array}{l} \text{Here } \lambda_0 = \lambda \\ \lambda_1 = \lambda, \; \mu_1 = \mu \end{array}\right.$$

$$\text{or } \; \frac{p_0(t+\Delta t) - p_0(t)}{\Delta t} = -\lambda p_0(t) + \mu p_1(t) + \frac{O(\Delta t)}{\Delta t} \; \text{ for } \; n = 0.$$

Taking the limit as $\Delta t \to 0$, we have

$$p_0'(t) = -\lambda p_0(t) + \mu p_1(t) \; \text{ for } \; n = 0. \tag{15.36}$$

To have $n$ customers in the system at time $t + \Delta t$, there may be the following three cases:

(i) $n$ customers in the system at time $t$, no arrival during time $\Delta t$, no service during time $\Delta t$
(ii) $(n - 1)$ customers in the system at time $t$, one arrival during time $\Delta t$, no service during time $\Delta t$
(iii) $(n + 1)$ customers in the system at time $t$, no arrival during time $\Delta t$, one service during time $\Delta t$, and so on.

Therefore, the probability of $n$ customers in the system at time $(t + \Delta t)$ is given by

$$\begin{aligned} p_n(t+\Delta t) &= p_n(t)(1-\lambda_n\Delta t)(1-\mu_n\Delta t)+p_{n-1}(t)\lambda_{n-1}\Delta t(1-\mu_{n-1}\Delta t)\\ &\quad +p_{n+1}(t)(1-\lambda_{n+1}\Delta t)\mu_{n+1}\Delta t+O(\Delta t)\\ \text{or}\ \ p_n(t+\Delta t)-p_n(t) &= -(\lambda_n+\mu_n)p_n(t)\Delta t+\lambda_{n-1}p_{n-1}(t)\Delta t+\mu_{n-1}p_{n+1}(t)\Delta t+O(\Delta t)\\ \text{or}\ \ \frac{p_n(t+\Delta t)-p_n(t)}{\Delta t} &= -(\lambda_n+\mu_n)p_n(t)+\lambda_{n-1}p_{n-1}(t)+\mu_{n+1}p_{n+1}(t)+\frac{O(\Delta t)}{\Delta t}. \end{aligned} \tag{15.37}$$

(a) For $0 < n < c$,

$$\lambda_{n-1}=\lambda_n=\lambda,\quad \mu_n=n\mu,\ \ \mu_{n+1}=(n+1)\mu.$$

Hence, from (15.37), we have

$$\frac{p_n(t+\Delta t)-p_n(t)}{\Delta t}=-(\lambda+n\mu)p_n(t)+\lambda p_{n-1}(t)+(n+1)\mu p_{n+1}(t)+\frac{O(\Delta t)}{\Delta t}.$$

Taking the limit as $\Delta t \to 0$ of both sides, we have

$$p_n'(t)=-(\lambda+n\mu)p_n(t)+\lambda p_{n-1}(t)+(n+1)\mu p_{n+1}(t). \tag{15.38}$$

(b) For $c \le n < N$,

$$\lambda_{n-1}=\lambda, \lambda_n=\lambda,\ \mu_n=c\mu,\ \mu_{n+1}=c\mu.$$

Hence, from (15.37), we have

$$\frac{p_n(t+\Delta t)-p_n(t)}{\Delta t}=-(\lambda+c\mu)p_n(t)+\lambda p_{n-1}(t)+c\mu p_{n+1}(t)+\frac{O(\Delta t)}{\Delta t}.$$

Taking the limit as $\Delta t \to 0$ of both sides, we have

$$p_n'(t) = -(\lambda + c\mu)p_n(t) + \lambda p_{n-1}(t) + c\mu p_{n+1}(t) \quad \text{for} \quad c \le n < N \tag{15.39}$$

(c) For $n = N$, $\lambda_n = \lambda_N = 0$, $\lambda_{n-1} = \lambda_{N-1} = \lambda\, \mu_n = c_N = c\mu$, $p_{n+1}(t) = 0$ .

Hence, from (15.37), we have

$$\frac{p_n(t+\Delta t) - p_n(t)}{\Delta t} = -(o + c\mu)p_n(t) + \lambda p_{n-1}(t) + \frac{O(\Delta t)}{\Delta t}.$$

Taking the limit as $\Delta t \to 0$ of both sides, we have

$$\begin{aligned} p_n'(t) &= -c\mu p_n(t) + \lambda p_{n-1}(t) \\ p_N'(t) &= -c\mu p_N(t) + \lambda p_{N-1}(t) \quad \text{for} \quad n = N \end{aligned} \tag{15.40}$$

Under a steady state condition of the system, i.e. $\lim_{t\to\infty} p_n(t) = p_n$ and $\lim_{t\to\infty} p_n'(t) = 0$, Eqs. (15.36), (15.38)–(15.40) reduce to:

$$-\lambda p_0 + \mu p_1 = 0 \quad \text{for} \quad n = 0 \tag{15.41}$$

$$\lambda p_{n-1} - (\lambda + n\mu)p_n + (n+1)\mu p_{n+1} = 0 \quad \text{for} \quad 0 < n < c \tag{15.42}$$

$$\lambda p_{n-1} - (\lambda + c\mu)p_n + c\mu p_{n+1} = 0 \quad \text{for} \quad c \le n < N \tag{15.43}$$

$$\lambda p_{N-1} - c\mu p_N = 0 \quad \text{for} \quad n = N. \tag{15.44}$$

***Solution of the Model***

$$p_n = \begin{cases} \frac{1}{n!}(c\rho)^n p_0, & \text{for} \quad 0 \le n < c \\ \frac{c^c \rho^n}{c!} p_0, & \text{for} \quad c \le n \le N \end{cases},$$

where

$$p_0 = \begin{cases} \left[ \sum_{n=0}^{c-1} \frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!} \frac{1-\rho^{N-c+1}}{1-\rho} \right]^{-1} & \text{for} \quad \rho = \frac{\lambda}{c\mu} \ne 1 \\ \left[ \sum_{n=0}^{c-1} \frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!} (N - c + 1) \right]^{-1} & \text{for} \quad \rho = \frac{\lambda}{c\mu} = 1 \end{cases}.$$

***Characteristics of the Model***

(i) ***Average queue length***

The average queue length is

$$
\begin{aligned}
L_q &= \sum_{n=c+1}^{N}(n-c)p_n = \sum_{n=c}^{N}(n-c)p_n = \sum_{n=c}^{N}(n-c)\frac{c^c\rho^n}{\underline{|c}}p_0 \\
&= \frac{c^c}{\underline{|c}}p_0\sum_{n=c}^{N}(n-c)\rho^n = \frac{c^c}{\underline{|c}}p_0\rho^c\sum_{n=c}^{N}(n-c)\rho^{n-c} \\
&= \frac{(c\rho)^c}{\underline{|c}}p_0\sum_{x=0}^{N-c}x\rho^x \text{ putting } n-c=x \\
&= \frac{(c\rho)^c}{\underline{|c}}p_0\sum_{n=0}^{N-c}\rho\frac{\mathrm{d}}{\mathrm{d}\rho}(\rho^x) \\
&= \frac{(c\rho)^c}{\underline{|c}}p_0\rho\frac{\mathrm{d}}{\mathrm{d}\rho}\left\{\sum_{x=0}^{N-c}\rho^x\right\} \\
&= \frac{(c\rho)^c}{\underline{|c}}p_0\rho\frac{\mathrm{d}}{\mathrm{d}\rho}\left[\frac{1-\rho^{N-c+1}}{1-\rho}\right] \\
&= \begin{cases} \dfrac{(c\rho)^c}{\underline{|c}}p_0\rho\dfrac{1-\rho^{N-c+1}-(1-\rho)(N-c+1)\rho^{N-c}}{(1-\rho)^2} & \text{for } \rho=\dfrac{\lambda}{c\mu}\neq 1 \\[2ex] p_0\dfrac{c^c(n-c)(N-c+1)}{L\underline{|c}} & \text{for } \lambda=\dfrac{\lambda}{c\mu}=1 \end{cases}
\end{aligned}
$$

(ii) ***Average number of customers in the system***

$L_s$ = average number of customers in the system

$$
\begin{aligned}
&= E(n) = \sum_{n=0}^{N}np_n = \sum_{n=0}^{c-1}np_n + \sum_{n=c}^{N}np_n \\
&= \sum_{n=c}^{N}(n-c)p_n + \sum_{n=c}^{N}cp_n + \sum_{n=0}^{c-1}np_n \\
&= L_q + c\left[\sum_{n=0}^{N}p_n - \sum_{n=0}^{c-1}p_n\right] + \sum_{n=0}^{c-1}np_n \\
&= L_q + c - c\sum_{n=0}^{c-1}p_n + \sum_{n=0}^{c-1}np_n = L_q + c + \sum_{n=0}^{c-1}(n-c)p_n \\
&= L_q + c - p_0\sum_{n=0}^{c-1}\frac{(c-n)(\rho c)^n}{\underline{|n}}
\end{aligned}
$$

***Calculation of $\lambda_{\mathbf{eff}}$***: $\lambda_{eff} = \lambda(1-p_N)$

Let $\bar{c}$ = expected number of idle servers

$$\therefore\ \bar{c} = \sum_{n=0}^{c} (c-n)p_n$$

$$\therefore\ c - \bar{c} = \text{the expected number of busy servers.}$$

$$\therefore\ \lambda_{\text{eff}} = \mu(c-\bar{c})$$

$$W_{\text{s}} = \frac{\lambda s}{\lambda_{\text{eff}}} \quad \text{and} \quad W_{\text{q}} = \frac{L_{\text{q}}}{\lambda_{\text{eff}}}$$

Now, $W_{\text{s}} = \frac{\lambda s}{\lambda_{\text{eff}}}$ and $W_{\text{q}} = \frac{L_{\text{q}}}{\lambda_{\text{eff}}}$.

**Special cases**:

(i) Let us take $c = 1,\ N \to \infty$ and consider $\rho < 1$. Then we get

$$p_0 = 1 - \rho$$

and $p_n = \rho^n p_0$, $n = 0, 1, 2, \ldots$

This result is exactly the same as that of model $(M/M/1){:}(\infty/\text{FCFS}/\infty)$.

(ii) Let us take $c$ = 1. Then we get

$$p_0 = \begin{cases} \frac{1-\rho}{1-\rho^{N+1}}, & \rho = \frac{\lambda}{c\mu} \neq 1 \\ \frac{1}{1+N}, & \rho = \frac{\lambda}{c\mu} = 1 \end{cases}$$

and $p_n = \rho^n p_0,\ \ 1 \le n \le N$.

This result corresponds to that of the queueing model $(M/M/1){:}(N/\text{FCFS}/\infty)$.

(iii) Let us take $N \to \infty$ and consider $\rho = \frac{\lambda}{c\mu} < 1$.

Then we get

$$p_0 = \left[\sum_{n=0}^{c-1} \frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!} \frac{1-\rho^{N-c+1}}{1-\rho}\right]^{-1}$$

and $p_n = \begin{cases} \dfrac{1}{n!}(c\rho)^n p_0, & 0 \le n < c \\ \dfrac{1}{c!} c^c \rho^n p_0, & c \le n \le N \end{cases},$

which is the steady state solution of the queueing model $(M/M/c){:}(N/\text{FCFS}/\infty)$.

Thus, we can consider this model as a general queueing model, as we can get the results of previous queueing models from this one.

**Machine Repairing Problem**

Whenever a machine breaks down, it will result in a great loss to an organization. Thus, machine repair problems are very important problems in queueing theory. When a machine breaks down, at that point the machine starts to be repaired. During this time, if another machine breaks down, then it will be attended after completion of the repair of the first machine. Thus, the broken-down machines form a queue and wait for their repair. There are various situations. There may be one or more than one mechanic. If there is one mechanic, we have a problem of single channel; if there is more than one mechanic, we have a multichannel problem. The machines may be repaired in a single phase or in $k$ phases.

### 15.3.6 *Model (***M/M/R***):(***K/***GD)* K > R

**Assumptions** In this model, we shall discuss a queueing system under the following assumptions:

(i) This is a queueing system in which there are $R$ service channels.
(ii) The maximum limit imposed on incoming customers, say $k$, is fixed.
(iii) The arrivals (breakdowns) and services are assumed to follow a Poisson distribution with parameters $\lambda$ and $\mu$ respectively.
(iv) This system has a finite calling source.

***Model Formulation***
Let $n$ be the number of machines in a breakdown situation.
Hence, the approximate probability of a single service during the time $\Delta t$ is $\mu_n \Delta t$, where

$$\mu_n = \begin{cases} n\mu & \text{for } n \le R \\ R\mu & \text{for } R \le n \le k \\ 0 & \text{for } n > k \end{cases}.$$

If $\lambda$ is the rate of breakdown per machine, then the probability of a single arrival during the time $\Delta t$ (when there are $n$ machines in a breakdown situation) is approximately $\lambda_n \Delta t$ for $n \le k$, where

$$\lambda_n = \begin{cases} (k-n)\lambda \text{ for } 0 \le n < k \\ 0 \quad \text{for } n \ge k \end{cases}.$$

To have 0 customer in the system at time $t + \Delta t$, there may be the following three cases:

(i) No customer in the system at time $t$, no arrival in time $\Delta t$

(ii) One customer in the system at time $t$, no arrival in time $\Delta t$, one service in time $\Delta t$

(iii) No customer in the system at time $t$, one arrival in time $\Delta t$, one service in time $\Delta t$, and so on.

Therefore, the probability of no customers in the system at time $t + \Delta t$ is given by

$$\begin{aligned} p_0(t+\Delta t) &= p_0(t)(1-\lambda_0\Delta t)+p_1(t)(1-\lambda_1\Delta t)\mu_1\Delta t+O(\Delta t)\\ &= p_0(t)(1-k\lambda\Delta t)+p_1(t)\{1-(k-1)\lambda\Delta t\}\mu\Delta t+O(\Delta t)\\ \text{or}\quad \frac{p_0(t+\Delta t)-p_0(t)}{\Delta t} &= -k\lambda p_0(t)+\mu p_1(t)+\frac{O(\Delta t)}{\Delta t}. \end{aligned} \tag{15.45}$$

To have $n$ (>0) customers in the system at time $t + \Delta t$, there may be the following three cases:

(i) $n$ customers in the system at time $t$, no arrival in time $\Delta t$, no service in time $\Delta t$

(ii) $(n - 1)$ customers in the system at time $t$, one arrival in time $\Delta t$, no service in time $\Delta t$

(iii) $(n + 1)$ customers in the system at time $t$, no arrival in time $\Delta t$, one service in time $\Delta t$, and so on.

Therefore, the probability of $n$ customers in the system at time $(t + \Delta t)$ is given by

$$\begin{aligned} p_n(t+\Delta t) &= p_n(t)(1-\lambda_n\Delta t)(1-\mu_n\Delta t)+p_{n-1}(t)\lambda_{n-1}\Delta t(1-\mu_{n-1}\Delta t\\ &\quad +p_{n+1}(t)(1-\lambda_{n+1}\Delta t)\mu_{n+1}\Delta t+O(\Delta t)\\ \text{or}\quad \frac{p_n(t+\Delta t)-p_n(t)}{\Delta t} &= -(\lambda_n+\mu_n)p_n(t)\\ &\quad +\lambda_{n-1}p_{n-1}(t)+\mu_{n+1}p_{n+1}(t)+\frac{O(\Delta t)}{\Delta t}\\ &\quad [\text{dividing both sides by } \Delta t]. \end{aligned} \tag{15.46}$$

When $0 < n < R$, then

$$\begin{aligned} \lambda_{n-1} &= \{k-(n-1)\}\lambda\\ \lambda_n &= (k-n)\lambda,\ \lambda_{n+1}=\{k-(n+1)\}\lambda\\ \mu_n &= n\mu,\ \mu_{n+1}=(n+1)\mu \end{aligned}$$

Hence, from (15.46), we have

$$\begin{aligned} \frac{p_n(t+\Delta t)-p_n(t)}{\Delta t} &= -\{(k-n)\lambda+n\mu\}p_n(t)+\{k-(n-1)\}\lambda p_{n-1}(t)\\ &\quad +(n+1)\mu p_{n+1}(t)+\frac{O(\Delta t)}{\Delta t} \end{aligned} \tag{15.47}$$

For $R \le n < k$,

$$\lambda_{n-1} = \{k - (n-1)\}\lambda, \lambda_n = (k-n)\lambda, \mu_n = R\mu, \mu_{n+1} = R\mu,$$

Hence, from (15.46), we have

$$\frac{p_n(t+\Delta t) - p_n(t)}{\Delta t} = -\{(k-n)\lambda + R\mu\}p_n(t) + \{k-(n-1)\}\lambda p_{n-1}(t) + R\mu p_{n+1}(t) + \frac{O(\Delta t)}{\Delta t} \quad \text{for } R \le n < k \tag{15.48}$$

For $n = k$,

$$\begin{aligned}\lambda_{n-1} &= \{k-(n-1)\}\lambda = \{k-(k-1)\}\lambda = \lambda[\because n = k]\\ \lambda_n &= (k-n)\lambda = 0\\ \mu_n &= R\mu\end{aligned}$$

Hence, from (15.46), we have

$$\frac{p_n(t+\Delta t) - p_n(t)}{\Delta t} = -(0 + R\mu)p_n(t) + \lambda p_{n-1}(t) + \frac{O(\Delta t)}{\Delta t} \quad \text{for } n = k,$$

$$\text{i.e. } \frac{p_k(t+\Delta t) - p_k(t)}{\Delta t} = \lambda p_{k-1}(t) + R\mu p_k(t) + \frac{O(\Delta t)}{\Delta t}. \tag{15.49}$$

Now taking the limit as $\Delta t \to 0$ in (15.45), (15.47)–(15.49), we have

$$\begin{aligned}&p_n'(t) = -\mu\lambda p_0(t) + \mu p_1(t) \quad \text{for } n = 0\\ &p_n'(t) = -\{(k-n)\lambda + n\mu\}p_n(t) + \{k-(n-1)\}\lambda p_{n-1}(t) + (n+1)\mu p_{n+1}(t) \quad \text{for } 0 < n < R\\ &p_n'(t) = -\{(k-n)\lambda + R\mu\}p_n(t) + \{k-(n-1)\}\lambda p_{n-1}(t) + R\mu p_{n+1}(t) \quad \text{for } R \le n < k\\ &p_n'(t) = \lambda p_{k-1}(t) + R\mu p_k(t) \quad \text{for } n = k.\end{aligned}$$

Under a steady state condition of this system, i.e. $\lim_{t\to\infty} p_n(t) = p_n$ and $\lim_{t\to\infty} p_n'(t) = 0$, the preceding four equations reduce to

$$-k\lambda p_0 + \mu p_1 = 0 \quad \text{for } n = 0 \tag{15.50}$$

$$\{k-(n-1)\}\lambda p_{n-1} - \{(k-n)\lambda + n\mu\}p_n + (n+1)\mu p_{n+1} = 0 \quad \text{for } 0 < n < R \tag{15.51}$$

$$\{k-(n-1)\}\lambda p_{n-1} - \{(k-n)\lambda + R\mu\}p_n + R\mu p_{n+1} = 0 \quad \text{for } R \le n < k \tag{15.52}$$

$$\lambda p_{k-1} - R\mu p_k = 0 \quad \text{for } n = k. \tag{15.53}$$

**Solution**
From (15.50), $\mu p_1 = k\lambda p_0$

$$\text{or } p_1 = k\frac{\lambda}{\mu}p_0 \text{ or } p_1 = k\rho p_0 \quad \text{where } \rho = \frac{\lambda}{\mu}.$$

Now setting $n = 1$ in (15.51), we have

$$\begin{aligned} & k\lambda p_0 - \{(k-1)\lambda + \mu\}p_1 + 2\mu p_2 = 0 \\ \text{or} \quad & 2\mu p_2 = \{(k-1)\lambda + \mu\}p_1 - k\lambda p_0 \\ \text{or} \quad & 2p_2 = \left\{(k-1)\frac{\lambda}{\mu} + 1\right\}p_1 - k\frac{\lambda}{\mu}p_0 \\ & = \{(k-1)\rho + 1\}p_1 - k\rho p_0 \left[\because \ \rho = \frac{\lambda}{\mu}\right] \\ & = (k-1)\rho p_1 + p_1 - p_1 \left[\because \ p_1 = k\rho p_0\right] \\ & = (k-1)\rho p_1 = (k-1)\rho k\rho p_0 = k(k-1)\rho^2 p_0 \\ & p_2 = \frac{k(k-1)}{2}\rho^2 p_0 = \binom{k}{2}\rho^2 p_0. \end{aligned}$$

Again, setting $n = 2$ in (15.51), we have

$$\begin{aligned} 3p_3 &= \frac{(k-2)(k-1)k\rho^3 p_0}{2} \\ \text{or} \quad p_3 &= \frac{(k-2)(k-1)k\rho^3 p_0}{3.2} = \binom{k}{3}\rho^3 p_0. \end{aligned}$$

By induction, $p_n = \binom{k}{n}\rho^n p_0$ for $0 \le n \le R$.

Setting $n = R$ in (15.52), we have

$$\{k-(R-1)\}\lambda p_{R-1} - \{(k-R)\lambda + R\mu\}p_R + R\mu p_{R+1} = 0$$

$$\begin{aligned}
\text{or}\;\; Rp_{R+1} &= \{(k-R)\rho + R\}p_R - (k-R+1)\rho p_{R-1} \\
&= \{(k-R)\rho + R\}\binom{k}{R}\rho^R p_0 - (k-R+1)\rho\binom{k}{R-1}\rho^{R-1}p_0 \\
&= \{(k-R)\rho + R\}\binom{k}{R}\rho^R p_0 - R\binom{k}{R}\rho^R p_0 \\
&= \binom{k}{R}\rho^R p_0[(k-R)\rho + R - R] \\
&= \binom{k}{R}(k-R)\rho^{R+1}p_0 \\
&= \binom{k}{R+1}(R+1)\rho^{R+1}p_0 \\
\therefore\; p_{R+1} &= \binom{k}{R+1}\frac{R+1}{R}\rho^{R+1}p_0
\end{aligned}$$

Again, setting $n = R + 1$ in (15.52), we have

$$\begin{aligned}
&(k-R)\rho p_R - \left\{(k-\overline{R+1})\rho + R\right\}p_{R+1} + Rp_{R+2} = 0 \\
\text{or}\quad Rp_{R+2} &= \{(k-R-1)\rho + R\}p_{R+1} - (k-R)\rho p_R \\
&= (k-R-1)\rho p_{R+1} + Rp_{R+1} - (k-R)\rho p_R \\
\therefore\; p_{R+2} &= \frac{(k-R-1)(R+1)}{R^2}\binom{k}{R+1}\rho^{R+2}p_0 \\
&= \frac{(k-R-1)(R+1)}{R^2}\frac{\underline{|k}}{\underline{|R}+1\underline{|K-R}-1}\rho^{R+2}p_0 \\
&= \frac{(R+1)(R+2)}{R^2\underline{|R}+2}\frac{\underline{|k}}{\underline{|k}-(R+2)}\rho^{R+2}p_0 \\
&= \binom{k}{R+2}\frac{(R+1)(R+2)}{R^2}\rho^{R+2}p_0 = \binom{k}{R+2}\frac{\underline{|R}+2}{\underline{|R}R^2}\rho^{R+2}p_0 \\
\therefore\; p_{R+i} &= \binom{k}{R+i}\frac{\underline{|R}+i}{\underline{|R}R^i}\rho^{R+i}p_0 \;\; \text{for}\;\; R \le R+i < k.
\end{aligned}$$

From (15.53), we have

$$
\begin{aligned}
&R\mu p_k = \lambda p_{k-1} \\
&\text{or } \; Rp_k = \frac{\lambda}{\mu} p_{k-1} = \rho \binom{k}{k-1} \frac{\underline{|k} - 1}{\underline{|R} R^{k-R-1}} \rho^{k-1} p_0 \\
&\binom{k}{1} \frac{\underline{|k} - 1}{\underline{|R} R^{k-R-1}} \rho^k p_0 = \binom{k}{k} \frac{k\underline{|k} - 1}{\underline{|R} R^{k-R-1}} \rho^k p_0 = \binom{k}{k} \frac{\underline{|k}}{\underline{|R} R^{k-R-1}} \rho^k p_0. \\
&\therefore \; p_k = \binom{k}{k} \frac{\underline{|k}}{\underline{|R} R^{k-R}} \rho^k p_0.
\end{aligned}
$$

Thus,

$$
p_n = \begin{cases} \binom{k}{n} \rho^n p_0 & \text{for } \; 0 \le n \le R \\ \binom{k}{n} \frac{\underline{|n}}{\underline{|R} R^{n-R}} \rho^n p_0 & \text{for } \; R \le n \le k \end{cases}.
$$

$$
\begin{aligned}
&\text{Now,} \quad \sum_{n=0}^{n} p_n = 1 \\
&\text{or} \qquad \sum_{n=0}^{R-1} p_n + \sum_{n=R}^{k} p_n = 1 \\
&\text{or} \qquad \sum_{n=0}^{R-1} \binom{k}{n} \rho^n p_0 + \sum_{n=R}^{k} \binom{k}{n} \frac{\underline{|n}}{\underline{|R} R^{n-R}} \rho^n p_0 = 1 \\
&\text{or} \qquad p_0 = \frac{1}{\sum_{n=0}^{R-1} \binom{k}{n} \rho^n + \sum_{n=R}^{n} \binom{k}{n} \frac{\rho^n \underline{|n}}{R^{n-R} \underline{|R}}}.
\end{aligned}
$$

### *Characteristics of the Model*

(i) *Average number of customers in the system*

The average number of customers in the system is given by

$$
\begin{aligned}
L_s = E(n) &= \sum_{n=0}^{k} np_n = \sum_{n=0}^{R-1} np_n + \sum_{n=R}^{k} np_n \\
&= \sum_{n=0}^{R-1} n \binom{k}{n} \rho^n + \frac{1}{\underline{|R}} \sum_{n=R}^{k} n \binom{k}{n} \frac{\rho^n \underline{|n}}{\underline{|R} R^{n-R}} \\
&= p_0 \left[ \sum_{n=0}^{n-1} n \binom{k}{n} \rho^n + \frac{1}{\underline{|R}} \sum_{n=R}^{k} n \binom{k}{n} \frac{\rho^n \underline{|n}}{R^{n-R}} \right]
\end{aligned}
$$

(ii) *Expected queue length*

The expected queue length is given by

$$
\begin{aligned}
L_{\mathrm{q}} &= \sum_{n=R+1}^{k} (n-R)p_n = \sum_{n=R+1}^{k} np_n - \sum_{n=R+1}^{k} np_n \\
&= \sum_{n=0}^{k} np_n - \sum_{n=0}^{R} np_n - R\left\{\sum_{n=0}^{k} p_n - \sum_{n=0}^{R} p_n\right\} \\
&= L_{\mathrm{s}} - \sum_{n=0}^{R} np_n - R\left\{1 - \sum_{n=0}^{R} p_n\right\} \quad \left[\because \sum_{n=0}^{k} p_n = 1\right] \\
&= L_{\mathrm{s}} - R - \sum_{n=0}^{R} np_n + R\sum_{n=0}^{R} p_n \\
&= L_{\mathrm{s}} - R + \sum_{n=0}^{R} (R-n)p_n \\
&= L_{\mathrm{s}} - (R - \overline{R})
\end{aligned}
$$

where $\overline{R} = \sum_{n=0}^{R} (R-n)p_n$ = expected number of idle machine repairmen.

(iii) *Effective arrival rate* ($\lambda_{\mathrm{eff}}$)

In this queueing system, the arrivals occur with a rate $\lambda$, but all arrivals do not join the system. This situation occurs when the maximum allowable queue length is reached. In that case, no new arrivals are allowed to join the queue. So, we shall define λ considering those arrivals which join the system. This arrival rate is known as the effective arrival rate, and it is denoted by $\lambda_{\mathrm{eff}}$.
Hence, the effective arrival rate is given by

$$\lambda_{\mathrm{eff}} = \mu(R - \overline{R})$$

or

$$
\begin{aligned}
\lambda_{\mathrm{eff}} &= \sum_{n=0}^{k} \lambda_n p_n = E[\lambda_n] = E[\lambda(k-n)][\text{Since } \lambda_n = \lambda(k-n)] \\
&= \lambda E[(k-n)] = \lambda[k - E(n)] = \lambda[k - L_{\mathrm{s}}].
\end{aligned}
$$

(iv) *Expected waiting time*

The expected waiting time of a customer in the system is given by

$$W_s = \frac{L_s}{\lambda_{\text{eff}}},$$

whereas the expected waiting time of a customer in the queue is given by

$$W_q = \frac{L_q}{\lambda_{\text{eff}}}.$$

*Example 6* There are 5 machines, each of which, when running, suffers breakdowns at an average rate of 2 per hour. There are two servicemen, and only one man can work on a machine at a time. If $n$ machines are out of order when $n > 2$, then $n - 2$ of them wait until a serviceman is free. Once a serviceman starts work on a machine, the time to complete the repair has an exponential distribution with mean 5 min. Find the distribution of the number of machines out of action at a given time. Also, find the average time an out-of-action machine has to spend waiting for the repairs to start.

**Solution**
Here

$k$ = total number of machines in the system = 5
$R$ = number of servicemen = 2
$\lambda$ = 2 per hour, $\mu$ = 1 per 5 min per hour.

Hence, $\rho = \lambda/\mu = \frac{2}{12} = \frac{1}{6}$.
Let $n$ be the number of machines out of order.
Hence,

$$p_n = \begin{cases} \binom{5}{n}\left(\frac{1}{6}\right)^n p_0 & \text{for } 0 \le n \le 2 \\ \binom{5}{n}\frac{|\underline{n}}{2^{n-2}|\underline{2}}\left(\frac{1}{6}\right)^n p_0 & \text{for } 2 \le n \le 5 \end{cases}$$
$$= \begin{cases} \binom{5}{n}\left(\frac{1}{6}\right)^n p_0 & \text{for } 0 \le n \le 2 \\ \binom{5}{n}2|\underline{n}\left(\frac{1}{12}\right)^n p_0 & \text{for } 2 \le n \le 5 \end{cases},$$

where

$$p_0 = \left[\sum_{n=0}^{2-1}\binom{5}{n}\left(\frac{1}{6}\right)^n + \sum_{n=2}^{5}\binom{5}{n}2|\underline{n}\left(\frac{1}{12}\right)^n\right]^{-1} = \frac{648}{1493}.$$

The average number of machines out of action is given by

$$L_q = \sum_{n=2+1}^{5} (n-2)p_n = \sum_{n=3}^{5} (n-2)p_n = p_3 + 2p_4 + 3p_5 = \frac{165}{1493}.$$

The average time an out-of-action machine has to spend waiting for repairs to start is $W_q = L_q/\lambda_{\text{eff}}$.
But

$$\lambda_{\text{eff}} = \sum_{n=0}^{k} \lambda_n p_n = \sum_{n=0}^{k} \lambda(k-n)p_n = \lambda \sum_{n=0}^{5} (5-n)p_n = \frac{6 \times 2050}{1493}.$$

Hence, $W_q = 55/4100$, h = 33/41 min.

### 15.3.7 Power Supply Model

#### Model Formulation

The power supply problem can be analysed as a queueing problem. The objective of this problem is to ensure the better utilization of a power supply to the customers. Let there be an electrical circuit supplying power to customers. Further, let the requirement and supply schedules follow a Poisson distribution with parameters $\lambda$ and $\mu$ respectively.

Let there be $n$ customers in the queue at any time $t$ and let $C$ be the total number of customers.

Then,

$$\left.\begin{aligned} \lambda_n &= (c-n)\lambda \\ \text{and} \quad \mu_n &= n\mu \end{aligned}\right\} \quad \text{for} \quad 0 \le n \le c.$$

Proceeding similarly as in other models, the probabilities that

(i) there is no customer and (ii) there are $n$ customers in the system at time $t + \Delta t$ are given by

$$p_0(t+\Delta t) = p_0(t)(1-\lambda_0\Delta t) + p_1(t)(1-\lambda_1\Delta t)\mu_1\Delta t + O(\Delta t) \tag{15.54}$$

and

$$\begin{aligned} p_n(t+\Delta t) = {} & p_{n-1}(t)\lambda_{n-1}\Delta t(1-\mu_{n-1}\Delta t) + p_n(t)(1-\lambda_n\Delta t)(1-\mu_n\Delta t) \\ & + p_{n+1}(t)(1-\lambda_{n+1}\Delta t)\mu_{n+1}\Delta t + O(\Delta t). \end{aligned} \tag{15.55}$$

From (15.1), we have

$$\begin{aligned} & p_0(t+\Delta t) = p_0(t)(1-c\lambda\Delta t) + p_1(t)\{1-(c-1)\lambda\Delta t\}\mu\Delta t + O(\Delta t) \\ \text{or} \quad & \frac{p_0(t+\Delta t)-p_0(t)}{\Delta t} = -c\lambda p_0(t) + \mu p_1(t) + \frac{O(\Delta t)}{\Delta t}. \end{aligned} \tag{15.56}$$

From (15.55), we have

$$\frac{p_n(t+\Delta t)-p_n(t)}{\Delta t} = \lambda_{n-1}p_{n-1}(t) - (\lambda_n+\mu_n)p_n(t) + \mu_{n+1}p_{n+1}(t) + \frac{O(\Delta t)}{\Delta t}. \tag{15.57}$$

When $0 < n < c$,

$$\lambda_{n-1} = \{c-(n-1)\}\lambda, \quad \lambda_n = (c-n)\lambda, \mu_n = n\mu, \quad \mu_{n+1} = (n+1)\mu.$$

Hence, from (15.57), we have

$$\begin{aligned} \frac{P_n(t+\Delta t)-p_n(t)}{\Delta t} &= (a-n+1)\lambda p_{n-1}(t) - \{(c-n)\lambda + n\mu\}p_n(t) \\ &\quad + (n+1)\mu p_{n+1}(t) + \frac{O(\Delta t)}{\Delta t}, \ 0 < n < c. \end{aligned} \tag{15.58}$$

When $n = c$, $\lambda_{n-1} = \{c-(c-1)\}\lambda = \lambda$, $\lambda_n = (c-c)\lambda = 0$

$$\mu_n = n\mu = c\mu, \ p_{n+1}(t) = 0.$$

Hence, from (15.57), we have

$$\frac{p_c(t+\Delta t)-p_c(t)}{\Delta t} = \lambda p_{c-1}(t) - c\mu p_c(t) + \frac{O(\Delta t)}{\Delta t}. \tag{15.59}$$

Taking the limit as $\Delta t \to 0$, from (15.56), (15.58) and (15.59), we have

$$p_0'(t) = -c\lambda p_0(t) + \mu p_1(t) \text{ for } n = 0 \tag{15.60}$$

$$\begin{aligned} p_n'(t) &= (c-n+1)\lambda p_{n-1}(t) - \{(c-n)\lambda + n\mu\}p_n(t) + (n+1)\mu p_{n+1}(t) \\ &\quad \text{for } 0 < n < c \end{aligned} \tag{15.61}$$

$$p_c'(t) = \lambda p_{c-1}(t) - c\mu p_c(t) \text{ for } n = c. \tag{15.62}$$

Under a steady state condition of this system, i.e. $\lim_{t\to\infty} p_n(t) = p_n$ and $\lim_{t\to\infty} p_n'(t) = 0$, the preceding three equations reduce to

$$-c\lambda p_0 + \mu p_1 = 0 \quad \text{for } n = 0 \tag{15.63}$$

$$(c-n+1)\lambda p_{n-1} - \{(c-n)\lambda + n\mu\}p_n + (n+1)\mu p_{n+1} = 0, \quad 0<n<c \tag{15.64}$$

$$\lambda p_{c-1} - c\mu p_c = 0 \quad \text{for} \quad n = c. \tag{15.65}$$

Equations (15.63)–(15.65) are the steady state equations of the system.

**Solution of the Model**

From (15.63), we have $p_1 = c\left(\frac{\lambda}{\mu}\right)p_0$.

Now, setting $n = 1$ in (15.64), we have

$$c\lambda p_0 - \{(c-1)\lambda + \mu\}p_1 + 2\mu p_2 = 0$$

$$\text{or} \quad 2\mu p_2 = -c\lambda p_0 + \{(c-1)\lambda + \mu\}c\left(\frac{\lambda}{\mu}\right)p_0$$

$$\text{or} \quad p_2 = \frac{c(c-1)}{\underline{|2}}\left(\frac{\lambda}{\mu}\right)^2 p_0.$$

Again, for $n = 2$,

$$p_3 = \frac{c(c-1)(c-2)}{\underline{|3}}\left(\frac{\lambda}{\mu}\right)^3 p_0.$$

In general,

$$p_n = \frac{c(c-1)\cdots(c-n+1)}{\underline{|n}}\left(\frac{\lambda}{\mu}\right)^n p_0 \quad \text{for } 0<n<c.$$

From (15.65), we have

$$p_c = \frac{1}{c}\frac{\lambda}{\mu}p_{c-1}$$

or $p_c = \frac{1}{c}\frac{\lambda}{\mu}.\frac{c(c-1)\cdots 2}{\underline{|c-1}}\left(\frac{\lambda}{\mu}\right)^{c-1}p_0 = \left(\frac{\lambda}{\mu}\right)^c p_0.$

But

$$\sum_{n=0}^{c} p_n = 1$$

$$\text{or} \quad p_0 + p_1 + p_2 + \ldots + p_c = 1$$

$$\text{or} \quad p_0 \left[ 1 + \frac{c\lambda}{\mu} + \frac{c(c-1)}{\underline{|2}} \left(\frac{\lambda}{\mu}\right)^2 + \cdots + \left(\frac{\lambda}{\mu}\right)^c \right] = 1$$

$$\text{or} \quad p_0 \left(1 + \frac{\lambda}{\mu}\right)^c = 1$$

$$\text{or} \quad p_0 = \left(\frac{\mu}{\lambda + \mu}\right)^c.$$

Hence,

$$\begin{aligned} p_n &= \frac{c(c-1)\cdots(c-n+1)}{\underline{|n}} \left(\frac{\lambda}{\mu}\right)^n \left(\frac{\mu}{\lambda+\mu}\right)^c \\ &= \binom{c}{n} \left(\frac{\lambda}{\lambda+\mu}\right)^n \left(\frac{\mu}{\lambda+\mu}\right)^{c-n}, \end{aligned}$$

which is a binomial distribution.

## 15.4 Non-poisson Queueing Models

### *15.4.1 Introduction*

When the arrivals and/or departures of a queueing system do not follow a Poisson distribution, they are called non-Poisson queueing systems. The development of these systems is more complicated, as the Poisson axioms do not hold. In studying non-Poisson queues, usually the following techniques are employed:

(i) Phase technique
(ii) Imbedded Markov chain technique
(iii) Supplementary variable technique.

In the phase technique, service is provided to a customer in several phases, say $k$ in number. The technique by which non-Markovian queues are reduced to Markovian queues is termed the imbedded Markov chain technique. In the supplementary variable technique, one or more random variables are added to convert a non-Markovian process into a Markovian one. These techniques have been applied in studying queueing systems like $M/E_k/1$, $M/G/1$, $G1/G/c$, etc. In this chapter, we shall discuss $M/E_k/1$ and $M/G/1$ queueing systems only.

### 15.4.2 Model (M/E$_k$/1):($\infty$/FCFS/$\infty$)

**Assumptions** In this model, we shall discuss a single server non-Poisson queueing system under the following assumptions:

(i) The arrivals follow a Poisson distribution.
(ii) The service time has an Erlang type $k$ distribution. Even though the service may not actually consist of $k$ phases, it is convenient in analysing this system to consider the Erlang as being made up of $k$ exponential phases, each with mean $\frac{1}{k\mu}$, where $\mu$ is the number of customers served per unit of time.
(iii) The service discipline is first come, first served.
(iv) The capacity of the system is infinite.
(v) This system consists of a single service channel in which there are $k$ phases in the system (waiting or in service).
(vi) A new arrival creates $k$ phases of service, and departure of one customer reduces $k$ phases of service.
(vii) The distribution of the total service time of a customer in the system will be some combined distribution of time in all these phases.
(viii) Each customer is served in $k$ phases one by one, and a new service does not start until all $k$ phases have been completed; i.e. each arrival increases the number of phases by $k$ in the system.

#### *Model Formulation*

If at any time there are $m$ customers waiting in the queue with one customer in service which has to still pass through $s$ phases, then the total number of phases $n$ in the system (waiting and in service) will be $n = mk + s$. Figure 15.5 shows the multiphase service-based model.

As $\mu$ is the number of customers served per unit time, then $k\mu$ will be the number of phases served per unit time and $\frac{1}{k\mu}$ will be the average time taken by the server at each phase.

Therefore, $\lambda_n = \lambda$, constant arrival per unit time
with $\mu_n = k\mu$ phases served per unit time.

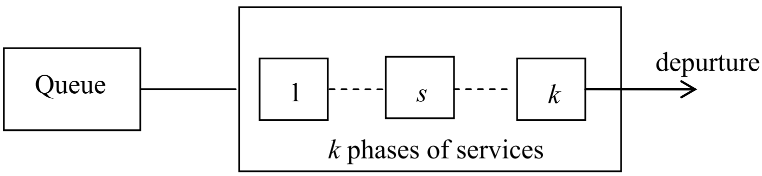



**Fig. 15.5** Multiphase service-based queueing system

Hence, the probability of one arrival during time $\Delta t$ is $\lambda\Delta t + O(\Delta t)$ and the probability of one departure (of $k$ phases) during time $\Delta t$ is $k\mu\Delta t + O(\Delta t)$.

Let $P_n(t)$ be the probability that there are $n$ phases in the system at time $t$. This model consists of a single service channel in which there are $n$ phases in the system (waiting or in service). According to the assumption of the model, a new arrival creates $k$ phases of service, and departure of one customer reduces $k$ phases of service.

To have zero phase in the system at time $t + \Delta t$, there may be the following cases:

(i) No phase in the system at time $t$, no arrival in time $\Delta t$
(ii) One phase in the system at time $t$, no arrival in time $\Delta t$, service of one unit (of $k$ phase) in time $\Delta t$, and so on.

Therefore, the probability of no phase in the system at time $t + \Delta t$ is given by

$$\begin{aligned} p_0(t+\Delta t) &= p_0(t)(1-\lambda\Delta t) + p_1(t)(1-\lambda\Delta t)k\mu\Delta t + O(\Delta t) \\ \text{or}\quad \frac{p_0(t+\Delta t)-p_0(t)}{\Delta t} &= -\lambda p_0(t) + k\mu p_1(t) + \frac{O(\Delta t)}{\Delta t}. \end{aligned} \tag{15.66}$$

To have $n$ (>0) phases in the system at time $t + \Delta t$, there may be the following cases:

(i) $(n - k)$ phases in the system at time $t$, one arrival in time $\Delta t$, no service (of any phase) in time $\Delta t$
(ii) $n$ phases in the system at time $t$, no arrival in time $\Delta t$, no service (of any phase) in time $\Delta t$
(iii) $(n + 1)$ phases in the system at time $t$, no arrival in time $\Delta t$, service of one unit (of $k$ phase) in time $\Delta t$, and so on.

Therefore, the probability $P_n(t + \Delta t)$ that there are $n$ (>0) phases in the system at time $t + \Delta t$ is given by

$$\begin{aligned} p_n(t+\Delta t) &= p_{n-k}(t)\lambda\Delta t(1-k\mu\Delta t) + p_n(t)(1-\lambda\Delta t)(1-k\mu\Delta t) \\ &\quad + p_{n+1}(t)(1-\lambda\Delta t)k\mu\Delta t + O(\Delta t)\text{or} \\ \frac{p_n(t+\Delta t)-p_n(t)}{\Delta t} &= \lambda p_{n-k}(t) - (\lambda+k\mu)p_n(t) + k\mu p_{n+1}(t) + \frac{O(\Delta t)}{\Delta t}. \end{aligned} \tag{15.67}$$

Now taking the limit as $\Delta t \to 0$ in (15.1) and (15.67), we get the differential difference equations for the system as follows:

$$p'_n(t) = \lambda p_{n-k}(t) - (\lambda + k\mu)p_n(t) + k\mu p_{n+1}(t) \tag{15.68}$$

$$\text{and}\quad p'_0(t) = -\lambda p_0(t) + k\mu p_1(t). \tag{15.69}$$

Under the steady state condition of the system, i.e. $\lim_{t\to\infty} p_n(t) = p_n$ and $\lim_{t\to\infty} p'_n(t) = 0$, these two equations reduce to

$$\lambda p_{n-k} - (\lambda + k\mu)p_n + k\mu p_{n+1} = 0, \quad n \geq 1 \tag{15.70}$$

and

$$-\lambda p_0 + k\mu p_1 = 0, \quad n = 0. \tag{15.71}$$

Let $\rho = \frac{\lambda}{k\mu}$; then Eqs. (15.70) and (15.71) reduce to

$$p_1 = \rho p_0, \quad n = 0 \tag{15.72}$$

and

$$(1+\rho)p_n = \rho p_{n-k} + p_{n+1}, n \geq 1. \tag{15.73}$$

Now, to solve Eq. (15.73), we shall use the generating function technique.

***Solution of the Model***

Let the generating function be $G(x)$, which is defined by

$$G(x) = \sum_{x=0}^{\infty} p_n x^n, \quad |x| \leq 1. \tag{15.74}$$

Multiplying both sides of (15.73) by $x^n$ and then summing over from 1 to $\infty$ (since the equation holds for $n \geq 1$), we have

$$(1+\rho)\sum_{n=1}^{\infty} p_n x^n = \rho \sum_{n=1}^{\infty} p_{n-k}x^n + \sum_{n=1}^{\infty} p_{n+1}x^n$$

or $$(1+\rho)\sum_{n=1}^{\infty} p_n x^n + \rho p_0 = p_1 + \rho \sum_{n=1}^{\infty} p_{n-k}x^n + \sum_{n=1}^{\infty} p_{n+1}x^n \quad [\because \rho p_0 = p_1]$$

or $$(1+\rho)\left[\sum_{n=1}^{\infty} p_n x^n + p_0\right] - p_0 = \rho \sum_{n=1}^{\infty} p_{n-k}x^n + \left[p_1 + \sum_{n=1}^{\infty} p_{n+1}x^n\right]$$

or $$(1+\rho)\sum_{n=0}^{\infty} -p_n x^n - p_0 = \rho \sum_{n=k}^{\infty} p_{n-k}x^n + \frac{1}{x}\sum_{n=0}^{\infty} p_{n+1}x^{n+1}$$

[since $p_{n-k} = 0$, for $n - k < 0$, i.e. for $n < k$]

or $$(1+\rho)\sum_{n=0}^{\infty} p_n x^n - p_0 = \rho \sum_{j=0}^{\infty} p_j x^{j+k} + \frac{1}{x}\sum_{i=1}^{\infty} p_i x^i$$

[substituting $n-k=j$, i.e. $=j+k$ and $n+1=i$, i.e.$n=i-1$]

$$\text{or}\quad (1+\rho)\sum_{n=0}^{\infty} p_n x^n - p_0 = \rho x^k \sum_{j=0}^{\infty} p_j x^j + \frac{1}{x}\left[\sum_{i=0}^{\infty} p_i x^i - p_0\right]$$

$$\text{or}\quad (1+\rho)G(x) - p_0 = \rho x^k G(x) + \frac{1}{x}[G(x) - p_0]$$

$$\begin{aligned}\text{or}\quad G(x) &= \frac{p_0(1-x)}{(1-x)-\rho x(1-x^k)} \qquad |x| \le 1 \\ &= \frac{p_0}{1-\rho x\left(\frac{1-x^k}{1-x}\right)} = p_0\left[1-\rho x\left(\frac{1-x^k}{1-x}\right)\right]^{-1} \\ &= p_0 \sum_{m=0}^{\infty} (\rho x)^m \left(\frac{1-x^k}{1-x}\right)^m \end{aligned} \tag{15.75}$$

[by binomial expansion].

$$\therefore\ G(x) = p_0 \sum_{m=0}^{\infty} (x\rho)^m \left(1+x+x^2+\cdots+x^{k-1}\right)^m$$

$$\left[\text{Since } \frac{1-x^k}{1-x} = 1+x+x^2+\cdots+x^{k-1}\right]$$

$$\text{or}\quad G(x) = p_0 \sum_{m=0}^{\infty} \rho^m \left(x+x^2+\cdots+x^k\right)^m.$$

Now we shall find the value of $p_0$.
From (15.74) and (15.75), we have

$$\sum_{n=0}^{\infty} p_n x^n = p_0 \sum_{m=0}^{\infty} \rho^m \left(x+x^2+\cdots+x^k\right)^m. \tag{15.76}$$

This is an identity. Now setting $x=1$ on both sides, we have

$$\sum_{n=0}^{\infty} p_n = p_0 \sum_{m=0}^{\infty} \rho^m (1+1+\cdots+k \text{ times})^m$$

$$\text{or}\quad 1 = p_0 \sum_{m=0}^{\infty} \rho^m k^m$$

$$\text{or}\quad 1 = p_0\left[1+\rho k+(\rho k)^2+(\rho k)^3+\cdots\right]$$

$$\text{or}\quad 1 = p_0 \frac{1}{1-\rho k},\quad \text{if}\ \ \rho k < 1 \text{ i.e., if } \frac{\lambda}{\mu} < 1 \text{ i.e., } \lambda < \mu$$

$$\text{or}\quad p_0 = 1 - k\rho.$$

Substituting $p_0 = 1 - k\rho$ in (15.76), we have

$$\sum_{n=0}^{\infty} p_n x^n = (1 - k\rho) \sum_{m=0}^{\infty} (\rho x)^m \left( \frac{1 - x^k}{1 - x} \right)^m$$
$$\text{or} \quad \sum_{n=0}^{\infty} p_n x^n = (1 - k\rho) \sum_{m=0}^{\infty} (\rho x)^m (1 - x^k)^m (1 - x)^{-m} \tag{15.77}$$

Now $(1 - x^k)^m = \sum_{r=0}^{m} (-1)^r \binom{m}{r} x^{rk}$
and $(1 - x)^{-m} = \sum_{s=0}^{\infty} \binom{m+s-1}{s} x^s.$
Hence, from (15.77), we have

$$\sum_{n=0}^{\infty} p_n x^n = (1 - k\rho) \sum_{n=0}^{\infty} \left[ \rho^m x^m \sum_{r=0}^{m} (-1)^r \binom{m}{r} x^{rk} \sum_{s=0}^{\infty} \binom{m+s-1}{s} x^s \right]$$
$$\text{or} \quad \sum_{n=0}^{\infty} p_n x^n = (1 - k\rho) \sum_{m=0}^{\infty} \rho^m \left[ \sum_{r=0}^{m} \sum_{s=0}^{\infty} (-1)^r \binom{m}{r} \binom{m+s-1}{s} x^{m+rk+s} \right].$$

Comparing the coefficients of $x^n$ from both sides, we have

$$p_n = (1 - k\rho) \sum_{m,r,s} \rho^m (-1)^r \binom{m}{r} \binom{m+s-1}{s}.$$

[The summation is taken over all such values of $m$, $r$, $s$ $(0 \le m < \infty,\ 0 \le r \le m,\ 0 \le s < \infty)$ for which $m + rk + s = n$.]

***Characteristics of the Model***

(i) ***Average (expected) number of phases in the system***

The average number of phases in the system is given by

$$L_{\mathrm{p}} = \sum_{n=0}^{\infty} n p_n$$

From (15.7), we have

$$(1 + \rho) p_n = \rho p_{n-k} + \rho_{n+1}, \quad n \ge 1.$$

Now multiplying both sides by $n^2$ and then summing from 1 to $\infty$, we have

$$(1+\rho)\sum_{n=1}^{\infty} n^2 p_n = \rho \sum_{n=1}^{\infty} n^2 p_{n-k} + \sum_{n=1}^{\infty} n^2 p_{n+1}$$

$$\text{or} \quad (1+\rho)\sum_{n=1}^{\infty} n^2 p_n = \rho \sum_{n=k}^{\infty} n^2 p_{n-k} + \sum_{n=0}^{\infty} n^2 p_{n+1}$$

$$[\text{since } p_{n-k} = 0 \text{ for } n-k<0, \text{i.e. for } n<k].$$

Now taking $n + 1 = m$ and $n - k = m$ in the first and second summations on the right-hand side respectively, we have

$$\begin{aligned}
(1+\rho)\sum_{n=0}^{\infty} n^2 p_n &= \rho \sum_{m=0}^{\infty} (m+k)^2 p_m + \sum_{m=1}^{\infty} (m-1)^2 p_m \\
&= \rho \sum_{n=0}^{\infty} (n+k)^2 p_n + \sum_{n=0}^{\infty} (n-1)^2 p_n - p_0 \\
&= \sum_{n=0}^{\infty} \left[\rho\left(n^2 + 2nk + k^2\right) + n^2 - 2n + 1\right] p_n - p_0 \\
&= (1+\rho)\sum_{n=0}^{\infty} n^2 p_n - 2(1-\rho k)\sum_{n=0}^{\infty} n p_n + \left(1+\rho k^2\right)\sum_{n=0}^{\infty} p_n - p_0
\end{aligned}$$

$$\text{or} \quad 2(1-\rho k)\sum_{n=0}^{\infty} n p_n = \left(1+\rho k^2\right) - p_0 \quad \left[\because \sum_{n=0}^{\infty} p_n = 1\right]$$

$$\text{or} \quad 2(1-\rho k)\sum_{n=0}^{\infty} n p_n = \left(1+\rho k^2\right) - (1-\rho k) \quad [\because p_0 = 1-\rho k]$$

$$\text{or} \quad \sum_{n=0}^{\infty} n p_n = \frac{\rho k(k+1)}{2(1-\rho k)}$$

$$\text{or} \quad L_{\text{p}} = \frac{\rho k(k+1)}{2(1-\rho k)} = \frac{\frac{\lambda}{\mu}(k+1)}{2\left(1-\frac{\lambda}{\mu}\right)} \qquad \left[\begin{aligned} &\because \rho = \frac{\lambda}{\mu k} \\ &\text{i.e.,} \quad \rho k = \frac{\lambda}{\mu} \end{aligned}\right]$$

$$= \frac{k+1}{2} \cdot \frac{\lambda}{\mu - \lambda}.$$

(ii) ***Average (expected) number of units in the queue***

We know that the average number of phases of one unit in service $= \frac{k+1}{2}$.

Therefore, the time taken for their service $= \frac{k+1}{2}\frac{1}{\mu}$, as $\frac{1}{\mu}$ is the mean service time per unit.

Hence, the expected number of phases in service
= expected number of phases arrived during the time $\frac{k+1}{2\mu}$

$= \frac{k+1}{2\mu}\lambda.$
Therefore, the expected number of units in the queue is given by

$$\begin{aligned} L_q &= \frac{1}{k}\left[L_p - \text{expected number of phases in service}\right] \\ &= \frac{1}{k}\left[\frac{k+1}{2}\frac{\lambda}{\mu-\lambda} - \frac{k+1}{2\mu}\lambda\right] \\ &= \frac{k+1}{2k}\cdot\frac{\lambda^2}{\mu(\mu-\lambda)}. \end{aligned}$$

(iii) *Average (expected) number of units in the system*

From the relation between the expected number of units in the system and in the queue, we have

$$L_s = L_q + \frac{\lambda}{\mu}$$
$$\text{or} \quad L_s = \frac{k+1}{2k}\frac{\lambda^2}{\mu(\mu-\lambda)} + \frac{\lambda}{\mu}.$$

(iv) *Expected waiting time in the system*

The expected waiting time in the system is given by

$$\begin{aligned} W_s &= \frac{1}{\lambda}L_s \\ &= \frac{k+1}{2k}\frac{\lambda}{\mu(\mu-\lambda)} + \frac{1}{\mu}. \end{aligned}$$

(v) *Expected waiting time in the queue*

The expected waiting time in the queue is given by

$$\begin{aligned} W_q &= \frac{1}{\lambda}L_q \\ &= \frac{k+1}{2k}\frac{\lambda}{\mu(\mu-\lambda)}. \end{aligned}$$

*Example 7* In a factory cafeteria, the customers have to pass through three counters. The customers buy coupons at the first counter, select and collect the snacks at the second counter and collect tea at the third. The server at each counter takes on average 1.5 min, although the distribution of service time is approximately exponential. If the arrival of customers to the cafeteria is approximately Poisson at an average rate of 6 per hour, calculate:

(i) The average time that a customer spends waiting in the cafeteria
(ii) The average time of getting the service
(iii) The most probable time getting the service.

**Solution**

This is a queueing problem of the form $(M/E_k/1)$:$(\infty/\text{FCFS}/\infty)$.

Here $k$ = number of phases = 3

Service time per phase = 1.5 min

∴ Service time per customer = 1.5 × 3 = 4.5 min

$\therefore\ \mu = \frac{1}{4.5}$ customers/minute = $\frac{40}{3}$ customers/h.

Here $\lambda = 6$ customers/h.

(i) The average waiting time spent by a customer in the cafeteria is given by

$$W_q = \frac{k+1}{2k} \cdot \frac{\lambda}{\mu(\mu-\lambda)} = \frac{3+1}{2\times 3} \frac{6}{\frac{40}{3}\left(\frac{40}{3}-6\right)}\ \text{h}$$
$$= \frac{9}{220}\ \text{h} = 2.45\ \text{min}.$$

(ii) The average time of getting the service

$$= \frac{1}{\mu} = \frac{3}{40}\ \text{h} = 4.5\ \text{min}$$

(iii) The most probable time spent in getting the service
= the modal value of service time $t$
= the modal value of the Erlang distribution.
Hence, the most probable time spent $= \frac{k-1}{\mu k} = \frac{3-1}{\frac{40}{3}\times 3} = \frac{3-1}{\frac{40}{3}\times 3} = \frac{1}{20}$ h = 3 min.

[The probability density function of the Erlang distribution with parameters $\mu$ and $k$ is given by

$$f(t,\mu,k) = \frac{(\mu k)^k}{(k-1)!} t^{k-1} e^{-k\mu t}, \quad 0 \le t < \infty, k = 1, 2, \ldots.$$

The mode of the distribution is at $\frac{k-1}{k\mu}$ and the mean $= \frac{1}{\mu}$, with variance $= \frac{\mu^2}{k}$.]

### 15.4.3 Model (M/G/1):($\infty$/GD)

A precise description of this model is as follows:

(i) Customers arrive according to a Poisson fashion at an average rate of $\lambda$.
(ii) The service time $t$ is represented by any probability distribution with mean $E(t) = \frac{1}{\mu}$ and variance $v(t) = \sigma^2$.
(iii) This system consists of a single server.
(iv) The service discipline is a general service discipline such as FCFS, LCFS, SIRO, etc.
(v) An arriving customer who finds the servers idle begins to take the service immediately.
(vi) All blocked customers wait until served.
(vii) The server cannot be idle when there is a waiting customer.

Here the probability that a customer leaves the system in $(t,\, t + \Delta t)$ is a quantity dependent on when the service of that particular customer begins. That is, it would involve not only $t$, but also another time variable, say $t'$, measured from the beginning of the current service.

Let $q_0$ denote the queue length when the service of the $n$th customer terminates and let $q_1$ be the queue length when the service of the $(n+1)$ th customer terminates. Further, let $k(k = 0,\, 1,\, 2,\, \ldots)$ denote the number of customers who arrive during the service of the $(n+1)$th customer. Then the relation between $q_0$, $q_1$ and $k$ is given by

$$q_1 = \begin{cases} k \quad \text{if} \quad q_0 = 0, \\ \text{i.e. the queue length is zero after serving } n\text{th customer.} \\ (q_0 - 1) + k, \quad \text{if} \quad q_0 > 0, \\ \text{i.e. the queue has } q_0 \text{ units after serving } n\text{-th customer.} \end{cases}$$

This relation can be written as

$$q_1 = q_0 - \delta + k, \tag{15.78}$$

where $\delta = \begin{cases} 0 & \text{if} \quad q_0 = 0 \\ 1 & \text{if} \quad q_0 > 0 \end{cases}$.

Now, taking expectation on both sides of (15.78), we have

$$E(q_1) = E(q_0) - E(\delta) + E(k). \tag{15.79}$$

In a steady state system, $E(q_1) = E(q_0)$.
Hence, from (15.79), we have

$$E(\delta) = E(k). \tag{15.80}$$

Now, squaring both sides of (15.78), we have

$$\begin{aligned} & q_1^2 = q_0^2 + k^2 + \delta^2 + 2kq_0 - 2\delta q_0 - 2\delta k \\ \text{or} \quad & q_1^2 = q_0^2 + k^2 + \delta + 2kq_0 - 2q_0 - 2k\delta \\ & \begin{bmatrix} \text{since } \delta \text{ can take values } 0 \text{ and } 1 \text{ only,} \\ \text{so } \delta^2 = \delta \text{ and } \delta q_0 = q_0 \end{bmatrix}. \end{aligned} \tag{15.81}$$

Now taking expectation on both sides, we have

$$E(q_1^2) = E(q_0^2) + E(k^2) + E(\delta) + 2E(kq_0) - 2E(q_0) - 2E(k\delta). \tag{15.82}$$

Again, since $E(q_1^2) = E(q_0^2)$ is in a steady state system, from (15.82) we have

$$\begin{aligned} & 2E(q_0) - 2E(kq_0) = E(k^2) + E(\delta) - 2E(k\delta) \\ \text{or} \quad & 2E(q_0) - 2E(k)E(q_0) = E(k^2) + E(k) - 2\{E(k)\}^2 \\ & \begin{bmatrix} \text{From (3), } E(\delta) = E(k) \\ \text{and } q_0, k \text{ and } \delta \text{ are independent} \end{bmatrix} \\ \text{or} \quad & E(q_0) = \frac{E(k^2) + E(k) - 2\{E(k)\}^2}{2\{1 - E(k)\}}. \end{aligned} \tag{15.83}$$

Now, in order to determine $E(q_0)$, the values of $E(k)$ and $E(k^2)$ must be computed.

Since the arrivals in time $t$ follow a Poisson distribution,

$$E(k/t) = \lambda t, \quad E(k^2/t) = (\lambda t)^2, \quad V(k/t) = E(k^2/t) - \{E(k/t)\}^2,$$

where $\lambda$ is the mean arrival rate.

Hence,

$$E(k) = \int_0^\infty E(k/t) f(t) \mathrm{d}t = \int_0^\infty \lambda t f(t) \mathrm{d}t = \lambda \int_0^\infty t f(t) \mathrm{d}t = \lambda E(t)$$

Also,

$$\begin{aligned} E(k^2) &= \int_0^\infty E(k^2/t) f(t) \mathrm{d}t = \int_0^\infty \left[ (\lambda t)^2 + \lambda t \right] f(t) \mathrm{d}t \\ &= \lambda^2 E(t^2) + \lambda E(t) = \lambda^2 V(t) + \lambda^2 \{E(t)\}^2 + \lambda E(t). \end{aligned}$$

Substituting the values of $E(k)$ and $E(k^2)$ in (15.5), we have

$$\begin{aligned}
E(q_0) &= \frac{\lambda^2 V(t) + \lambda^2\{E(t)\}^2 + \lambda E(t) + \lambda E(t) - 2\lambda^2\{E(t)\}^2}{2\{1-\lambda E(t)\}} \\
&= \frac{\lambda^2 V(t) + \lambda^2\{E(t)\}^2 + 2\lambda E(t) - 2\lambda^2\{E(t)\}^2}{2\{1-\lambda E(t)\}} \\
&= \frac{\lambda^2\left[V(t) + \{E(t)\}^2\right]}{2\{1-\lambda E(t)\}} + \lambda E(t) \\
&= \frac{\lambda^2\left[\sigma^2 + \left(\frac{1}{\mu}\right)^2\right]}{2\left\{1-\lambda\cdot\frac{1}{\mu}\right\}} + \lambda\cdot\frac{1}{\mu}\left[\begin{array}{l}\text{As } E(t) = \dfrac{1}{\mu} \text{ where } \mu \text{ is the mean} \\ \text{service time and } V(t) = \sigma^2\end{array}\right] \\
&= \frac{\lambda^2\sigma^2 + \rho^2}{2(1-\rho)} + \rho\left[\text{As } \rho = \frac{\lambda}{\mu}\right] \\
\text{or}\quad L_s &= \rho + \frac{\lambda^2\sigma^2 + \rho^2}{2(1-\rho)}.
\end{aligned} \tag{15.84}$$

This is known as the Pollaczek-Khintchine (P-K) formula.

**Alternative form of Pollaczek formula**

If the $(n + 1)$th customer has to wait for the time $w$ units before taking the service, the length of the queue must be $q_1$ during the time $w + t$.
Hence, $E(q_1) = \lambda E(w+t) = \lambda E(w) + \lambda E(t)$
or $\lambda E(w) = E(q_1) - \lambda E(t)$,
i.e. the average waiting time $= E(w) = \frac{1}{\lambda}\left[E(q_1) - \frac{\lambda}{\mu}\right]$ where $E(t) = \frac{1}{\mu}$

$$\begin{aligned}
\text{or}\quad E(w) &= \frac{1}{\lambda}\left[P + \frac{\lambda^2\sigma^2+\rho^2}{2(1-\rho)} - \rho\right] \quad \left[\begin{array}{l}\text{since } E(q_1) = E(q_0) \\ \text{in the steady state}\end{array}\right] \\
&= \frac{\lambda^2\sigma^2+\rho^2}{2(1-\rho)\lambda}. \\
\therefore\quad \frac{E(w)}{E(t)} &= \frac{\mu\left(\lambda^2\sigma^2+\rho^2\right)}{2(1-\rho)\lambda} \\
\text{or}\quad \frac{E(w)}{E(t)} &= \frac{\rho}{2(1-\rho)}\left(1+\mu^2\sigma^2\right).
\end{aligned} \tag{15.85}$$

This formula was also suggested by Pollaczek.
Hence, the average number of customers in the system is given by

$$L_s = \rho + \frac{\lambda^2\sigma^2 + \rho^2}{2(1-\rho)}.$$

The average number of customers in the queue is given by

$$L_q = L_s - \frac{\lambda}{\mu} = L_s - \rho = \frac{\lambda^2\sigma^2 + \rho^2}{2(1-\rho)}.$$

*Example 8* In a heavy machine shop, the overhead crane is 75% utilized. Time study observations gave the average slinging time as 10.5 min with a standard deviation of 8.8 min. What is the average calling rate for the service of the crane, and what is the average delay in getting service? If the average service time is cut to 8.0 min with a standard deviation of 6.0 min, how much reduction will occur, on average, in the delay of getting service?

**Solution**
This is an $(M/G/1)$:$(\infty/\text{FCFS}/\infty)$ queueing system.
Since the overhead crane is utilized 75%, $\rho = 0.75$.
Again,

$$\mu = \frac{60}{10.5} = 5.71 \text{ per h}$$
$$\lambda = \rho\mu = 0.75 \times 5.71 = 4.29 \text{ per h}$$
$$\sigma = 8.8 \text{ min} = \frac{8.8}{60} \text{ h}.$$

Now. the average delay in getting the service is given by

$$L_q = \frac{\lambda^2\sigma^2 + \rho^2}{2\lambda(1-\rho)} = \frac{(4.29)^2 \times \left(\frac{8.8}{60}\right)^2 + (0.75)^2}{2 \times 4.29 \times (1 - 0.75)} = 0.4468 \text{ h} = 26.8 \text{ min}.$$

If the service time is cut to 8 min, then

$$\mu = \frac{60}{8} = 7.5 \text{ per hour and } \rho = \frac{4.29}{7.5} = 0.571.$$

In this case, the standard deviation = 6.0 min $= \frac{6}{60} = 0.1$ h.

Then

$$L_q = \frac{(4.29)^2 \times (0.1)^2 + (0.571)^2}{2 \times 4.29 \times (1 - 0.571)} \text{ h} = 0.1386 \text{ h} = 8.3 \text{ min}.$$

In the second case, a reduction of (26.8–8.3), i.e. 18.5 min, will occur on average in the delay of getting service.
The average waiting time of a customer in queue is given by

$$W_q = \frac{L_q}{\lambda} = \frac{\lambda^2\sigma^2 + \rho^2}{2\lambda(1-\rho)}.$$

The average waiting time that a customer spends in the system is given by

$$\begin{aligned} W_s &= W_q + \text{service time} \\ &= \frac{\lambda^2\sigma^2 + \rho^2}{2\lambda(1-\rho)} + \frac{1}{\mu}. \end{aligned}$$

**Efficiency of this queueing system**

The efficiency of this queueing system is measured by the ratio $\frac{W_q}{E(t)}$ as follows:

$$\frac{W_q}{E(t)} = \frac{\text{average waiting time in the queue}}{\text{average service time}}.$$

From this ratio, we conclude that the system will be better for a smaller ratio.
**Note**: In particular, let $G = M$. Then

$$\mathrm{Var}(t) = \frac{1}{\mu^2}, \quad E(t) = \frac{1}{\mu}.$$

$$L_s = \frac{\lambda}{\mu} + \frac{\lambda^2\left[\frac{1}{\mu^2} + \frac{1}{\mu^2}\right]}{2\left(1 - \frac{\lambda}{\mu}\right)} = \rho + \frac{\rho^2}{1-\rho} = \frac{\rho}{1-\rho},$$

which is the same result as obtained earlier for the queueing model ($M/M/1$):($\infty$/FCFS/$\infty$).

*Example 9* In a queueing system with a single server, arrivals are Poisson distributed with mean 5 per hour. Find the expected waiting time in the system if the service time distribution is normal with mean 3 min and standard deviation 2 min.

**Hints**:

Here $\lambda$ = 5 per hour, $\mu = \frac{60}{3} = 20$ per hour.

$$\sigma = 2\,\text{min} = \frac{2}{60}\ \text{h}.$$

The expected waiting time in the system

$$= L_s = \frac{\lambda^2\sigma^2 + \rho^2}{2\lambda(1-\rho)} + \frac{1}{\mu}.$$

## 15.5 Mixed Queueing Model and Cost

### 15.5.1 Introduction

In this section, we shall discuss a queueing model with random arrival rate and deterministic service rate. This type of queueing model is known as a mixed queueing model. We shall also discuss the cost models of an *M*/*M*/1 system with finite and infinite capacity to find the optimum service level by minimizing the cost of the system.

### 15.5.2 Mixed Queueing Model

For a queueing model, if either the arrival rate or the service rate is fixed and the other one is random, then the model is known as a mixed queueing model.

**Model**: **(M/D/1)**:**(∞/FCFS/∞)**

This queueing system involves a single server, Poisson arrivals and a deterministic distribution. The capacity of the system is infinite. Here the service time is not a random variable but a constant. Without loss of generality, the service time can be taken as unity, $\mu = 1$.

It is also assumed that a steady state condition prevails, with $\rho = \lambda < 1$ as $\mu = 1$.

In this scheme, the system is observed just after the completion of service of a customer. Hence, if $n$ customers are in the system at the observation time, then at the time of the next observation, the number $n'$ in the system is given by

$$n' = \begin{cases} k & \text{if } n = 0 \\ n - 1 + k, & \text{if } n > 0 \end{cases},$$

where $k = 0,\ 1,\ 2,\ \ldots$ is the number of arrivals during the service time which is taken as unity. Let $q_k$ be the probability that $k$ persons arrive during the service time.

Since the arrival rate follows a Poisson distribution,

$$q_k = \frac{e^{-\lambda}(\lambda)^k}{k!}, \quad k = 0, 1, 2, \ldots$$

**Table 15.2** Transaction matrix

| $n' \to n$ | 0 | 1 | 2 | 3 | 4 | … | … |
|---|---|---|---|---|---|---|---|
| 0 | $q_0$ | $q_1$ | $q_2$ | $q_3$ | $q_4$ | … | … |
| 1 | $q_0$ | $q_1$ | $q_2$ | $q_3$ | $q_4$ | … | … |
| 2 | 0 | $q_0$ | $q_1$ | $q_2$ | $q_3$ | … | … |
| 3 | 0 | 0 | $q_0$ | $q_1$ | $q_2$ | … | … |
| 4 | 0 | 0 | 0 | $q_0$ | $q_1$ | … | … |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

where $\lambda$ is the mean arrival rate.
For different values of $n$ and $n'$, the transaction matrix between the end and the beginning of the service is given by Table 15.2.

In the transaction matrix, $q_0$ is the probability of no arrival during the service time.

Let $p_n(n = 0,\ 1,\ 2,\ \ldots)$ be the steady state probability that there are $n$ customers in the system. Then the probabilities are given by

$$\begin{aligned}
p_0 &= p_0q_0 + p_1q_0 \\
p_1 &= p_0q_1 + p_1q_1 + p_2q_0 \\
p_2 &= p_0q_2 + p_1q_2 + p_2q_1 + p_3q_0 \\
\cdots & \quad \cdots \quad \cdots \\
\cdots & \quad \cdots \quad \cdots
\end{aligned}$$

The $(n+1)$th equation of the above is given by

$$p_n = p_0q_n + p_1q_n + p_2q_{n-1} + p_3q_{n-3} + \cdots + p_{n+1}q_0, \quad n = 0, 1, 2, \ldots$$

$$\text{or} \quad p_n = p_0q_n + \sum_{j=1}^{n+1} p_jq_{n+1-j}, \quad n = 0, 1, 2, \ldots. \tag{15.86}$$

Applying a $Z$-transform to Eq. (15.86), we have

[$Z$-transform of $p_n$ is defined by $Z(p_n) = \sum_{n=0}^{\infty} z^n p_n = P(z)$ (say)]

$$\sum_{n=0}^{\infty} z^n p_n = \sum_{n=0}^{\infty} z^n p_0 q_n + \sum_{n=0}^{\infty} z^n \left\{ \sum_{j=1}^{n+1} p_j q_{n+1-j} \right\}$$

$$= \sum_{n=0}^{\infty} z^n p_n q_n + \sum_{n=0}^{\infty} z^{n+1-1} \left\{ \sum_{j=0}^{n+1} p_j q_{n+1-j} \right\}$$

$$\text{or} \quad \sum_{n=0}^{\infty} z^n p_n = p_0 \sum_{n=0}^{\infty} z^n q_n + \frac{1}{z} \sum_{m=1}^{\infty} z^m \left\{ \sum_{j=0}^{m} p_j q_{m-j} - p_0 q_m \right\}, \quad \text{where} \quad n = m+1. \tag{15.87}$$

Let us consider

$$r_m = \sum_{j=0}^{m} p_j q_{m-j} \tag{15.88}$$
$$\text{or} \quad r_m = p_0 q_m + p_1 q_{m-1} + p_2 q_{m-2} + \cdots + p_m q_0,$$

i.e. $r_m$ is the convolution of $p_m$ and $q_m$.
Hence, $Z(r_m) = Z(q_m)Z(p_m)$

$$\text{or} \quad R(z) = Q(z)P(z), \quad \text{where} \quad R(z) = Z(r_m) \text{ and } Q(z) = Z(q_m). \tag{15.89}$$

Then, from (15.87), we have

$$\begin{aligned}
P(z) &= p_0 Q(z) + \frac{1}{z} \sum_{m=1}^{\infty} z^m r_m - \frac{p_0}{z} \sum_{m=1}^{\infty} z^m q_m \\
&= p_0 Q(z) + \frac{1}{z} \{R(z) - r_0\} - \frac{p_0}{z} \{Q(z) - q_0\} \\
&= p_0 Q(z) + \frac{1}{z} R(z) - \frac{p_0}{z} Q(z) \quad [\text{as } r_0 = p_0 q_0] \\
&= p_0 Q(z) + \frac{1}{z} P(z)Q(z) - \frac{p_0}{z} Q(z) \quad [\text{as } R(z) = P(z)Q(z)]
\end{aligned} \tag{15.90}$$

$$\text{or} \quad P(z) = p_0 \left[ \frac{(z-1)Q(z)}{z - Q(z)} \right].$$

Now,

$$\begin{aligned}
P(1) &= \lim_{z \to 1} P(z) = p_0 \lim_{z \to 1} \left[ \frac{(z-1)Q(z)}{z - Q(z)} \right] \left( = \frac{0}{0} \text{ form} \right) \\
&= p_0 \lim_{z \to 1} \frac{(z-1)Q'(z) + Q(z)}{1 - Q'(z)} \qquad [\text{Using L'Hospital's rule}] \\
&= p_0 \frac{Q(1)}{1 - Q'(1)} \\
&= \frac{p_0}{1 - Q'(1)} \; [\text{as } Q(1) = 1].
\end{aligned}$$

Hence, $\frac{p_0}{1-Q'(1)} = 1$

or $p_0 = 1 - Q'(1)$.
Now we have to compute $Q'(1)$.

$$\therefore Q(z) = \sum_{n=0}^{\infty} q_n z^n, \text{ where } q_n \text{ is the probability of new arrivals.}$$

$$\therefore Q(z) = \sum_{n=0}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} z^n \quad \left[\text{Since the distribution of arrivals is Poisson}, q_n = \frac{e^{-\lambda}\lambda^n}{n!}\right]$$

$$= e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda z)^n}{n!}$$

$$= e^{\lambda(z-1)}.$$

Hence, $Q'(1) = \lambda$ i.e., $p_0 = 1 - \lambda$.
From (15.90), we have

$$P(z) = \frac{(1-\lambda)(z-1)Q(z)}{z - Q(z)} = \frac{(1-\lambda)(z-1)e^{\lambda(z-1)}}{z - e^{\lambda(z-1)}} = \frac{(1-\lambda)(1-z)}{1 - ze^{-\lambda(z-1)}}.$$

**Determination of $p_n$**

In obtaining $p_n$ we observe that a general expression cannot be easily obtained in this case because of the complexity of $P(z)$. However, $p_n$ can be obtained by using the formula

$$P^n(0) = n! p_n,$$

where $P^n(0) = \left[\frac{\partial^n P(z)}{\partial z^n}\right]_{z=0}$ $\left[\text{as } P(z) = \sum_{n=0}^{\infty} p_n z^n\right]$.
Therefore, $p_n = \frac{1}{n!} P^n(0)$.
Again, the expected number of customers in the system is given by

$$L_s = \sum_{n=0}^{\infty} n p_n = P'(1) \quad \left[\begin{array}{l}\text{since } P(z) = \sum_{n=0}^{\infty} p_n z^n \\ \therefore P'(z) = \sum_{n=0}^{\infty} n p_n z^{n-1}\end{array}\right]$$

$$\text{or} \quad L_s = \frac{\lambda + Q''(1)}{2(1-\lambda)} = \frac{\lambda + \lambda^2}{2(1-\lambda)} \qquad \left[\text{As } P(z) = p_0\left[\frac{(z-1)Q(z)}{Z - Q(z)}\right]\right].$$

## 15.6 Cost Models in Queueing System

In a queueing system, there are two types of conflicting costs:

(i) Cost of offering the service to the customers
(ii) Cost incurred due to delay in offering the service to the customers.

These two types of costs conflict in nature. An increase in the existing service facility would reduce the customers' waiting time; on the other hand, a decrease in the level of service would increase the customers' waiting time. This means that a higher cost of offering the service would reduce the cost of waiting to the customers and vice versa.

Therefore, the service level in a queueing system is a function of the service rate $\mu$, which balances the two conflicting costs. The objective is to find the optimum service level such that the sum of these costs is minimized. Hence, the total cost $TC$ of the queueing system is given by

$$\begin{aligned} TC = {} & <\text{cost for providing service facility}> \\ & + <\text{cost for waiting of customers}> \end{aligned}$$

$$TC = C_1\mu + C_2 L_s$$

where $C_1$ = cost of service per customer per unit time
$C_2$ = cost for waiting of customers per unit time
$L_s$ = average number of customers in the queueing system.

Now, for the queueing model $(M/M/1)$:$(\infty$ /FCFS/$\infty)$,

$$L_s = \frac{\lambda}{\mu - \lambda}$$

and

$$TC = C_1\mu + C_2\frac{\lambda}{\mu - \lambda}. \tag{15.91}$$

Here $TC$ is a function of the continuous variable $\mu$.

Then, the necessary condition for $TC$ to be optimum is given by

$$\frac{\mathrm{d}(TC)}{\mathrm{d}\mu} = 0,$$

which implies $C_1 - \frac{C_2\lambda}{(\mu-\lambda)^2} = 0$

$$\Rightarrow \quad \mu = \lambda + \sqrt{\lambda C_2 / C_1}. \tag{15.92}$$

Then $\frac{\mathrm{d}^2 TC}{\mathrm{d}\mu^2} = \frac{2\lambda C_2}{(\mu-\lambda)^3}$, which is positive for $\mu = \lambda + \sqrt{\lambda C_2 / C_1}$.

Hence, the optimum value of $\mu$ is

$$\mu^* = \lambda + \sqrt{\lambda C_2 / C_1}. \tag{15.93}$$

This result shows that the optimum value of $\mu$ is dependent on $C_1$, $C_2$ and also on the arrival rate $\lambda$.

Again, for the queueing model ($M/M/1$):($N$/FCFS/$\infty$), the capacity of the system is finite, and its value is $N$ when the system is fulfilled; then newly arrived customers cannot join in the queue. As a result, the larger the value of $N$, the smaller will be the number of lost customers. In this case, the total cost $TC$ of the system is given by

$$\begin{aligned} TC &= \langle \text{cost of service rate} \rangle + \langle \text{cost of waiting customer} \rangle \\ &\quad + \langle \text{cost of servicing customers} \rangle + \langle \text{cost of lost customers} \rangle \\ \text{or} \quad TC &= C_1\mu + C_2 L_s + C_3 N + \lambda p_N C_4, \end{aligned} \tag{15.94}$$

where

$C_3$ = cost of servicing to each customer per unit time

$C_4$ = cost per lost customer

and $\lambda p_N$ = number of lost customers per unit time.

**Case 1** When $\rho \neq 1$

$$L_s = \frac{\rho[1-(N+1)\rho^N + N\rho^{N+1}]}{(1-\rho)(1-\rho^{N+1})} \tag{15.95}$$

and

$$p_N = \frac{(1-\rho)\rho^N}{1-\rho^{N+1}}. \tag{15.96}$$

In this case,

$$TC = C_1\mu + \frac{C_2\rho[1-(N+1)\rho^N + N\rho^{N+1}]}{(1-\rho)(1-\rho^{N+1})} + C_3 N + \lambda C_4 (1-\rho)\rho^N / \left(1-\rho^{N+1}\right). \tag{15.97}$$

**Case 2** When $\rho = 1$

$$L_s = \frac{N}{2} \text{ and } p_N = \frac{1}{N+1}.$$

In this case,

$$TC = C_1\mu + C_2N/2 + C_3N + \lambda C_4/(N+1). \tag{15.98}$$

As $N$ is given, $TC$ is a function of the continuous variable $\mu$ only. The optimum value of $\mu$ can be obtained easily by minimizing the total cost $TC$ given by either (15.97) or (15.98).

## 15.7 Exercises

1. Arrivals at a telephone booth are considered to be Poisson, with an average time of 10 min between one arrival and the next. The length of a phone call is assumed to be distributed exponentially with mean 3 min.
   (i) What is the probability that a person arriving at the booth will have to wait?
   (ii) Find the average number of units in the system.
   (iii) Estimate the fraction of a day that the phone will be in use.
   (iv) What is the probability that it will take a person more than 10 min altogether to wait for the phone and complete his call?
2. At a certain airport it takes exactly 5 min to land an aeroplane, once it is given the signal to land. Although incoming planes have scheduled arrival times, the wide variability in arrival times produces an effect which makes the incoming planes appear to arrive in a Poisson fashion at an average rate of 6 per hour. This produces occasional congestion in the airport, which can be dangerous and costly. Under these circumstances, how much time will a pilot expect to spend circling the field waiting to land?
3. Patients arrive at a clinic according to a Poisson distribution at a rate of 30 patients per hour. The waiting room does not accommodate more than 14 patients. The examination time per patient is exponential with a mean rate of 20 per hour.
   (i) Find the effective arrival rate at the clinic.
   (ii) What is the probability that an arriving patient will not wait?
   (iii) What is the expected waiting time until a patient is discharged from the clinic?

4. A barber with a one-man operation takes exactly 25 min to complete one haircut. If customers arrive in a Poisson fashion at an average rate of one every 40 min, how long on average must a customer wait for service? Also find the average time that a customer spends in the barbershop.
5. A road transport company has one reservation clerk on duty at a time. He handles information on bus schedules and makes reservations. Customers arrive at a rate of 8 per hour, and the clerk can service 12 customers on average per hour.

    (i) What is the average number of customers waiting for the service of the clerk?
    (ii) What is the average time a customer has to wait before getting service?
    (iii) The management is contemplating installing a computer system to handle the information and reservations. This is expected to reduce the service time from 5 to 3 min. The additional cost of having the new system works out to Rs. 400.00 per day. If the cost of goodwill of having to wait is estimated to be Re. 1.00 per minute spent waiting to be served, should the company install the computer system? Assume that the working time is 8 h per day.

6. At a one-man barbershop, customers arrive according to a Poisson distribution with a mean arrival rate of 5 per hour, and the hair-cutting time is exponentially distributed, with an average haircut taking 10 min. It is assumed that because of his excellent reputation, customers are always willing to wait for this barber. Calculate:

    (i) The average number of customers in the shop and the average number of customers waiting for a haircut
    (ii) The percentage of time an arrival can walk right in without having to wait
    (iii) The percentage of customers who have to wait prior to getting into the barber's chair.

7. At a certain filling station, customers arrive in a Poisson process with an average time of 12 per hour. The time intervals between services follow an exponential distribution, and as such the mean time taken to service a unit is 2 min. Evaluate:

    (i) The probability that there is no customer at the counter
    (ii) The probability that there are more than two customers at the counter
    (iii) The probability that there is no customer to be served
    (iv) The probability that a customer is being served, but nobody is waiting
    (v) The expected number of customers in the waiting line
    (vi) The expected time a customer spends in the system.

8. A supermarket has a single cashier. During peak hours, customers arrive at a rate of 20 customers per hour. The average number of customers who can be processed by the cashier is 24 per hour. Calculate:

(i) The probability that the cashier is idle
(ii) The average number of customers in the queueing system
(iii) The average time a customer spends in the system
(iv) The average number of customers in the queue
(v) The average time a customer spends in the queue waiting for service.

9. An airlines organization has one reservation clerk on duty in its local branch at any given time. The clerk handles information regarding passenger reservations and flight schedules. Assume that the number of customers arriving during any given period is Poisson distributed with an arrival rate of 8 per hour and that the reservation clerk can serve a customer in 6 min on average, with an exponentially distributed service time.

   (i) What is the probability that the system is busy?
   (ii) What is the average time that a customer spends in the system?
   (iii) What is the average length of the queue, and what is the number of customers in the system?

10. At a public telephone booth in a post office, arrivals are considered to be Poisson with an average inter-arrival time of 12 min. The length of a phone call may be assumed to be distributed exponentially with an average of 4 min.

   (i) What is the probability that a fresh arrival will not have to wait for the phone?
   (ii) What is the probability that an arrival will have to wait more than 10 min before the phone is free?
   (iii) What is the average length of queues that form from time to time?

11. A bank plans to open a single server drive-in banking facility at a particular centre. It is estimated that 20 customers will arrive each hour on average. If, on average, it requires 2 min to process a customer's transaction, determine:

   (i) The proportion of time that the system will be idle
   (ii) The average of how long a customer will have to wait before reaching the server
   (iii) The fraction of customers who will have to wait.

12. A warehouse has only one loading dock manned by a three-person crew. Trucks arrive at the loading dock at an average rate of 4 trucks per hour, and the arrival rate is Poisson distributed. The loading of a truck takes 10 min on average and can be assumed to be exponentially distributed. The operating cost of a truck is Rs. 50 per hour, and the members of the loading crew are paid Rs. 15 each per hour. Would you advise the truck owner to add another crew of three persons?
13. In a factory, the machine breakdown on an average rate is 10 machines per hour. The idle time cost of a machine is estimated to be Rs. 100 per hour. The factory works 8 h a day. The factory manager is considering 2 mechanics for repairing the machines. The first mechanic *A* takes about 5 min on average to

repair a machine and demands wages Rs. 50 per hour. The second mechanic $B$ takes about 4 min in repairing a machine and demands wages at the rate of Rs. 75 per hour. Assuming that the rate of machine breakdown is Poisson distributed and the repair rate is exponentially distributed, which of the two mechanics should be engaged?

14. In the production shop of a company, the breakdown of the machines is found to be Poisson distributed with an average rate of 3 machines per hour. The breakdown time at one machine costs Rs. 200 per hour to the company. The company is considering two repairmen for hire. One of the repairmen is slow but cheap, the other fast but expensive. The slow-cheap repairman demands Rs. 100 per hour and will repair the broken-down machines exponentially at the rate of 4 per hour. The fast-expensive repairman demands Rs. 150 per hour and will repair machines exponentially at an average rate of 6 per hour. Which repairman should be hired?
15. At a railway station, only one train is handled at a time. The railway yard is sufficient only for waiting of two trains to wait; then a newly arriving train is given the signal to leave the station. Trains arrive at the station at an average rate of 6 per hour, and the railway station can handle them on an average of 12 per hour. Assuming Poisson arrivals and an exponential service distribution, find the steady state probabilities for the various numbers of trains in the system. Also find the average waiting time of a new train coming into the yard.
16. Assume that the goods trains are coming into a yard at the rate of 30 trains per day, and suppose that the inter-arrival times follow an exponential distribution. The service time for each train is assumed to be exponential with an average of 36 min. If the yard can admit 9 trains at a time (there being 10 lines, one of which is reserved for shunting purposes), calculate the probability that the yard is empty and find the average queue length.
17. If for a period of 2 h in the day (8 to 10 a.m.) trains arrive at the yard every 20 min but the service time continues to remain 36 min, then calculate for this period: (a) the probability that the yard is empty, (b) the average number of trains in the system, on the assumption that the line capacity of the yard is limited to 4 trains only.
18. A petrol station has a single pump and space for not more than 3 cars (2 waiting, 1 being served). A car arriving when the space is filled to capacity goes elsewhere for petrol. Cars arrive according to a Poisson distribution at a mean rate of one every 8 min. Their service time has an exponential distribution with a mean of 4 min.

    The proprietor has the opportunity of renting an adjacent piece of land, which would provide space for an additional car to wait (he cannot build another pump). The rent would be Rs. 10 per week. The expected net profit from each customer is Re. 0.50, and the station is open 10 h every day. Would it be profitable to rent the additional space?
19. A barbershop has two barbers and three chairs for waiting customers. Assume that customers arrive in a Poisson fashion at a rate of 5 per hour and that each

barber services customers according to an exponential distribution with a mean of 15 min. Further, if a customer arrives and there are no empty chairs in the shop, he will leave. Find the steady state probabilities. What is the probability that the shop is empty? What is the expected number of customers in the shop?

20. A car servicing station has two bays where service can be offered simultaneously. Due to space limitations, only 4 cars are accepted for servicing. The arrival pattern is Poisson with 12 cars per day. The service time in both bays is exponentially distributed with $\mu = 8$ cars per day per bay. Find the average number of cars in the service station, the average number of cars waiting to be serviced and the average time a car spends in the system.

21. A mechanic repairs four machines. The mean time between service requirements is 5 h for each machine and forms an exponential distribution. The mean repair time is 1 h and also follows the same distribution pattern. Machine downtime costs Rs. 25 per hour, and the mechanic costs Rs. 55 per day. Determine the following:

    (a) The probability that the service facility will be idle
    (b) The probability of various numbers of machines (0–4) to be repaired, and being repaired
    (c) The expected number of machines waiting to be repaired, and being repaired
    (d) The expected downtime cost per day.

    Would it be economical to engage two mechanics, each repairing only two machines?

22. A repairman is to be hired to repair machines which break down at an average rate of 6 per hour. The breakdown follows a Poisson distribution. The non-productive time of a machine is considered to cost Rs. 20 per hour. Two repairmen, Mr. X and Mr. Y, have been interviewed for this purpose. Mr. X charges Rs. 10 per hour, and he services breakdown machines at the rate of 8 per hour. Mr. Y demands Rs. 14 per hour, and he services at an average rate of 12 per hour. Which repairman should be hired? (Assume an 8-h shift per day.)

23. A TV repairman finds that the time spent on his jobs has an exponential distribution with mean 30 min. If he repairs sets in the order in which they come in, and if the arrival of sets is approximately Poisson with an average of 10 per 8-h day, what is the repairman's expected idle time each day? How many jobs are ahead of the average set just brought in?

24. An insurance company has three claim adjusters in its branch office. People with claims against the company are found to arrive in a Poisson fashion at an average rate of 20 per 8-h day. The amount of time that an adjuster spends with a claimant is found to have an exponential distribution with a mean service time of 40 min. Claimants are processed in the order of their appearance. How many hours a week can an adjuster except to spend with claimants?

25. At a port, let there be six unloading berths and four unloading crews. When all the berths are full, arriving ships are diverted to an overflow capacity 20 km

down the river. Tankers arrive according to a Poisson process with a mean of one every 2 h. It takes an unloading crew, on average, 10 h to unload a tanker, the unloading time following an exponential distribution.

(a) On average, how long does a tanker spend at the port?
(b) What is the average arrival rate at the overflow capacity?

26. There is congestion on the platform of a railway station. The trains arrive at the rate of 30 trains per day. The waiting time for any train to hump is exponentially distributed with an average of 36 min. Calculate the following:

(a) The mean queue size
(b) The probability that the queue size exceeds 9.

27. A colliery working one shift per day uses a large number of locomotives which break down at random intervals; on average one fails per 8-h shift. The fitter carries out a standard maintenance schedule on each faulty locomotive. Each of the five main parts of this schedule takes on average $\frac{1}{2}$ h, but the time varies widely. How much time will the fitter have for the other tasks, and what is the average time a locomotive is out of service?
28. If, for a period of 2 h in a day, trains arrive at the yard every 24 min and the service time is 36 min, then calculate for this period:

(a) The probability that the yard is empty
(b) The average queue length on the assumption that the line capacity of the yard is limited to 4 trains only.

29. At what average rate must a clerk at a supermarket work in order to ensure a probability of 0.09 that the customer will not have to wait longer than 10 min? It is assumed that there is only one counter, to which customers arrive in a Poisson fashion at an average of 15 per hour. The length of service by the clerk has an exponential distribution.
30. Customers arrive at a one-window drive-in bank according to a Poisson distribution with mean 10 per hour. The service time per customer is exponential with mean 5 min. The space in front of the window, including that for the serviced car, can accommodate a maximum of three cars. Other cars can wait outside the space.

(a) What is the probability that an arriving customer can drive directly to the space in front of the window?
(b) What is the probability that an arriving customer will have to wait outside the indicated space?
(c) How long is an arriving customer expected to wait before starting service?
(d) How many spaces should be provided in front of the window so that all the arriving customers can wait in front of the window at least 90% of the time?

31. In a self-service facility, arrivals occur according to a Poisson distribution with mean 50 per hour. The service time per customer is exponentially distributed with mean 5 min.
    (a) Find the expected number of customers in service.
    (b) What is the percentage of time the facility is idle?
32. An oil refinery receives crude oil at an average rate of one tanker per day. The unloading facilities, which operate 24 h per day, can handle one tanker at a time, but can unload tankers at an average rate of 2 per day. Under the usual assumptions of Poisson arrival and exponential service times, determine the:
    (a) Average number of tankers in the system
    (b) Average time spent by a tanker in the system
    (c) Average waiting time of a tanker in the queue
    (d) Percentage of time in which exactly two tankers are in the system.
33. A group of engineers has two terminals available to aid in their calculations. The average computing job requires 20 min of terminal time, and each engineer requires some computation about once every 0.5 h; that is, the mean time between a call for service is 0.5 h. Assume these jobs are distributed with an exponential distribution. If there are six engineers in the group, find:
    (i) The expected number of engineers waiting to use one of the terminals
    (ii) The total lost time per day.
34. In a factory cafeteria, the customers (employees) have to pass through three counters. The customers buy coupons at the first counter, select and collect the snacks at the second counter and collect tea at the third. The server at each counter takes on average 1.5 min, although the distribution of service time is approximately Poisson at an average rate of 6 per hour. Calculate:
    a. The average time of getting the service
    b. The average time a customer spends waiting in the cafeteria
    c. The most probable time in getting the service.
35. At a certain airport it takes exactly 5 min to land an aeroplane, once it is given the signal to land. Although incoming planes have scheduled arrival times, the wide variability in arrival times produces an effect which makes the incoming planes appear to arrive in a Poisson fashion at an average rate of 6 per hour. This produces occasional congestion at the airport, which can be dangerous and costly. Under these circumstances, how much time will a pilot expect to spend circling the field waiting to land?
36. In a car manufacturing plant, a loading crane takes exactly 10 min to load a car into a wagon and again come back to position to load another car. If the arrival of cars is a Poisson stream at an average of one every 20 min, calculate the average waiting time of a car.

37. A barber runs his own saloon. It takes him exactly 25 min to complete one haircut. Customers arrive in a Poisson fashion at an average rate of one every 35 min.

    (a) For what percentage of time would the barber be idle?
    (b) What is the average time that a customer spends in the shop?

38. The repair of a lathe requires four steps to be completed one after another in a certain order. The time taken to perform each step follows an exponential distribution with a mean of 5 min and is independent of the other steps. The machine breakdown follows a Poisson process with a mean rate of 2 breakdowns per hour. Answer the following:

    (i) What is the expected idle time of the machine, assuming there is only one repairman available in the workshop?
    (ii) What is the average waiting time of a broken-down machine in the queue?
    (iii) What is the expected number of broken-down machines in the queue?

# Chapter 16
# Flow in Networks

## 16.1 Objectives

The objectives of this chapter are to:

- Present ideas on graphs and their properties
- Study flow and potential in networks
- Discuss the solution procedure of a minimum path problem when all arc lengths are either non-negative or unrestricted in sign
- Present ideas on network flow, maximum flow problems and generalized maximum flow problems
- Discuss the source, the sink of a flow, the cut and its capacity in a network flow problem and also the proof of the max-flow min-cut theorem
- Discuss the solution procedure of maximum flow problems
- Derive the LPP formulation of a network.

## 16.2 Introduction

Network is an important topic in the areas of engineering as well as in applied mathematics. It arises in the different sectors of our daily life, such as transportation, electrical circuits and communication. Its representation is also widely used for problems in diverse areas like production, distribution, project planning, facilities location, resource management and financial planning. In fact, a network representation provides such a powerful visual and conceptual aid for portraying the relationship between the components of the system that it is used in virtually every field of scientific, social and economic endeavours. One of the most exciting developments in operations research in recent years has been the rapid advancement in both the methodology and applications of network optimization models.

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_16

A number of algorithmic breakthroughs have had a major impact, employing ideas from computer science concerning data structure and efficient data manipulation. Consequently, algorithms and software are now available and are being used to solve numerous problems on a routine basis that would have been completely intractable two or three decades ago.

A network, in its more generalized and abstract sense, is called a graph. In recent years, graph theory has been a subject of much study and research by mathematicians and has found ever-increasing applications in diverse areas. In the field of operations research, graph theory plays an important role in finding the optimal solution of the problem of choosing the best sequence of operations from a finite number of alternatives.

## 16.3 Graphs

A graph is a set of points (commonly called vertices or nodes) in space that are interconnected by a set of lines (called edges or arcs). For a graph $G$, the vertex set is denoted by $V$ and the edge set by $E$. Common nomenclature denotes the number of vertices $|V|$ by $n$ and the number of edges $|E|$ by $m$. In Fig. 16.1, a graph $G$ with $V = \{v_1, v_2, v_3, v_4, v_5\}$, $E = \{e_1, e_2, \ldots, e_7\}$, $n = 5$ and $m = 7$ is shown.

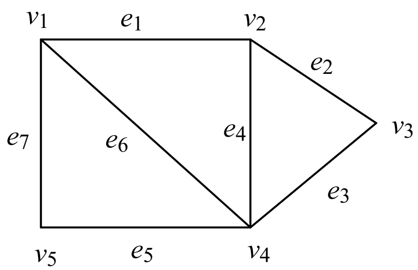
The elements of $E$ are denoted either by $e_k, k = 1, 2, \ldots, m$ or $(v_i, v_j)$. If the values of both $n$ and $m$ are finite, then the graph $G$ is said to be finite. The degree of a vertex $v$ [denoted by $d(v)$] is the number of edges that have $v$ as an end point. For example, in the graph of Fig. 16.1, there are two vertices $v_1$ and $v_2$ of odd degree (both have degree 3).

A self-loop is an edge $(v_i, v_j)$ where $v_i = v_j$. Two edges $(v_i, v_j)$ and $(v_r, v_s)$ are parallel edges if $v_i = v_r$ and $v_j = v_s$.

An arc with an arrow denoting the direction from $v_i$ to $v_j$ is called a directed arc. A graph with directed arcs is called a directed graph or digraph (Fig. 16.2).

A subgraph of $G(V, E)$ is defined as a graph $G_1(V_1, E_1)$ with $V_1 \subseteq V$ and $E_1$ containing all those arcs of $G$ which connect the vertices of $G_1$. For example, in Fig. 16.1, if $V_1 = (v_1, v_2, v_4)$ and $E_1 = (e_1, e_4, e_6)$, then $G_1(V_1, E_1)$ is a subgraph of $G$.

**Fig. 16.1** Graph



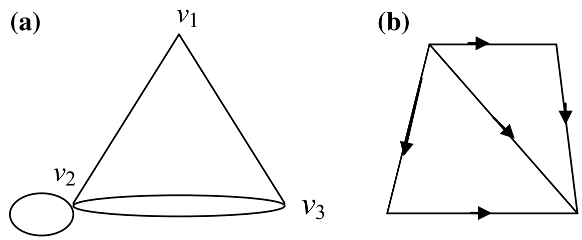



**Fig. 16.2** **a** Undirected graph, **b** directed graph

A partial graph of $G(V, E)$ is a graph $G_2(V_2, E_2)$ which contains all the vertices of $G$ and some of its arcs $(E_2 \subseteq E)$.

An arc (directed or undirected) is said to be incident with a vertex in which it joins to some other vertex; thus, it connects two vertices. The directed arc $e_k = (v_i, v_j)$ is said to be incident from or going from $v_i$ and incident to or going to $v_j$. We call $v_i$ as the initial vertex and $v_j$ as the terminal vertex of the arc $(v_i, v_j)$.

Let $V_1$ and $V_2$ be two subsets of $V$ such that they have no common vertex, and let $e_k = (v_i, v_j)$ be an arc such that $v_i \in V_1,\ v_j \in V_2$. Then $e_k$ is said to be incident from or going from $V_1$ and incident to or going to $V_2$. It is incident with both $V_1$ and $V_2$ and is said to connect them. In Fig. 16.3, if $V_1 = \{v_2, v_3\}$ and $V_2 = \{v_6, v_7\}$, then $e_4$ connects $V_1$ and $V_2$. It runs from $V_1$ to $V_2$ and is incident with both. Now we denote by $\Omega(V_k)$ the set of arcs of $G(V, E)$ incident with a subset $V_k$ of $V$, by $\Omega^+(V_k)$ the set of arcs incident to $V_k$ and by $\Omega^-(V_k)$ the set of arcs incident from $V_k$. Hence,

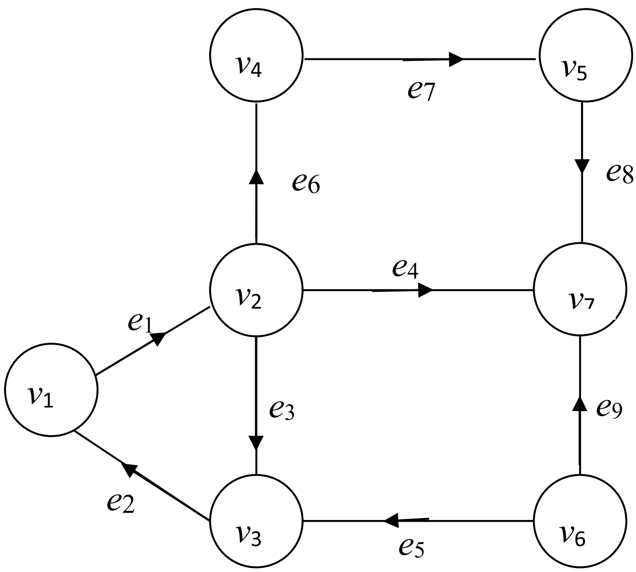



**Fig. 16.3** Directed graph

$$\Omega^+(V_1) = \{e_1, e_5\}, \quad \Omega^-(V_1) = \{e_2, e_4, e_6\}, \;\; \Omega(V_1) = \{e_1, e_2, e_4, e_5, e_6\}.$$

Similarly, $\Omega^+(V_2) = \{e_4, e_8\}$, $\Omega^-(V_2) = \{e_5\}$, $\Omega(V_2) = \{e_4, e_5, e_8\}$.

For a sequence of arcs $(e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_m)$ of a graph, if every intermediate arc $e_k$ has one vertex common with the arc $e_{k-1}$ and another common with $e_{k+1}$, then this sequence is called a chain. For example, the sequence $(e_1, e_6, e_7)$ in Fig. 16.3 is a chain. We may also denote a chain by the vertices which it connects; for example, this chain may also be written as $(v_1, v_2, v_4, v_5)$.

A chain becomes a cycle if, in the sequence of arcs, no arc is used twice, and the first arc has a common vertex with the last arc and this vertex is not common with any intermediate arc. For example, the chain $(e_3, e_4, e_9, e_5)$ in Fig. 16.3 is a cycle.

A path is a chain in which all the arcs are directed in the same sense such that the terminal vertex of the preceding arc is the initial vertex of the succeeding arc. In Fig. 16.3, the sequence of arcs $(e_1, e_6, e_7, e_8)$ is a path. We may also denote the path in terms of the vertices as $(v_1, v_2, v_4, v_5, v_7)$. Note that every path is a chain, but every chain is not a path.

A circuit is a cycle in which all the arcs are directed in the same sense. The cycle $(e_1, e_3, e_2)$ is a circuit.

A graph is said to be connected if for every pair of vertices there is a chain connecting the two. A graph is strongly connected if there is a path connecting every pair of vertices in it.

A tree is defined as a connected graph with at least two vertices and no cycles. It can be proved that a tree with $n$ vertices has $(n - 1)$ arcs and that every pair of vertices is joined by one and only one chain. If we delete an arc from a tree, the resulting graph is not connected, and if we add an arc, a cycle is formed. As the name indicates, a natural tree is the best example of a graphical tree, with the branches forming the arcs and the extremities of the branches forming the vertices (Fig. 16.4).

A vertex which is connected to every other vertex of the graph by a path is called a centre of the graph. A graph may or may not have a centre, or it may have many centres. Every vertex of a strongly connected graph is a centre.

A tree with a centre is called an arborescence. In Fig. 16.5, the centre is marked. In arborescence all the arcs incident with the centre are going away from it, and all other arcs are directed in the same sense.

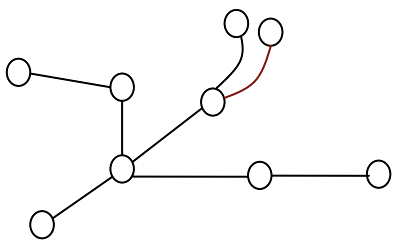


**Fig. 16.4** Tree

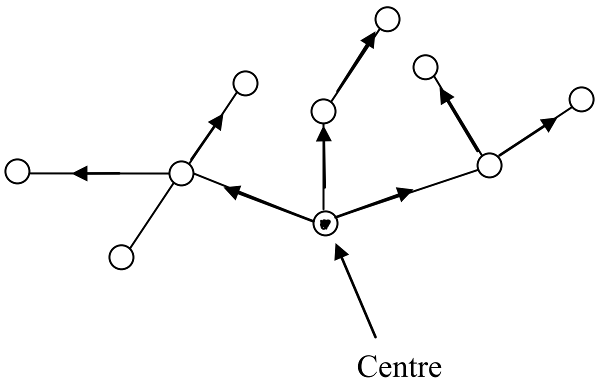



**Fig. 16.5** Graph with centre

## 16.4 Minimum Path Problem

Let us consider a real number $x_{ij}$ which is associated with each arc $\left(v_i, v_j\right)$ of a graph $G(V, E)$, and also let $v_a$ and $v_b$ be two vertices of the graph. There may be a number of paths from $v_a$ to $v_b$. For each path, the length of the path is defined as $\sum x_{ij}$, where the summation is over the sequence of arcs forming the path. In this case, the objective of the problem is to find the path of smallest length.

The term 'length' is used here in a general sense of any real number associated with the arc and should not be regarded as a geometrical distance. It may be distance, time, cost, etc. There may be some abstract situations in which the length is not even non-negative. In general, $x_{ij}$ is a real number, unrestricted in sign. Several techniques/algorithms have been suggested for solving the minimum path problem. Here, we shall describe two different techniques for solving the minimum path problem when (i) $x_{ij} \geq 0$ and (ii) $x_{ij}$ is unrestricted in sign.

**Case I** When $x_{ij} \geq 0$, i.e. when all arc lengths are non-negative.

Let $f_i$ denote the minimum length of the path from the vertex $v_a$ to $v_i$. We have to find $f_b$. Obviously, $f_a = 0$. Let $V_p$ be a subset of $V$ such that $v_a \in V_p$ and $v_b \notin V_p$. First, we have to determine $f_i$ for every $v_i$ in $V_p$ and then $f_i + x_{ij}$ for every $v_i$ in $V_p$ and $v_j$ not in $V_p$ such that $\left(v_i, v_j\right)$ is an arc incident from $V_p$.

$$\text{Let} \quad f_r + x_{rs} = \operatorname*{Min}_i \left\{f_i + x_{ij}\right\},$$
$$\text{where} \quad v_r \in V_p \;\; \text{and} \;\; v_s \notin V_p.$$

Then the length of the minimum path from $v_a$ to $v_s$ is given by $f_s = f_r + x_{rs}$.

Now, from an enlarged subset $V_{p+1}$ of $V$ denoted by $V_{p+1} = V_p \cup \{v_s\}$ we repeat the process.

Suppose we start with $p = 0$ with $V_0$ consisting of a single vertex $v_a$ and $f_a = 0$. Following the procedure, the sets $V_1, V_2, \ldots, V_p, V_{p+1}, \ldots$ are formed. As soon as we arrive at a set in this sequence which includes $v_b$ and $f_b$ has been obtained. If no such set can be found, there is no path connecting $v_a$ to $v_b$.

We note that for solving these types of problems, drawing of the graph is not essential either to describe the problem or to find its solution. The problem is completely enunciated if all the vertices, arcs and arc lengths are specified. In fact, in a large problem with vertices and arcs, drawing of a graph is neither practicable nor necessary.

**Case II:** When $x_{ij}$ is unrestricted in sign, i.e. when arc lengths are unrestricted in sign.

Let $v_a$, $v_b$ be two vertices in the graph $G(V, E)$ whose arc lengths are real numbers, positive, negative or zero. We have to find the minimum path from $v_a$ to $v_b$. Let us assume that there is no circuit in the graph whose arc length is negative. If there is any such circuit, one can go around and around it and decrease the length of the path without limit, getting an unbounded solution.

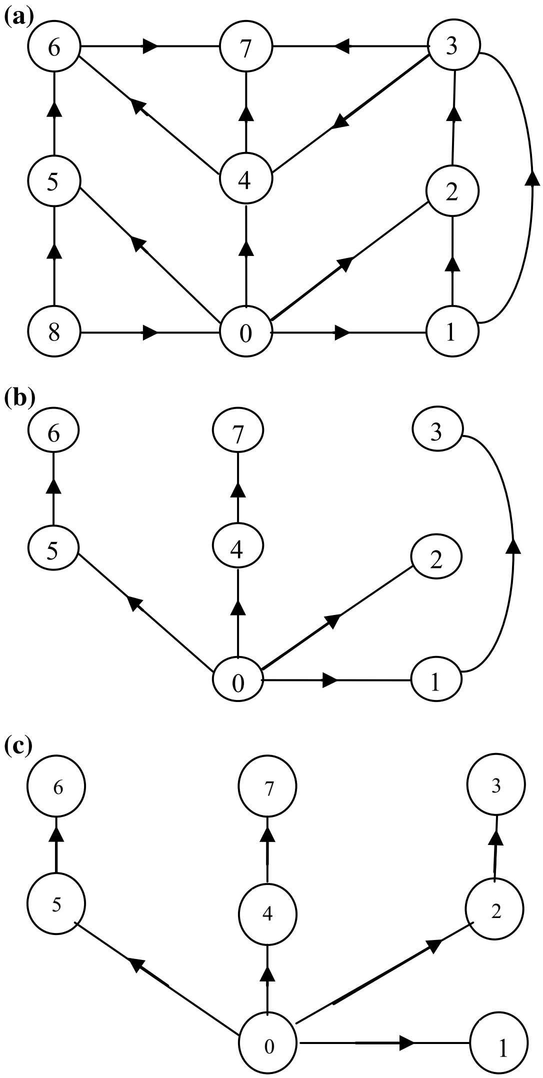
For solving the problem, we have to construct an arborescence $A_1(V_1, E_1)$, $V_1 \subseteq V$, $E_1 \subseteq E$, with centre $v_a$ and $V_1$ containing all the vertices of $V$ which can be reached from $v_a$ along a path and $E_1$ containing from $v_a$ along a path and some of the arcs of $E$ which are necessary to construct the arborescence. If $V_1$ contains $v_b$, a path connects $v_a$ to $v_b$. In a particular arborescence, this path is unique. There may be several arborescences; therefore, several paths exist. If in any problem only one arborescence is possible, there is only one path from $v_a$ to $v_b$, and that is the solution. If $V_1$ does not contain $v_b$, there is no path from $v_a$ to $v_b$, and the problem has no solution.

We now discuss the method of construction of an arborescence. Mark out the arcs going from $v_b$. From the vertices so reached, mark out the arcs (not necessarily all of them) going out to other vertices. No vertex should be reached by more than one arc; that is, not more than one arc should be incident to any vertex. If there is a vertex to which no arc is incident, it cannot be reached from $v_a$ and so is left out. No arc incident to $v_a$ should be drawn (Fig. 16.6).

Let $f_i$ denote the length of the path from $v_a$ to any vertex $v_i$ in the arborescence. The arborescence determines $f_i$ uniquely for each $v_i$ in $V$, but $f_i$ is not necessarily minimum. Let $(v_j, v_i)$ be an arc in $G$ but not in arborescence $A_1$. Then consider the length $f_j + x_{ji}$ and compare it with $f_i$. If $f_i \leq f_j + x_{ji}$, make no change. If $f_i > f_j + x_{ji}$, delete the arc incident to $v_i$ in $A_1$ and include the arc $(v_j, v_i)$. This modifies the arborescence from $A_1$ to $A_2$ and reduces $f_i$ to its new value $f_j + x_{ji}$, the reduction in the value of $f_i$ being $f_i - f_j - x_{ji}$. The lengths of the paths to the vertices going through $v_i$ are also reduced by the same amount. These adjustments are made, and thus the new values of $f_i$ for all $v_i$ in $A_2$ are calculated.

Now repeat the operations in $A_2$; i.e. select a vertex and check whether any alternative arc gives a smaller path to it. If yes, modify $A_2$ to $A_3$ and adjust $f_i$

**Fig. 16.6** Graph with centre




accordingly. Finally, we reach an arborescence $A_r$ which cannot further be changed by the described procedure.

**Theorem-1** *The arborescence $A_r$ marks out the minimum path to each $v_i$ from $v_a$, and $f_b$ in this arborescence is the minimum path to $v_b$.*

*Proof* Let $(v_a, v_1, v_2, \ldots, v_b)$ be any path in $G$ from $v_a$ to $v_b$. Its length is $x_{a1} + x_{12} + \cdots + x_{pb}$. The vertices in this path are in $A_r$ because $A_r$ contains all those vertices of $G$ which can be reached from $v_a$. By the property of $A_r$, for every vertex $v_i$ in $A_r$ and for every arc $(v_j, v_i)$ in $G$,

$$f_i \leq f_j + x_{ji}$$
$$\text{or} \quad f_i - f_j \leq x_{ji}$$

because otherwise $A_r$ could have been further modified. Writing these inequalities for all vertices of the above path, we have

$$f_1 - f_a \leq x_{a1}$$
$$f_2 - f_1 \leq x_{12}$$
$$\dots$$
$$f_b - f_p \leq x_{pb}.$$

Adding, we get

$$f_b - f_a \leq x_{a1} + a_{12} + \cdots + x_{pb}$$
$$\text{or} \quad f_b \leq x_{a1} + a_{12} + \cdots + x_{pb} \quad [\because f_a = 0].$$

Thus, it is proved that no path from $v_a$ to $v_b$ in $G$ can be smaller than $f_b$. Since the path length $f_b$ is also in $G$, this path is the minimum.

**Note** The path of maximum length can be found either by changing the signs of the lengths of all the arcs and then finding the minimum or by reversing the inequality $f_i > f_j + x_{ji}$ to $f_i < f_j + x_{ji}$.

## 16.5 Problem of Potential Difference

Potential plays an important role in electricity, fluid flow and other branches of mechanics. In graph theory, the potential in a network is defined as follows.

For each vertex in a graph $G(V, E)$, there is associated a real number $f_i$ called its potential.

For an arc of a graph, there is an associated potential difference. This potential difference is denoted by $x_{ij}$ for an arc $e_k = (v_i, v_j)$ and is defined by $x_{ij} = f_j - f_i$. Hence, the potential difference in a chain through the vertices $v_1, v_2, \dots, v_p$ is defined as

$$x_{12} + x_{23} + \cdots + x_{p-1,p} = f_p - f_1.$$

It follows immediately that the potential difference in a cycle is zero.

Again, it is obvious that $x_{ij} = -x_{ji}$.

The problem of maximum potential difference in a network is stated as follows:

The potential difference $x_{ab}$ of two vertices $v_a$ and $v_b$ of a graph $G(V, E)$ is to be maximized subject to the conditions $x_{ij} \leq c_{ij}$ for all the arcs $(v_i, v_j)$, where $c_{ij}$ are constants.

In terms of potential, these constraints may be written as

$$f_j - f_i \le c_{ij} \quad \text{for all the arcs } (v_i, v_j).$$

Hence, the problem reduces to the following:

Maximize $f_b - f_a$

subject to the constraints

$$f_j - f_i \le c_{ij} \quad \text{for all arcs } \left(v_i, v_j\right) \text{ of the graph } G(V, E).$$

Now, if we consider the potential of the vertex $v_a$ as zero, then this problem reduces to

Maximize $f_b$

subject to the constraints

$$f_j - f_i \le c_{ij} \quad \text{for all arcs } \left(v_i, v_j\right) \text{of the graph } G(V, E),$$

where $f_i$ is a real-valued function associated with the vertex $v_i$ in a graph with $f_a = 0$.

The problem is identical to the minimum path problem, as the inequalities of this problem are similar to the inequalities of the minimum path problem. The only difference between the two inequalities is that $x_{ij}$ of the minimum path problem is replaced by $c_{ij}$ here. Therefore, to solve the problem of potential difference, we have to solve the problem of minimum path from $v_a$ to $v_b$ with $c_{ij}$ as the length of the arc $(v_i, v_j)$.

A more general problem of maximum potential difference in a network is of the following type:

$$b_{ij} \le f_j - f_i \le c_{ij} \quad \text{for all arcs } \left(v_i, v_j\right).$$

The solution methodology remains the same, as each inequality of the above type can be written as

$$\begin{aligned} f_j - f_i &\le c_{ij} \\ f_i - f_j &\le -b_{ij}. \end{aligned}$$

## 16.6 Network Flow Problem

Like potential, flow in a network is also a familiar concept in electrical theory, flows of traffic through a network of roads, flows of liquid through a network of pipelines, etc. The basic principle of flow in a network is that, for every vertex, the total flow in is equal to the total flow out. To extend this idea in a more abstract situation, it is required to give a precise definition of flow in a graph.

Let a real number $x_k$ be associated with every arc $e_k$, $k$ = 1, 2, …, $m$ of a graph $G$ ($V$, $E$) such that for every vertex $v_j$,

$$\sum x_k = \sum\nolimits_1 x_k, \quad k = 1, 2, \ldots, m,$$

where the left-hand side summation $\sum$ is taken for all arcs going to $v_j$ and the right-hand side summation $\sum_1$ is taken for all arcs going from $v_j$. Then $x_k$ is said to be a flow in the arc $e_k$ and the set $\{x_k\}$, $k$ = 1, 2, …, $m$ is said to be a flow in the graph $G$.

Now we shall define a graph to state the problem of maximum flow in a network.

Let $G(V, E)$ be a graph with $V$ as the set of ($n$ + 2) vertices $v_a$, $v_1$, $v_2$, $v_3$, …, $v_n$, $v_b$ and $E$ as the set of ($m$ + 1) arcs $e_0$, $e_1$, $e_2$, …, $e_m$. The vertices $v_a$ and $v_b$ and the arc $e_0$ play a special role in this graph. $v_a$ is called the source, $v_b$ the sink, and the arc $e_0$ connects $v_b$ to $v_a$. It is the only arc going from $v_b$.

For every arc $e_k$, $k$ = 1, 2, …, $m$ (except $e_0$), there is associated a real number $c_k$ ($\geq 0$) called the capacity of the arc.

Let $\{x_k\}$ be a flow in the graph $G$ such that $0 \leq x_k \leq c_k$ ($k$ = 1, 2, …, $m$), where $c_k$ is the capacity of the flow of the arc $e_k$. In this context, note that $x_0$ as the flow in the arc $e_0$ is defined, but the capacity of the arc $e_0$ is not defined. So there is no constraint on $x_0$.

Since $x_0$ is the only flow out at $v_b$ and flow in at $v_a$, then by the definition of flow, we have

$$\begin{aligned} \text{total flow in at } v_b &= \text{total flow out at } v_b \\ &= x_0 \\ &= \text{total flow in at } v_a \\ &= \text{total flow out at } v_a \end{aligned}$$

Since all the flows out at $v_a$ are equal to the flows in at $v_b$, then $v_a$ and $v_b$ are called the source and sink respectively. The arc $e_0$ plays an important role in the flow network and is considered the return arc.

Hence, our problem is to determine the maximum flow out at the source (=-maximum flow in in the sink). More precisely, our problem is to determine the flow $\{x_k\}$ such that $x_0$ is maximum subject to the constraints $0 \leq x_k \leq c_k$, $k$ = 1, 2, …, $m$.

Now we have an important question: How can we solve the maximum flow problem? The procedure for solving such a problem is given in the following algorithm.

**Algorithm**

Step 1 Set an initial feasible flow by guessing. If this is not possible, initially set $x_k$ = 0 for all $k$.

Step 2 Divide the set $V$ of vertices into two subsets $W_1$ and $W_2$ such that each vertex is either in $W_1$ or in $W_2$ but not in both. Initially, consider $W_1 = \{v_a\}$ and all other vertices in $W_2$.

Step 3 Transfer a vertex from $W_2$ to $W_1$ by the following procedure:
For $v_i \in W_1$, $v_j \in W_2$,

(a) Transfer $v_j$ to $W_1$ if $(v_i, v_j)$ is an arc $e_k$ with $x_k < c_k$;
(b) Transfer $v_j$ to $W_1$ if $(v_j, v_i)$ is an arc $e_k$ with $x_k > 0$;
(c) Do not transfer $v_j$ to $W_1$ otherwise.

Repeat the process for transferring the vertices from $W_2$ to $W_1$. Finally, if $v_b$ is transferred to $W_1$ by this procedure, the flow is not optimal. Go to Step 4.

Step 4 Increase the value of $x_k$ in the arc of category (a) mentioned above in which $x_k < c_k$ and decrease the value of $x_k$ in the arc of category (b) in which $x_k > 0$ so that the flow remains feasible and at least one arc gets the capacity flow. Then go to Step 2.
After repetitive operations of Steps 2, 3 and 4, a situation will appear when $v_b$ cannot be transferred to $W_1$ by the operation of Step 3. In this situation, the flow is optimal.

**Definition** If in the graph $G(V, E)$ of the maximum flow problem, $W_2$ is a subset of $V$ such that $v_b \in \mathrm{W}_2$, $v_a \notin \mathrm{W}_2$, then the set of arcs $\Omega^+(W_2)$ (arcs incident to $W_2$) is said to be a cut. The capacity of the cut is the sum of the capacities of the arcs contained in the cut.

**Theorem 1** *For any feasible flow* $\{x_k\}$, $k = 1, 2, \ldots, m$, *in the graph*, *the flow* $x_0$ *in the return arc is not greater than the capacity of any cut in the graph.*

*Proof* Let $\Omega^+(W_2)$ be any cut. Now, let us consider the flow in the arcs going to and going from $W_2$. According to the definition of flow, the flow in is equal to the flow out.

Therefore,

$$\sum x_k = x_0 + \sum\nolimits_1 x_k,$$

where $\sum$ and $\sum_1$ denote the summations over the arcs going to and going from $W_2$ (except $e_0$).

Since $x_k \geq 0$ for all $k$, then $\sum x_k \geq x_0$.

Also, $x_k \leq c_k$ for all $k$.

Therefore $\sum c_k \geq x_0$, where $\sum c_k$ is the capacity of the cut $\Omega^+(W_2)$. This proves the theorem.

**Theorem 2** *The preceding algorithm solves the problem of the maximum flow.*

*Proof* Let us suppose that by successive applications of the algorithm, a situation will appear when no vertex of $W_2$ can be transferred to $W_1$ by the prescribed procedure and $v_b \in W_2$. Let the set of arcs $\Omega^+(W_2)$ be a cut. Also let $e_k \in \Omega^+(W_2)$. This implies that $e_k$ is an arc $(v_r, v_s)$, where $v_r \in W_1$, $v_s \in W_2$. In this case, the flow in this arc should be saturated, i.e. $x_k = c_k$; otherwise, it would have been possible to transfer $v_s$ from $W_2$ to $W_1$ by the criterion of Step 3(a), which is contrary to the hypothesis.

Again, let $e_i \in \Omega^+(W_2)$, $i \neq 0$. This means that $u_i$ is an arc $(v_p, v_q)$, where $v_p \in W_2$, $v_q \in W_1$. In this case, the flow in this arc should be zero; otherwise, it would have been possible to transfer $v_p$ from $W_2$ to $W_1$ by the criterion of Step 3(b), which is again contrary to the hypothesis.

Therefore, we conclude that the flow into $W_2$ is $\sum c_k$, the summation being over all $e_k \in \Omega^+(W_2)$, and the flow out of $W_2$ is only in the return arc $e_0$ as it is the only arc going from $W_2$ carrying a non-zero flow. Let the flow in $e_0$ be $y_0$. Then, from the definition of flow,

$$\sum c_k = y_0,$$

where $\sum c_k$ is the total capacity of the cut obtained by the application of the algorithm.

But from the theorem 'For any feasible flow $\{x_k\}$, $k = 1, 2, \ldots, m$, in the graph, the flow $x_0$ in the return arc is not greater than the capacity of any cut in the graph.', for any flow $x_0$ in $e_0$,

$$x_0 \leq \sum c_k,$$
$$\text{i.e. } x_0 \leq y_0,$$

where $\sum c_k$ is the capacity of any cut.

It follows that $y_0 = \text{Max } x_0$. This means that the maximal flow problem can be solved by the preceding algorithm.

### Max-Flow Min-Cut Theorem

**Theorem 3** *The maximum flow in a graph is equal to the minimum of the capacities of all possible cuts in it.*

*Proof* Let $v_a$ and $v_b$ be the source and sink of the graph $G(V, E)$ of a flow problem. Also, let $W_2$ be a subset of $V$ such that $v_b \in W_2$, $v_a \notin W_2$. Let $\Omega^+(W_2)$ be any cut.

Let us consider the flow in the arcs going to and going from $W_2$.

Now, from the definition of flow, the flow in should be equal to the flow out,

$$\text{i.e. } \sum x_k = x_0 + \sum\nolimits_1 x_k,$$

where $\sum$ and $\sum_1$ denote the summations over the arcs going to and going from $W_2$ (except $e_0$).

Since for all $k$, $x_k \leq 0$, $\sum x_k \geq x_0$.
Again, $x_k \leq c_k$ for all $k$.
Therefore, $\sum c_k \geq x_0$ where $\sum c_k$ is the total capacity of the cut $\Omega^+(W_2)$,

$$\text{or } x_0 \leq \sum c_k.$$
$$\text{Therefore, } \max x_0 \leq \min \sum c_k.$$

Again, we know that there is a cut corresponding to which the flow in $e_0$ is equal to the cut capacity. Necessarily this flow should be maximum, and the corresponding cut capacity should be the least of all cut capacities.

This proves the theorem.

## 16.7 Generalized Problem of Maximum Flow

The general form of the maximum flow problem is as follows:

Let $G(V, E)$ be a graph consisting of a source $v_a$, a sink $v_b$, vertices $v_i$, $i = 1, 2, \ldots, n$, arcs $e_k$, $k = 1, 2, \ldots, m$ and a return arc $e_0$. Find a flow $\{x_k\}$ in $G$ such that

$$x_0 \text{ is maximum}$$
$$\text{subject to } b_k \leq x_k \leq c_k.$$

Here $x_k$ can vary between an upper bound $c_k$ and a lower bound $b_k$, both of which are real numbers and not necessarily non-negative. Thus, negative flow $x_k$ in an arc is admissible and is interpreted as a flow-$x_k$ in the reverse direction.

The constraints on $x_k$ may be rewritten as

$$\begin{aligned}
&0 \leq x_k - b_k \leq c_k - b_k \\
\text{or}\quad &0 \leq y_k \leq c_k - b_k, \quad \text{where } x_k = y_k + b_k \\
\text{or}\quad &y_k = x_k - b_k.
\end{aligned}$$

In terms of the new variables $y_k$, the constraints are similar to those of the earlier maximum flow problem. However, if $\{x_k\}$ is a flow, $\{y_k\}$ is not necessarily a flow.

From the definition of flow, $\{x_k\}$ is a flow if

$$\sum x_k = \sum\nolimits_1 x_k$$

or if $\sum y_k + \sum b_k = \sum_1 y_k + \sum_1 b_k$ for every vertex $v_j$,

where $\sum$ and $\sum_1$ are summations on arcs going to and going from $v_j$.

From this equation, it is clear that for $\{y_k\}$ to be a flow, the relation $\sum b_k = \sum_1 b_k$ must hold. But there is no reason that this relation should hold, in general.

To overcome this difficulty, a new vertex $v_0$ in $G$, called a fictitious vertex, is introduced, and the necessary number of arcs, called fictitious arcs, are drawn according to the following rules.

**Rule 1**: For an arc $e_k = (v_i, v_j)$ with $b_k > 0$, form a cycle $(v_0, v_j, v_i, v_0)$ by drawing the arcs $(v_0, v_j)$ and $(v_i, v_0)$. Then getting a flow in this cycle, assume a flow $b_k$ in each of the arcs $(v_0, v_j)$, $(v_j, v_i)$, $(v_i, v_0)$.
**Rule 2**: For an arc $e_k = (v_i, v_j)$ with $b_k < 0$, form a cycle $(v_0, v_i, v_j, v_0)$ by drawing the arcs $(v_0, v_i)$ and $(v_j, v_0)$. Then getting a flow in this cycle, assume a flow $-b_k$ in each of the arcs $(v_0, v_i)$, $(v_i, v_j)$, $(v_j, v_0)$.

Due to the inclusion of the fictitious vertex and arcs in $G(V, E)$, we get a new graph $G_1(V_1, E_1)$ with constant flows in the cycles so formed. These flows together form a flow in $G_1$, as it is a linear combination of flows in the cycles. Now, we shall refer to this flow as the fictitious flow in $G_1$.

Let $\{x_k\}$ be any flow in $G$. It is also a flow in $G_1$. Let this flow be superimposed on the fictitious flow. The result is a flow, as it is a linear combination of two flows. Denoting this flow by $\{y_k\}$, we have

$$y_k = \begin{cases} x_k - b_k, \text{for } e_k \text{ in } G \\ \text{the fictitious flow for all fictitious arcs} \end{cases}.$$

Let us now determine the flow $\{y_k\}$ in $G_1$ such that

$$\begin{aligned} &y_0 \text{ is maximum subject to } 0 \le y_k \le c_k - b_k \text{ for } e_k \in G \\ &\text{and} \quad y_k = \text{fictitious flow for fictitious arcs.} \end{aligned}$$

This involves keeping the flow constant in some arcs of $G_1$ and varying it in others until $y_0$ is maximum. This can be done by our previous algorithm. Having determined the optimal flow $\{y_k\}$, we have to calculate $x_k = y_k + b_k$, and finally we get the optimal flow $\{x_k\}$ in $G$.

## 16.8 Formulation of an LPP of a Flow Network

Let us consider the flow network shown in Fig. 16.7.

Also, let $x_0$ be the flow of the return arc from $v_b$ to $v_a$. For different arcs, the flow variables are given in the following table:

| Arc | $(a, 1)$ | $(1, 2)$ | $(1, b)$ | $(2, 1)$ | $(2, 3)$ | $(2, 4)$ | $(2, b)$ | $(3, a)$ | $(3, 4)$ | $(4, b)$ | $(b, 2)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $x_9$ | $x_{10}$ | $x_{11}$ |

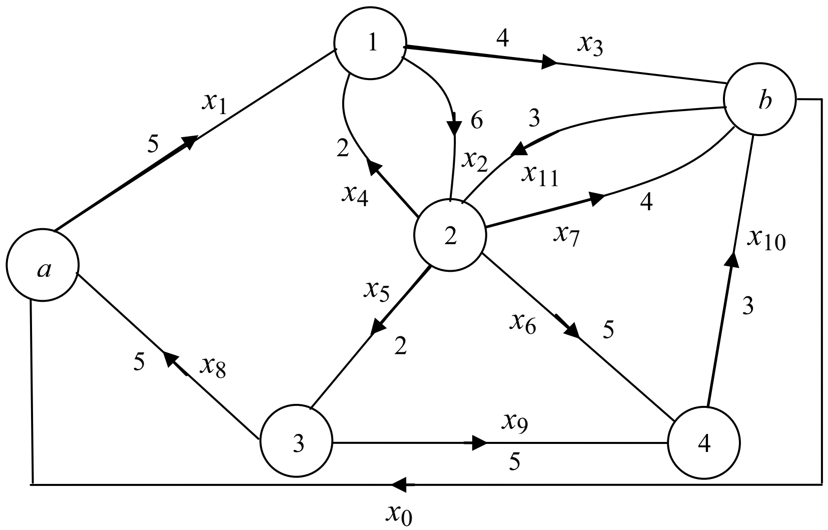



**Fig. 16.7** Representation of network flow

Here $(v_i, v_j)$ is denoted by $(i, j)$.

From the definition of flow, we know that the flow into the vertex is equal to the flow out from the vertex. Hence, for different vertices, the constraints are as follows:

For vertex $v_1$,

$$x_1 + x_4 = x_2 + x_3$$
$$\Rightarrow x_1 - x_2 - x_3 + x_4 = 0.$$

For vertex $v_2$,

$$x_2 + x_{11} = x_4 + x_5 + x_6 + x_7$$
$$\Rightarrow x_2 - x_4 - x_5 - x_6 + x_{11} = 0.$$

For vertex $v_3$,

$$x_5 = x_8 + x_9$$
$$\Rightarrow x_5 - x_8 - x_9 = 0.$$

For vertex $v_4$,

$$x_6 + x_9 = x_{10}$$
$$\Rightarrow x_6 + x_9 - x_{10} = 0.$$

For vertex $v_a$,

$$x_0 + x_8 = x_1$$
$$\Rightarrow x_0 - x_1 + x_8 = 0.$$

For vertex $v_b$,

$$x_3 + x_7 + x_{10} = x_0 + x_{11}$$
$$\Rightarrow x_0 + x_{11} - x_3 - x_7 - x_{10} = 0.$$

For the network, the capacity constraints are as follows:
$x_1 \le 5, x_2 \le 6, x_3 \le 4, x_4 \le 2, x_5 \le 2, x_6 \le 5, x_7 \le 4, x_8 \le 5, x_9 \le 5, x_{10} \le 3, x_{11} \le 3$ and $x_i \ge 0$; $i$ = 1, 2, 3, …, 11.

Hence, the LPP of the given network is as follows:

Maximize $f = x_0$

subject to

$$\begin{aligned}
x_1 - x_2 - x_3 + x_4 &= 0\\
x_2 - x_4 - x_5 - x_6 + x_{11} &= 0\\
x_5 - x_8 - x_9 &= 0\\
x_6 + x_9 - x_{10} &= 0\\
x_0 - x_1 + x_8 &= 0\\
x_0 - x_3 - x_7 - x_{10} + x_{11} &= 0
\end{aligned}$$

$x_1 \le 5$, $x_2 \le 6$, $x_3 \le 4$, $x_4 \le 2$, $x_5 \le 2$, $x_6 \le 5$, $x_7 \le 4$, $x_8 \le 5$, $x_9 \le 5$, $x_{10} \le 3$, $x_{11} \le 3$, $x_i \ge 0$; $i$ = 1, 2, 3, …, 11 and $x_0$ is unrestricted in sign.

**Example 1** Find the minimum path from $v_0$ to $v_8$ in the graph (Fig. 16.8) in which the number along a directed arc denotes its length.

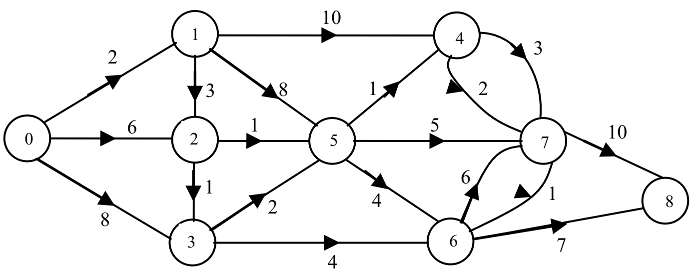



**Fig. 16.8** Representation of directed flow network

**Solution** Let $V$ be the set of all vertices of the given graph and $V_p$ a subset of $V$. Let $x_{ij}$ be the arc length of an arc $(v_i, v_j)$ and $f_i$ be the minimum length of the path from the vertex $v_0$ to a vertex $v_i$ of $V_p$. Also, let $f_s$ be the minimum length of the path from $V_p$ to another vertex $v_s$, not in $V_p$, through an arc incident from $V_p$,

$$\text{i.e. } f_s = \min_i \{f_i + x_{ij}\} \text{for all } v_i \in V_p \text{ and } (v_i, v_j) \in \Omega^-(V_p),$$

where $\Omega^-(V_p)$ is the set of arcs incident from $V_p$. We denote $(v_i, v_j)$ as $(i, j)$. Now to find the minimum path, we construct Table 16.1.

From Table 16.1, it is seen that the minimum path is either 0–1–2–3–6–8 or 0–1–2–5–6–8 and the length of each path is 17.

**Example 2** Find the minimum path from $v_0$ to $v_7$ in the graph of Fig. 16.9.

**Solution** From the graph, it is clear that there is no circuit whose length is negative.

Now, we draw an arborescence $A_1$ (see Fig. 16.10a) with centre $v_0$ consisting of all those vertices of the graph which can be reached from $v_0$ and the necessary number of arcs.

From the graph, it is also clear that a number of arborescences can be drawn. $A_1$ is one of them.

Now the lengths $f_i$ of the paths from $v_0$ to different vertices $v_i$ of $A_1$ are as follows:

$$f_0 = 0, \quad f_1 = 1, \;\; f_2 = -4, \;\; f_3 = -4 - 1 = -5, \;\; f_4 = 3,$$
$$f_7 = 3 + 2 = 5, \;\; f_5 = -2, \;\; f_6 = -2 + 2 = 0.$$

Let us consider the vertex $v_2$. There is an arc $(v_1, v_2)$ in $G$ which is not in $A_1$ such that

$$f_2 = -4 < f_1 + x_{12} = 1 + 2 = 3.$$

So, $A_1$ will be unchanged.

Now we consider the vertex $v_3$. There is an arc $(v_1, v_3)$ in $G$ which is not in $A_1$ such that

$$f_3 = -5 < f_1 + x_{13} = 1 + 2 = 3.$$

So, $A_1$ will be unchanged.

Now, for $v_4$ in $A_2$, arc $(v_3, v_4)$ is in $G$, but not in $A_2$ such that

$$f_4 = 3 > f_3 + x_{34} = -5 - 4 = -9.$$

So we delete the arc $(v_0, v_4)$, include $(v_3, v_4)$, get another arborescence $A_2$ (see Fig. 16.10b) with $f_4 = -9$ and consequently $f_7 = -9 + 2 = -7$.

For $v_5$ in $A_2$, arc $(v_4, v_5)$ is in $G$ but not in $A_2$ such that

**Table 16.1** Calculation of minimum path

| $p$ | $V_p$ | $f_i$ | $\Omega^{-}(V_p)$ | $x_{ij}$ | $f_i + x_{ij}$ | $f_s$ | $v_s$ | Minimum path up to $v_s$ |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | (0, 1) | 2 | 2 | 2 | 1 | 0–1 |
| | | | (0, 2) | 6 | 6 | | | |
| | | | (0, 3) | 8 | 8 | | | |
| 1 | 0 | 0 | (0, 2) | 6 | 6 | | | |
| | | | (0, 3) | 8 | 8 | | | |
| | 1 | 2 | (1, 2) | 3 | 5 | 5 | 2 | 0–1–2 |
| | | | (1, 4) | 10 | 12 | | | |
| | | | (1, 5) | 8 | 10 | | | |
| 2 | 0 | 0 | (0, 3) | 8 | 8 | | | |
| | 1 | 2 | (1, 4) | 10 | 12 | | | |
| | | | (1, 5) | 8 | 10 | | | |
| | 2 | 5 | (2, 3) | 1 | 6 | 6 | 3 | 0–1–2–3 |
| | | | (2, 5) | 1 | 6 | 6 | 5 | 0–1–2–5 |
| 3 | 0 | 0 | | | | | | |
| | 1 | 2 | (1, 4) | 10 | 12 | | | |
| | 2 | 5 | | | | | | |
| | 3 | 6 | (3, 6) | 4 | 10 | | | |
| | 5 | 6 | (5, 4) | 1 | 7 | 7 | 4 | 0–1–2–5–4 |
| | | | (5, 6) | 4 | 10 | | | |
| | | | (5, 7) | 5 | 11 | | | |
| 4 | 0 | 0 | | | | | | |
| | 1 | 2 | | | | | | |
| | 2 | 5 | | | | | | |
| | 3 | 6 | (3, 6) | 4 | 10 | 10 | 6 | 0–1–2–3–6 |
| | 4 | 7 | (4, 7) | 3 | 10 | 10 | 7 | 0–1–2–5–4–7 |
| | 5 | 6 | (5,6) | 4 | 10 | 10 | 6 | 0–1–2–5–6 |
| | | 6 | (5,7) | 5 | 11 | | | |
| 5 | 0 | 0 | | | | | | |
| | 1 | 2 | | | | | | |
| | 2 | 5 | | | | | | |
| | 3 | 6 | | | | | | |
| | 4 | 7 | | | | | | |
| | 5 | 6 | | | | | | |
| | 6 | 10 | (6, 8) | 7 | 17 | 17 | 8 | 0–1–2–3–6–8 |
| | | | | | | | | 0–1–2–5–6–8 |
| | 7 | 10 | (7, 8) | 10 | 20 | | | |

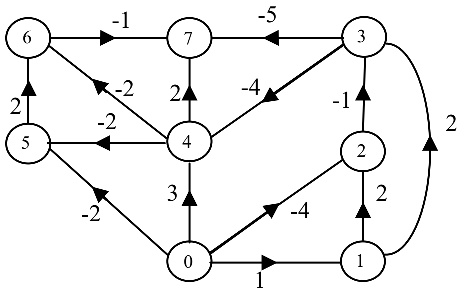



**Fig. 16.9** Graph with known edge

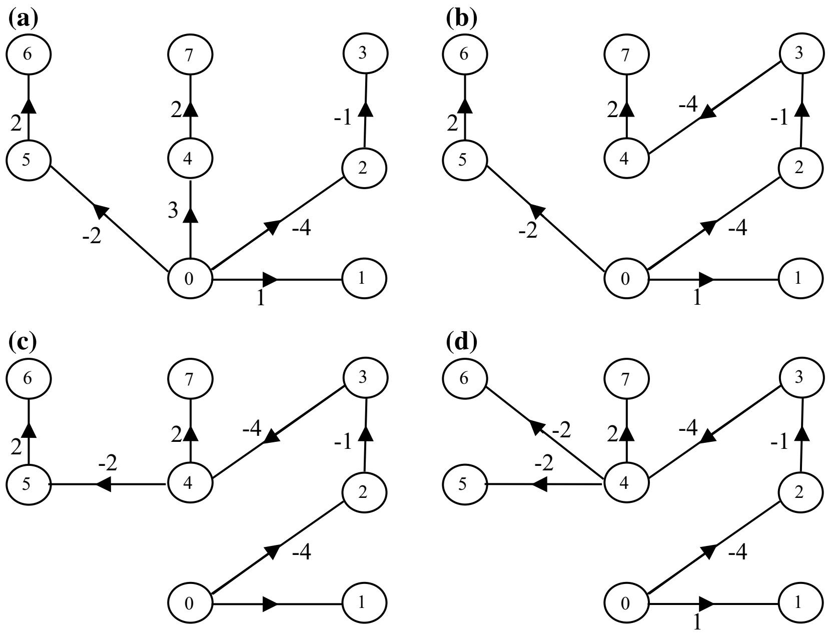


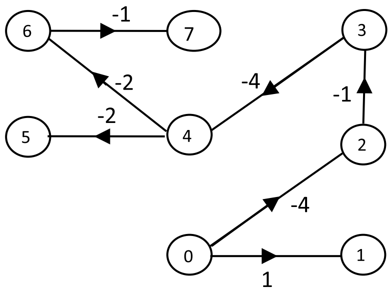

**Fig. 16.10** **a** Arborescence $A_1$, **b** arborescence $A_2$, **c** arborescence $A_3$, **d** arborescence $A_4$

$$f_5 = -2 > f_4 + x_{45} = -9 - 2 = -11.$$

So we delete $(v_0, v_5)$, include $(v_4, v_5)$, get another arborescence $A_3$ (see Fig. 16.10c) with $f_5 = -11$ and consequently $f_6 = -11 + 2 = -9$.

For $v_6$ in $A_3$, arc $(v_4, v_6)$ is in $G$ but not in $A_3$ such that

**Fig. 16.11** Arborescence $A_5$




$$f_6 = -9 > f_4 + x_{46} = -9 - 2 = -11.$$

So we delete $(v_5, v_6)$, include $(v_4, v_6)$ and get another arborescence $A_4$ (see Fig. 16.10d) with $f_6 = -11$.

For $v_7$ in $A_4$, arc $(v_6, v_7)$ is in $G$ but not in $A_4$ such that

$$f_7 = -7 > f_6 + x_{67} = -11 - 1 = -12.$$

So we delete $(v_4, v_7)$, include $(v_6, v_7)$ and get another arborescence $A_5$ (see Fig. 16.11) with $f_7 = -12$.

Again, for $v_7$ in $A_5$, arc $(v_3, v_7)$ is in $G$ but not in $A_5$ such that $f_7 = -12 < f_3 + x_{37} = -5 - 5 = -10$.
So, we leave $A_5$ unchanged.

From the above arborescence $A_5$, it is observed that $A_5$ cannot further be modified. Therefore, no alternative arc decreases the length of the path from $v_0$ to any vertex. Hence, the minimum path from $v_0$ to $v_7$ contains the vertices $v_0, v_2, v_3, v_4, v_6, v_7$ with length $-12$.

**Example 3** Find the minimum path from $v_1$ to $v_8$ in the graph with arc lengths as follows [here $(i, j)$ denotes the arc $(v_i, v_j)$]:

| Arc | (1, 2) | (1, 3) | (1, 4) | (2, 3) | (2, 6) | (2, 5) | (3, 5) | (3, 4) |
|---|---|---|---|---|---|---|---|---|
| Length | −1 | 4 | −11 | 2 | −8 | 7 | −3 | 7 |

| Arc | (4, 7) | (5, 6) | (5, 8) | (6, 3) | (6, 4) | (6, 7) | (6, 8) | (7, 3) | (7, 8) |
|---|---|---|---|---|---|---|---|---|---|
| Length | 3 | 1 | 12 | 4 | 2 | 6 | – 10 | –2 | 2 |

**Solution** For the given problem, the corresponding graph $G$ is given in Fig. 16.12.

From the graph, it is clear that there is no circuit whose length is negative. Now, we draw an arborescence $A_1$ (see Fig. 16.13) with centre $v_1$ consisting of all those vertices of the graph which can be reached from $v_1$ and the necessary arcs. Now the lengths $f_i$ of the paths from $v_1$ to different vertices $v_i$ of $A_1$ are as follows:

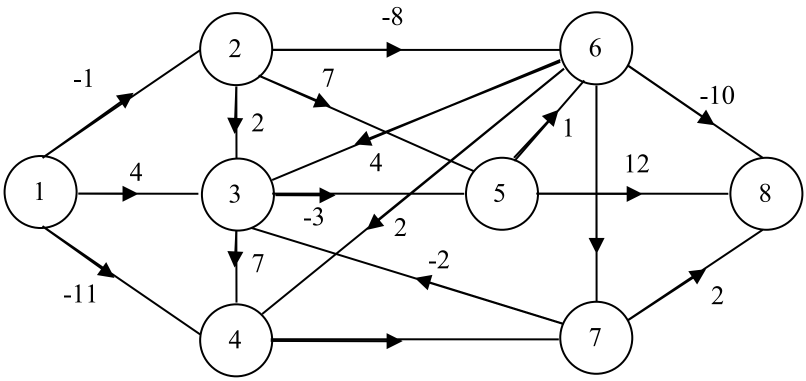
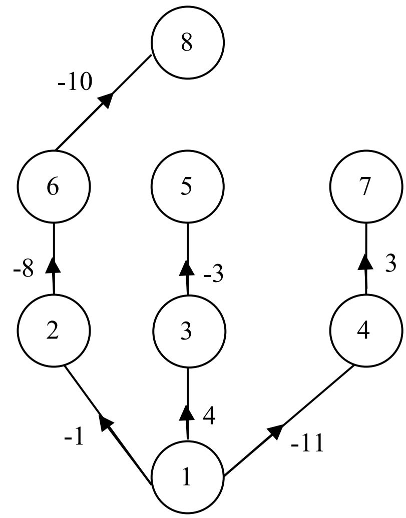



**Fig. 16.12** Graph

**Fig. 16.13** Arborescence $A_1$

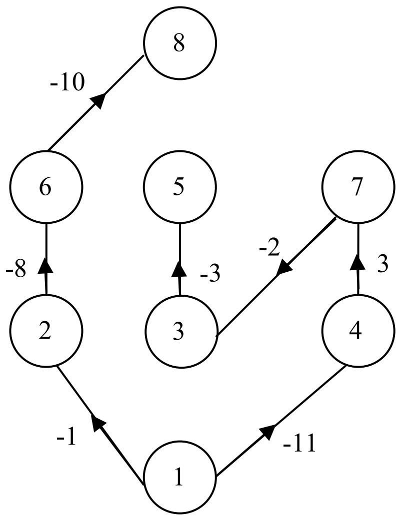



$$f_1 = 0, \quad f_2 = -1, \; f_3 = 4, \; f_4 = -11, \; f_5 = 1, \; f_6 = -9, \; f_7 = -8, \; f_8 = -19$$

Let us consider the vertex $v_3$. This vertex is connected by the arcs $(v_1, v_3)$, $(v_2, v_3)$, $(v_6, v_3)$ and $(v_7, v_3)$. Among these, arcs $(v_2, v_3)$, $(v_6, v_3)$ and $(v_7, v_3)$ are not in the arborescence $A_1$ such that

$$\begin{aligned} f_3 = 4 > f_2 + x_{23} &= -1 + 2 = 1 && \text{for the arc}(v_2, v_3) \\ f_3 = 4 > f_6 + x_{63} &= -9 + 4 = -5 && \text{for the arc}(v_6, v_3) \\ \text{and } f_3 = 4 > f_7 + x_{73} &= -8 - 2 = -10 && \text{for the arc}(v_7, v_3). \end{aligned}$$

**Fig. 16.14** Arborescence $A_2$




Hence, we delete the arc $(v_1, v_3)$, include $(v_7, v_3)$, get another arborescence $A_2$ (see Fig. 16.14) with $f_3 = -10$ and consequently $f_5 = -10 - 3 = -13$.

Again for the vertex $v_4$ in $A_2$, arcs $(v_3, v_4)$ and $(v_6, v_4)$ are in $G$ but not in $A_2$ such that

$$f_4 = -11 < f_3 + x_{34} = -10 + 7 = -3 \quad \text{for arc } (v_3, v_4)$$
$$\text{and } f_4 = -11 < f_6 + x_{64} = -9 + 2 = -7 \quad \text{for arc } (v_6, v_4).$$

So, $A_2$ will not be changed.

Now, for $v_5$ in $A_2$, arc $(v_2,\ v_5)$ is in $G$ but not in $A_2$ such that

$$f_5 = -13 < f_2 + x_{25} = -1 + 7 = 6.$$

So $A_2$ will not be changed.

Again, for $v_6$ in $A_2$, arc $(v_5, v_6)$ is in $G$ but not in $A_2$ such that

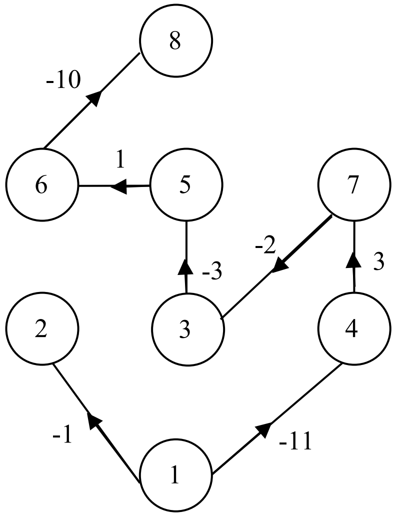
$$f_6 = -9 > f_5 + x_{56} = -13 + 1 = -12.$$

So we delete the arc $(v_2, v_6)$, include $(v_5, v_6)$, get another arborescence $A_3$ (see Fig. 16.15) with $f_6 = -12$ and consequently $f_8 = -12 - 10 = -22$.

Now for vertex $v_7$ in $A_3$, arc $(v_6, v_7)$ is in $G$ but not in $A_3$ such that

$$f_7 = -8 < f_6 + x_{67} = -12 + 6 = -6.$$

**Fig. 16.15** Arborescence $A_3$




So we leave $A_3$ unchanged.

Again for vertex $v_8$ in $A_3$, arcs $(v_5, v_8)$ and $(v_7, v_8)$ are in $G$ but not in $A_3$ such that

$$\begin{aligned} & f_8 = -22 < f_5 + x_{58} = -13 + 12 = -1 \\ \text{and} \quad & f_8 = -22 < f_7 + x_{78} = -8 + 2 = -6. \end{aligned}$$

So we leave $A_3$ unchanged.

From the arborescence $A_3$, it is seen that the arborescence $A_3$ cannot further be modified. Therefore, no alternative arc decreases the length of the path from $v_1$ to any other vertex. Hence the minimum path from $v_1$ to $v_8$ is $1-4-7-3-5-6-8$ with length $-22$.

*Example 4* Find the maximum potential difference $x_{14}$ between $v_1$ and $v_4$ of the graph with the following data subject to the condition that for each arc $x_{jk} \leq c_{jk}$:

| | | | | | | |
|---|---|---|---|---|---|---|
| $V$ | 1 | 2 | 3 | 4 | | |
| $E$ | $(1,2)$ | $(1,3)$ | $(2,3)$ | $(2,4)$ | $(3,4)$ | $(1,4)$ |
| $c_{jk}$ | 3 | 2 | $-2$ | 1 | 4 | $-1$ |

**Solution** Treating the values of $c_{jk}$ as the arc lengths of a graph, we construct the graph in Fig. 16.16.

Now we draw an arborescence $A_1$ with centre $v_1$ consisting of all those vertices of the graph which can be reached from $v_1$ and the necessary number of arcs (Fig. 16.17).

Now the lengths $f_j$ of the paths from $v_1$ to different vertices $v_j$ of $A_1$ are as follows:

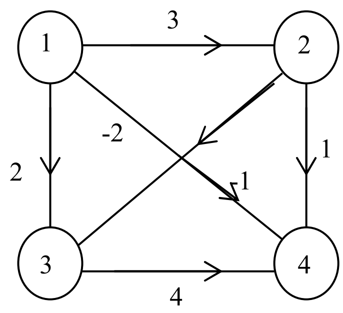



**Fig. 16.16** Graph of the problem of Example 4

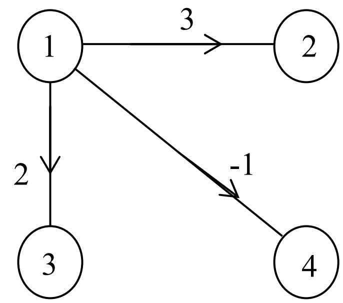



**Fig. 16.17** Arborescence $A_1$

$$f_1 = 0, \quad f_2 = 3, \;\; f_3 = 2, \;\; f_4 = -1$$

Now, for the vertex $v_4$ in $A_1$, arcs $(v_2, v_4)$ and $(v_3, v_4)$ are in $G$, but not in $A_1$ such that

$$f_4 = -1 < f_2 + c_{24} = 3 + 1 = 4$$
$$\text{and} \quad f_4 = -1 < f_3 + c_{34} = 2 + 4 = 6.$$

So we leave $A_1$ unchanged.

Hence, the maximum potential difference $x_{14} = -1$ with the optimal path $(v_1, v_4)$.

*Example 5* Find the maximum potential difference between $v_1$ and $v_4$ in the graph $G(V, E)$, where

| | | | | | | |
|---|---|---|---|---|---|---|
| $V$ | 1 | 2 | 3 | 4 | | |
| $E$ | $(1,2)$ | $(1,3)$ | $(2,3)$ | $(3,4)$ | $(4,2)$ | $(1,4)$ |

subject to the constraints

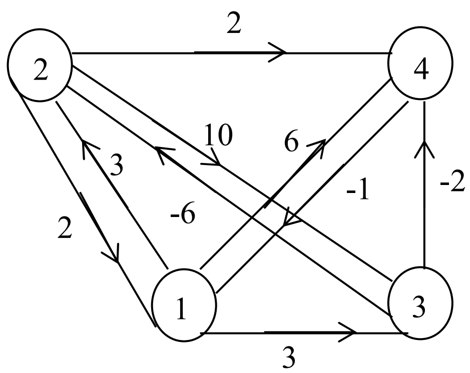



**Fig. 16.18** Graph of the problem of Example 5

$$-2 \leq f_2 - f_1 \leq 3, \quad 6 \leq f_3 - f_2 \leq 10, \; f_4 - f_3 \leq -2$$
$$-2 \leq f_2 - f_4, \quad 1 \leq f_4 - f_1 \leq 6, \; f_3 - f_1 \leq 7.$$

**Solution** The given constraints can be written as follows:

$$f_2 - f_1 \leq 3, \quad f_1 - f_2 \leq 2, \; f_3 - f_2 \leq 10, \; f_2 - f_3 \leq -6$$
$$f_4 - f_3 \leq -2, \quad f_4 - f_2 \leq 2, \; f_4 - f_1 \leq 6, \; f_1 - f_4 \leq -1, \; f_3 - f_1 \leq 7$$

For the given problem, the graph is given in Fig. 16.18.

The student may easily solve the problem. For this problem, the maximum potential difference between $v_1$ and $v_4$ is 3.

**Example 6** In the graph of Fig. 16.19, the numbers along the arcs are the values of $c_i$. Find the maximum flow in the graph.

For solving the given problem, let us assume that the initial flow is zero in each arc. Let $W_1 = \{v_a\}$ and $W_2$ be the set of other vertices. Now, for an arc $(v_a, v_1)$, $v_{av} \in W_1$, $v_I \in W_2$, the flow is zero, which is less than its capacity. So we transfer $v_1$ to $W_1$. Again, for an arc $(v_1, v_5)$, $v_1 \in W_1$, $v_5 \in W_2$, the flow is less than its capacity 4. So we transfer $v_5$ to $W_1$. For the arc $(v_5, v_b)$, $v_5 \in W_1$, $v_b \in W_2$, the flow is zero, which is less than its capacity. So $v_b$ is transferred to $W_1$. As a result, the solution is not optimal. In this process, the corresponding chain is $(v_a, v_1, v_5, v_b)$. The least capacity in this chain is 3. So, in each arc of this chain and also in the return arc $(v_b, v_a)$, we increase the flow in such a way that the flow of at least one arc coincides with its capacity, keeping the flows as same in all other arcs.

The modified flow is feasible, as in each arc it is either less than or equal to its capacity, and also at every vertex the total flow in is equal to the total flow out. The above process is repeated for every modified flow until it is not possible to transfer $v_b$ to $W_I$. The flows of all iterations are shown in Table 16.2. In each feasible flow the number in parentheses indicates the corresponding arc forming the chain in which $v_b$ is transferred to $W_1$. The asterisk indicates that the flow in the corresponding arc is equal to its capacity and will not further be increased.

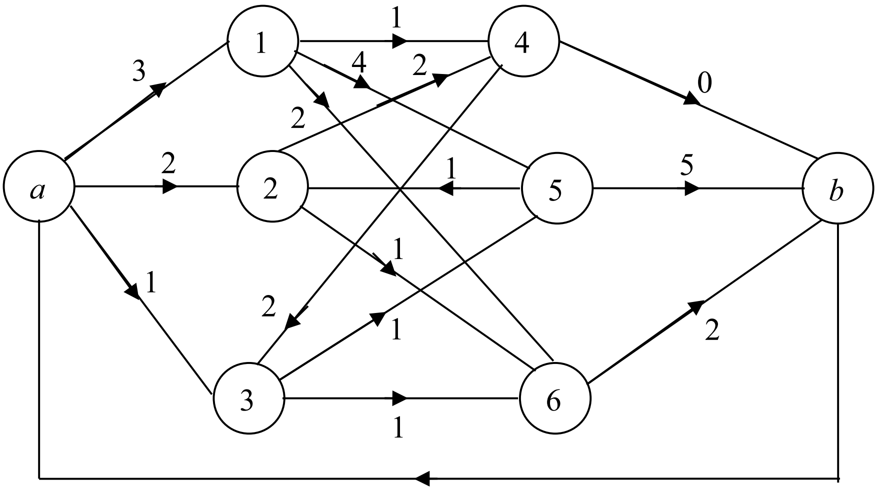



**Fig. 16.19** Flow in graph

**Table 16.2** Solution of flow problem of Example 6

| Arcs | Capacity ($c_i$) | Feasible flows | | | | | |
|---|---|---|---|---|---|---|---|
| | | I | II | III | IV | V | VI |
| ($a$, 1) | 3 | (0) | (1) | 3* | 3* | 3* | 3* |
| ($a$, 2) | 2 | 0 | 0 | (0) | (1) | 2* | 2* |
| ($a$, 3) | 1 | 0 | 0 | 0 | 0 | (0) | 1* |
| (1, 4) | 1 | (0) | 1* | 1* | 1* | (1*) | 0 |
| (1, 5) | 4 | 0 | (0) | 2 | 2 | (2) | 3 |
| (1, 6) | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| (2, 4) | 2 | 0 | 0 | 0 | (0) | 1 | 1 |
| (2, 6) | 1 | 0 | 0 | (0) | 1* | 1* | 1* |
| (3, 5) | 1 | (0) | 1* | 1* | 1* | 1* | 1* |
| (3, 6) | 1 | 0 | 0 | 0 | (0) | 1* | 1* |
| (4, 3) | 2 | (0) | 1 | 1 | (1) | (2) | 1 |
| (4, b) | 0 | 0* | 0* | 0* | 0* | 0* | 0* |
| (5, 2) | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| (5, $b$) | 5 | (0) | (1) | 3 | 3 | (3) | 4 |
| (6, $b$) | 2 | 0 | 0 | (0) | (1) | 2* | 2* |
| ($b$, $a$) | – | 0 | 1 | 3 | 4 | 5 | 6 |

From Table 16.2, it is seen that the maximum flow in the graph is 6.

**Note** The preceding problem can be solved in different ways of vertex transfer from $W_2$ to $W_1$. Two different ways of solving the problem have been shown in Tables 16.3 and 16.4.

**Example 7** Find the maximum flow in the graph with the following arcs and flow capacities. Arc $(v_j, v_k)$ is denoted as $(j, k)$. $v_a$ is the source and $v_b$, the sink.

| Flow | Arc | $(a, 1)$ | $(a, 2)$ | $(a, 3)$ | (1, 4) | (1, 5) | (1, 6) | (2, 4) | (2, 5) |
|---|---|---|---|---|---|---|---|---|---|
| | Capacity | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |

| Flow | Arc | (2, 6) | (3, 4) | (3, 5) | (3, 6) | $(4, b)$ | $(5, b)$ | $(6, b)$ |
|---|---|---|---|---|---|---|---|---|
| | Capacity | 1 | 1 | 1 | 1 | 2 | 2 | 2 |

For the given problem, first of all we draw the graph, which is shown in Fig. 16.20.

For solving the given problem, let us assume that the initial flow is zero in each arc. Let $W_1 = \{v_a\}$ and $W_2$ be the set of other vertices. Now, for an arc $(v_a, v_1)$, $v_a \in W_1$, $v_1 \in W_2$, the flow is zero, which is less than its flow capacity. So we transfer $v_1$ to $W_1$. Again, for an arc $(v_1, v_4)$, $v_1 \in W_1$, $v_4 \in W_2$, the flow is less than its flow capacity. So we transfer $v_4$ to $W_1$. For the arc $(v_4, v_b)$, $v_4 \in W_1$, $v_b \in W_2$, the flow is zero, which is less than its flow capacity. So $v_b$ is transferred to $W_1$. As a

**Table 16.3** Alternative solution of flow problem of Example 6

| Arcs | Capacity ($c_i$) | Feasible flows | | | | |
|---|---|---|---|---|---|---|
| | | I | II | III | IV | V |
| $(a, 1)$ | 3 | (0) | 3* | 3* | 3* | 3* |
| $(a, 2)$ | 2 | 0 | (0) | (1) | 2* | 2* |
| $(a, 3)$ | 1 | 0 | 0 | 0 | (0) | 1* |
| (1, 4) | 1 | 0 | 0 | 0 | 0 | 0 |
| (1, 5) | 4 | (0) | 3 | 3 | 3 | 3 |
| (1, 6) | 2 | 0 | 0 | 0 | 0 | 0 |
| (2, 4) | 2 | 0 | 0 | (0) | 1 | 1 |
| (2, 6) | 1 | 0 | (0) | 1* | 1* | 1* |
| (3, 5) | 1 | 0 | 0 | (0) | 1* | 1* |
| (3, 6) | 1 | 0 | 0 | 0 | (0) | 1* |
| (4, 3) | 2 | 0 | 0 | (0) | 1 | 1 |
| $(4, b)$ | 0 | 0* | 0* | 0* | 0* | 0* |
| (5, 2) | 2 | 0 | 0 | 0 | 0 | 0 |
| $(5, b)$ | 5 | (0) | 3 | (3) | 4 | 4 |
| $(6, b)$ | 2 | 0 | (0) | 1 | (1) | 2* |
| $(b, a)$ | – | 0 | 3 | 4 | 5 | 6 |

**Table 16.4** Alternative solution of flow problem of Example 6

| Arcs | Capacity ($c_i$) | Feasible flows | | | | |
|---|---|---|---|---|---|---|
| | | I | II | III | IV | V |
| ($a$, 1) | 3 | 0 | 0 | 0 | (0) | 3* |
| ($a$, 2) | 2 | (0) | (1) | 2* | 2* | 2* |
| ($a$, 3) | 1 | 0 | 0 | (0) | 1* | 1* |
| (1, 4) | 1 | 0 | 0 | 0 | 0 | 0 |
| (1, 5) | 4 | 0 | 0 | 0 | (0) | 3 |
| (1, 6) | 2 | 0 | 0 | 0 | 0 | 0 |
| (2, 4) | 2 | 0 | (0) | 1 | 1 | 1 |
| (2, 6) | 1 | (0) | 1* | 1* | 1* | 1* |
| (3, 5) | 1 | 0 | 0 | (0) | 1* | 1* |
| (3, 6) | 1 | 0 | (0) | 1* | 1* | 1* |
| (4, 3) | 2 | 0 | (0) | 1 | 1 | 1 |
| (4, $b$) | 0 | 0* | 0* | 0* | 0* | 0* |
| (5, 2) | 2 | 0 | 0 | 0 | 0 | 0 |
| (5, $b$) | 5 | 0 | 0 | (0) | (1) | 4 |
| (6, $b$) | 2 | (0) | (1) | 2* | 2* | 2* |
| ($b$, $a$) | – | 0 | 1 | 2 | 3 | 6 |

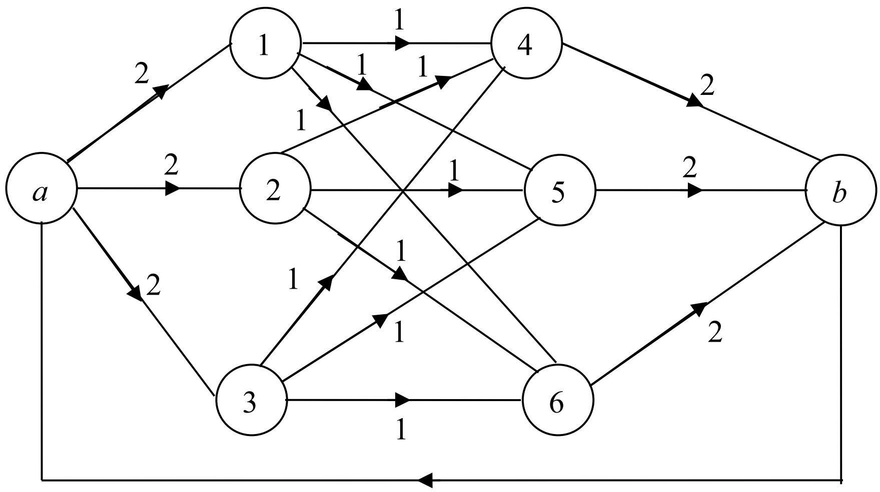



**Fig. 16.20** Maximum flow and capacity

result, the solution is not optimal. In this process, the corresponding chain is ($v_a$, $v_1$, $v_4$, $v_b$). The least flow capacity in this chain is 1. So, in each arc of this chain and also in the return arc ($v_b$, $v_a$), we increase the flow in such a way that the flow of at

**Table 16.5** Solution of flow problem of Example 7

| Arcs | Capacity | Feasible flows | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | I | II | III | IV | V | VI | VII |
| $(a, 1)$ | 2 | (0) | (1) | 2* | 2* | 2* | 2* | 2* |
| $(a, 2)$ | 2 | 0 | 0 | (0) | (1) | 2* | 2* | 2* |
| $(a, 3)$ | 2 | 0 | 0 | 0 | 0 | (0) | (1) | 2* |
| (1, 4) | 1 | (0) | 1* | 1* | 1* | 1* | 1* | 1* |
| (1, 5) | 1 | 0 | (0) | 1* | 1* | 1* | 1* | 1* |
| (1, 6) | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| (2, 4) | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| (2, 5) | 1 | 0 | 0 | (0) | 1* | 1* | 1* | 1* |
| (2, 6) | 1 | 0 | 0 | 0 | (0) | 1* | 1* | 1* |
| (3, 4) | 1 | 0 | 0 | 0 | 0 | 0 | (0) | 1* |
| (3, 5) | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| (3, 6) | 1 | 0 | 0 | 0 | 0 | (0) | 1* | 1* |
| $(4, b)$ | 2 | (0) | 1 | 1 | 1 | 1 | (1) | 2* |
| $(5, b)$ | 2 | 0 | (0) | (1) | 2* | 2* | 2* | 2* |
| $(6, b)$ | 2 | 0 | 0 | 0 | (0) | (1) | 2* | 2* |
| $(b, a)$ | | 0 | 1 | 2 | 3 | 4 | 5 | 6 |

least one arc coincides with its flow capacity, keeping the flows as same in all other arcs.

The modified flow is feasible, as in each arc it is either less than or equal to its flow capacity, and at every vertex the total flow in is equal to the total flow out. The above process is repeated for every modified flow until it is not possible to transfer $v_b$ to $W_1$. The flows of all iterations are shown in Table 16.5. In each feasible flow the number in parentheses indicates the corresponding arc forming the chain in which $v_b$ is transferred to $W_1$. The asterisk indicates that the flow in the corresponding arc is equal to its capacity and will not further be increased.

From Table 16.5, it is seen that the maximum flow in the graph is 6.

**Example 8** Find the maximal flow in the graph of Fig. 16.21 with the constraints

$$2 \leq x_1 \leq 10, \quad 4 \leq x_2 \leq 12, \quad -2 \leq x_3 \leq 4, \quad 0 \leq x_4 \leq 5, \quad 0 \leq x_5 \leq 10.$$

**Solution** Since the lower bounds of the first three variables $x_1$, $x_2$, $x_3$ are non-zero real numbers, we introduce a fictitious vertex $v_0$.

Now, for $e_1$, $b_1 = 2$. Therefore, we draw two fictitious arcs $(v_a, v_1)$ and $(v_a, v_0)$ and assume a flow 2 in each of the arcs of the cycle $(v_0, v_1, v_a, v_0)$. See Fig. 16.22. Similarly, we draw the arcs $(v_0, v_2)$ and $(v_a, v_0)$ and assume a flow 4 in each of the arcs of the cycle $(v_0, v_a, v_2, v_0)$. Since $b_3 = -2$, we draw the arcs $(v_0, v_2)$ and $(v_1, v_0)$ and assume a flow 2 in each of the arcs of the cycle $(v_0, v_2, v_1, v_0)$. In the other arcs $e_4$ and $e_5$, we assume the initial flow as zero.

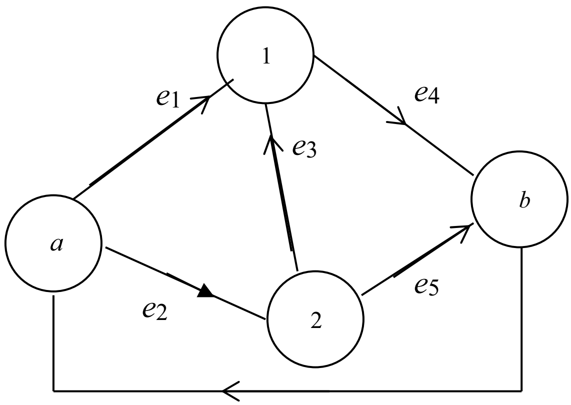



**Fig. 16.21** Graph with constraints

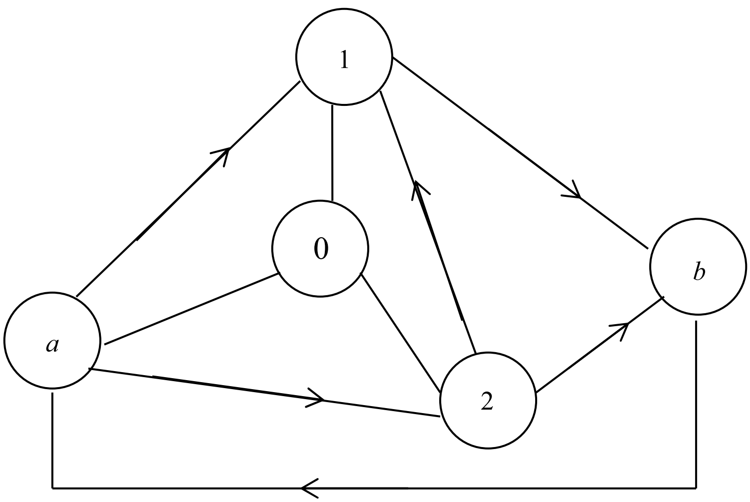



**Fig. 16.22** Maximum flow in modified graph

Then the initial flow is as follows:

| Arcs | $(v_a, v_0)$ | $(v_0, v_2)$ | $(v_0, v_1)$ | $(v_a, v_2)$ | $(v_a, v_1)$ | $(v_2, v_1)$ | $(v_1, v_b)$ | $(v_2, v_b)$ |
|---|---|---|---|---|---|---|---|---|
| Flow | 6 | 6 | 0 | -4 | -2 | 2 | 0 | 0 |

Now expressing the constraints in terms of $y_i$ (where $y_i = x_i - b_i$), we have $0 \le y_1 \le 8,\ 0 \le y_2 \le 8,\ 0 \le y_3 \le 6,\ 0 \le y_4 \le 5,\ 0 \le y_5 \le 10.$

**Table 16.6** Solution of flow problem of Example 8

| Arcs | Capacity ($c_i$) | Feasible flows | | |
|---|---|---|---|---|
| | | I | II | III |
| $(a, 0)$ | – | 6 | 6 | 6 |
| $(a, 1)$ | 8 | (−2) | 3 | 3 |
| $(a, 2)$ | 8 | −4 | (−4) | 6 |
| (0, 1) | – | 0 | 0 | 0 |
| (0, 2) | – | 6 | 6 | 6 |
| $(1, b)$ | 5 | (0) | 5* | 5* |
| (2, 1) | 6 | 2 | 2 | 2 |
| $(2, b)$ | 10 | 0 | (0) | 10* |
| $(b, a)$ | – | 0 | 5 | 15* |

Keeping the flow in the fictitious arcs as same, we shall apply the algorithm for finding the maximum flow in the modified graph. The iterations for determining the modified flow are shown in Table 16.6.

From Table 16.6, we get the optimal solution for the modified graph as

$$y_1 = 3, \quad y_2 = 6, \quad y_3 = 2, \quad y_4 = 5, \quad y_5 = 10.$$

Hence, $x_1 = y_1 + 2 = 5$, $x_2 = y_2 + 4 = 10$, $x_3 = -2 + y_3 = 0$, $x_4 = y_4 = 5$, $x_5 = y_5 = 10$ and the maximum flow = 15, satisfying the constraints of the original graph.

**Note** For this problem, there is also an alternative solution. This solution is $y_1 = 5$, $y_2 = 4$, $y_3 = 0$, $y_4 = 5$, $y_5 = 10$; i.e. $x_1 = 7$, $x_2 = 8$, $x_3 = -2$, $x_4 = 5$, $x_5 = 10$.

The negative flow −2 in the arc $(v_2, v_1)$ can be interpreted as the positive flow 2 in the arc $(v_1, v_2)$.

## 16.9 Exercises

1. Formulate the minimal cost and flows in a network. Hence, find the minimal cost of the network shown in Fig. 16.23, where, for $(c_{ij}, d_{ij})$, $c_{ij}$ represents the flow capacity from vertex $i$ to vertex $j$, and $d_{ij}$ represents the cost per unit flow from vertex $i$ to vertex $j$.
2. Formulate the LPP of the network $G$ shown in Fig. 16.24.
3. Find the maximum flow in the graph with the following arcs and arc capacities, where the flow in each arc is non-negative. Arc $(v_j, v_k)$ is denoted as $(j, k)$. $v_a$ is the source and $v_b$, the sink.

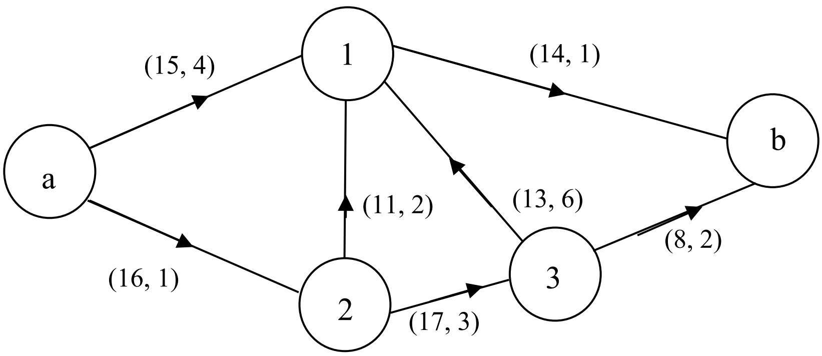



**Fig. 16.23** Flow network

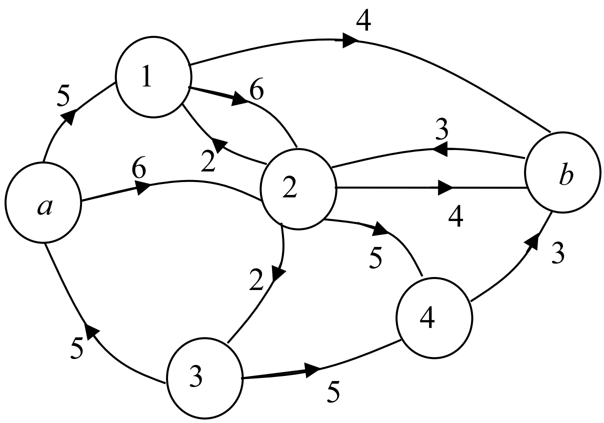



**Fig. 16.24** Flow network

| Arc | $(a, 1)$ | $(a, 2)$ | $(a, 3)$ | (1, 4) | (1, 5) | (1, 6) | (2, 4) | (2, 6) |
|---|---|---|---|---|---|---|---|---|
| Capacity | 3 | 2 | 1 | 1 | 4 | 2 | 2 | 1 |
| Arc | (3, 5) | (3, 6) | (4, 3) | $(4, b)$ | (5, 2) | $(5, b)$ | $(6, b)$ | |
| Capacity | 1 | 1 | 2 | 0 | 2 | 5 | 2 | |

4. Find the maximum flow in the graph with the following arcs and arc capacities, where the flow in each arc is non-negative. Arc ($v_j$, $v_k$) is denoted as ($j$, $k$). $v_a$ is the source and $v_b$ is the sink.

| Arc | ($a$, 1) | ($a$, 2) | ($a$, 3) | (1, 4) | (1, 5) | (1, 6) | (2, 4) | (2, 5) | (2, 6) |
|---|---|---|---|---|---|---|---|---|---|
| Capacity | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| Arc | (3, 4) | (3, 5) | (3, 6) | (4, $b$) | (5, $b$) | (6, $b$) | | | |
| Capacity | 1 | 1 | 1 | 2 | 2 | 2 | | | |

5. Find the maximum non-negative flow in the network described below. Arc ($v_j$, $v_k$) is denoted as ($j$, $k$). $v_a$ is the source and $v_b$ is the sink.

| Arc | ($a$, 1) | ($a$, 2) | (1, 2) | (1, 3) | (1, 4) | (2, 4) | (3, 2) | (3, 4) | (4, 3) |
|---|---|---|---|---|---|---|---|---|---|
| Capacity | 8 | 10 | 3 | 4 | 2 | 8 | 3 | 4 | 2 |
| Arc | (3, $b$) | (4, $b$) | | | | | | | |
| Capacity | 10 | 9 | | | | | | | |

6. Find the maximum flow in the network with the following data, where the flow in arcs is not necessarily non-negative. The arc ($v_j$, $v_k$) is denoted as ($j$, $k$), and the flow limit ($b$, $c$) means that the constraint on the flow $x_i$ is $b_i \leq x_i \leq c_i$. $v_a$ is the source and $v_b$ is the sink.

| Arc | ($a$, 1) | ($a$, 2) | (1, 2) | (1, 3) | (2, 4) | (3, 4) | (3, $b$) | (4, $b$) |
|---|---|---|---|---|---|---|---|---|
| ($b_i$, $c_i$) | (0, 10) | (0, 5) | (−2, 3) | (7, 10) | (−3, 5) | (−1, 1) | (0, 8) | (0, 4) |

# Chapter 17
# Inventory Control Theory

## 17.1 Objectives

The objectives of this chapter are to:

- Discuss the concepts of inventory as well as the various forms of inventory and reasons for maintaining inventories
- Define the different terminologies of inventory like demand, replenishment, constraints, shortages, lead time, planning/time horizon and various types of inventory costs
- Derive the single item purchasing and manufacturing inventory models with and without shortages
- Determine the optimal order quantity for single item price break inventory models and multi-item purchasing model with different constraints like investment, average inventory and space constraints
- Introduce the probabilistic models for discrete (the well-known newspaper boy problem) and continuous cases
- Discuss the fuzzy inventory model.

## 17.2 Introduction

In a broad sense, inventory is defined as an idle resource of an enterprise/company/manufacturing firm. It can be defined as a stock of physical goods, commodities or other economic resources which are used to meet customer demand or production requirements. This means that the inventory acts as a buffer stock between a supplier and a customer.

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_17

The inventory or stock of goods may be kept in any one of the following forms:

(i) Raw materials
(ii) Semi-finished goods (work-in-process inventory)
(iii) Finished (or produced) goods
(iv) Maintenance, repair and operating (MRO) supply items.

In any sector of an economy, the control and maintenance of inventory is a common problem to all organizations. Inventories of physical goods are maintained in government and non-government establishments, e.g. agriculture, industry, military, business, etc.

Some reasons for maintaining inventories are as follows:

(i) To conduct the smooth and efficient running of business
(ii) To provide customers service by meeting their demands from stock without delay
(iii) To avail price discounts for bulk purchasing
(iv) To maintain more stable operations and/or workforce
(v) To take financial advantage of transporting/shipping economics
(vi) To plan an overall operating strategy through decoupling of successive stages in the chain of acquiring goods, preparing products, shipping to branch warehouses and finally serving customers
(vii) To motivate customers to purchase more by displaying a large number of goods in the showroom/shop
(viii) To take advantage in purchasing of some raw materials and some commonly used physical goods (such as paddy, wheat, etc.) whose prices seasonally fluctuate. In this context, it is more profitable to procure a sufficient quantity of these raw materials/commonly used physical goods when their prices are low to be used later during the high price season or when need arises.

Production/inventory planning and control is essentially concerned with the design, operation and control of an inventory system in any sector of a given economy. The problem of inventory control is primarily concerned with the following fundamental questions:

(i) Which items should be carried in stock? Which items should be produced?
(ii) How many of each of these items should be ordered/produced?
(iii) When should an order be placed? Or when should an item be produced?
(iv) What type of inventory control system should be used?

In practice, it is a formidable task to determine a suitable inventory policy. Regarding these questions, an inventory problem is a problem of making optimal decisions. In other words, an inventory problem deals with decisions that optimize either the cost function (total or average cost) or the profit function (total or average profit) of the inventory system. However, there are certain types of problems, such as those relating to the storage of water in a dam, in which one has no control over the replenishment of inventory. The supply of inventory of water in a dam depends on rainfall; the organization operating the dam has no control over it.

This chapter presents mathematical models of different inventory systems. However, this type of model is based on different assumptions and approximations. It is difficult both to devise and operate with an exact/accurate model as nobody knows what the real situation is. Therefore, it is almost impossible to construct a realistic model with accuracy. For this reason, some approximations and simplifications must be considered during the model formulation. The solution of the inventory problem is a set of specific values of variables that minimizes the total (or average) cost of the system or maximizes the total (or average) profit of the system.

### 17.2.1 Types of Inventories

There are different types of inventories, namely: (i) transportation inventories, (ii) fluctuation inventories, (iii) anticipation inventories, (iv) decoupling inventories and (v) lot-size inventories.

*Transportation inventories*

These arise due to transportation of inventory items to various distribution centres and customers from the various production centres. When the transportation time is long, the items under transport cannot be served to customers. These inventories exist solely because of transportation time.

*Fluctuation inventories*

These have to be carried because sales and production times cannot be predicted accurately. In real-life problems, there are fluctuations in the demand and lead times that affect the production of items.

*Anticipation inventories*

These build up in advance by anticipating or foreseeing the future demand for the season of large sales due to a promotion programme or a plant shutdown period.

*Decoupling inventories*

These inventories are used to reduce the interdependence of various stages of the production system.

*Lot-size inventories*

Generally, the rate of consumption is different from the rate of production or purchasing. Therefore, items are produced in larger quantities which result in lot-size inventories, also called cycle inventories.

## 17.3 Basic Terminologies in Inventory

The inventory system depends on several system factors and parameters such as demand, replenishment rate, shortages, constraints, various types of costs, etc.

***Demand***

Demand is defined as the number of units of an item required by the customer in a unit time and has the dimension of a quantity. It may be known exactly or in terms of probabilities, or it may be completely unknown.

The demand pattern of items may be either deterministic or probabilistic. Problems in which demand is known and fixed are called deterministic problems. Problems in which the demand is assumed to be a random variable are called stochastic or probabilistic problems.

In the case of deterministic demand, it is assumed that the quantities needed over subsequent periods of time are known exactly. Further, the known demand may be fixed or variable with time, stock level or selling price of an item, etc.

Probabilistic demand occurs when requirements over a certain period of time are not known with certainty, but their pattern can be described by a known probability distribution.

In some cases, demand may also be represented by uncertain data in a non-stochastic sense, i.e. by vague/imprecise data. This type of demand is termed fuzzy demand, and the system is called a fuzzy system.

***Replenishment***

Replenishment refers to the amount of quantities that are scheduled to be put into inventories, at the time when decisions are made about ordering these quantities or at the time when they are actually added to stock. Replenishment can be categorized according to size, pattern and lead time. Replenishment size may be constant or variable, depending upon the type of inventory system. It may depend on time, demand and/or on-hand inventory level. The replenishment patterns are usually instantaneous, uniform or in batch. The replenishment quantity again may be probabilistic or fuzzy in nature.

***Constraints***

Constraints are the limitations imposed on the inventory system. They may be imposed on the amount of investment, available space, the amount of inventory held, average instantaneous expenditure, number of orders, etc.

***Fully backlogged/partially backlogged shortages***

During a stock-out period, sales or goodwill may be affected either by a delay or by complete refusal in meeting the demand. If the unfulfilled demand for the goods is satisfied completely at a later date, then it is a case of a fully backlogged shortage. That is, it is assumed that no customer balks during this period, and the demand of all the waiting customers is met at the beginning of the next period gradually after the commencement of the next production.

Generally, it is observed that during the stock-out period some of the customers wait for the product and others balk away. When this happens, the phenomenon is called a partially backlogged shortage.

***Lead time***

The time gap between the time of placing an order or production start and the time of arrival of goods in stock is called the lead time. It may be a constant or a variable. Again, variable lead time may be probabilistic or imprecise.

***Planning/time horizon***

The time period over which the inventory level will be controlled is called the time/planning horizon. It may be finite or infinite depending upon the nature of the inventory system for the commodity.

***Deterioration/damageability/perishability***

Deterioration is defined as decay, evaporation, obsolescence or loss of utility or marginal value of a commodity that results in a decreasing usefulness from the original condition. Vegetables, foodgrains and semiconductor chips are examples of products that can deteriorate.

Damageability is defined by the damage when items are broken or lose their utility due to accumulated stress, bad handling, a hostile environment, etc. The amount of damage by stress varies with the size of the stock and the duration for which the stress is applied. Items made of glass, china, clay, ceramic, mud, etc. are examples of damageable products.

Perishable items are those which have a finite lifetime (fixed or random). Fixed lifetime products (e.g. human blood) have a deterministic shelf life, while the random lifetime scenario is closely related to the case of an inventory which experiences continuous physical depletion due to deterioration or decay.

***Various types of inventory costs***

Inventory costs are the costs associated with the operation of an inventory system and result from action or lack of action on the part of management in establishing the system. They are basic economic parameters to any inventory decision model.

*Purchase or unit cost*

The purchase or unit cost of an item is the unit purchase price of obtaining the item from an external source or the unit production cost for the internal production. It may also depend upon the demand. When production is done in large quantities, it results in reduction of production costs per unit. Also, when quantity discounts are allowed for bulk orders, the unit price is reduced and dependent on the quantity purchased or ordered.

*Ordering/setup cost*

The ordering or setup cost originates from the experience of issuing a purchase order to an outside supplier or from internal production setup costs. The ordering cost includes clerical and administrative costs, telephone charges, telegrams, transportation costs, loading and unloading costs, etc. Generally, this cost is assumed to be independent of the quantity ordered or produced. In some cases, it may depend on the quantity of goods purchased because of price breaks quantity discounts or transportation costs.

*Holding or carrying cost*

The holding or carrying cost is the cost associated with the storage of the inventory until its use or sale. It is directly proportional to the amount/quantity in the inventory and the time for which the stocks are held. This cost generally includes costs related to insurance, taxes, obsolescence, deterioration, renting of warehouse, light, heat, maintenance and interest on the money locked up.

*Shortage cost or stock-out cost*

The shortage cost or stock-out cost is the penalty incurred for being unable to meet a demand when it occurs. This cost arises due to shortage of goods, lost sales for delay in meeting the demand or total inability to meet the demand. In the case where the unfulfilled demand for the goods can be satisfied at a later date (backlogging case), this cost depends on both the shortage quantity and the delaying time. On the other hand, if the unfulfilled demand is lost (no backlogging case), the shortage cost becomes proportional to the shortage quantity only. In both cases, there is a loss of goodwill, which cannot be quantified for the development of the mathematical model.

*Disposal cost*

When an amount of some units of an item remains in excess at the end of the inventory cycle, and if this amount is sold at a lower price in the next cycle to derive some advantages like clearing the stock or winding up the business, the revenue earned through such a process is called the disposal cost.

*Salvage values*

During storage, some units are partially spoiled or damaged; i.e. some units lose their utility partially. In a developing country, it is normally observed that some of these units are sold at a reduced price (less than the purchase price) to a section of customers, and this gives some revenue to management. This revenue is called the salvage value.

## 17.4 Classification of Inventory Models

Inventory problems (models) may be classified into two categories.

(i) **Deterministic inventory models**

These are the inventory models in which demand is assumed to be known constant or variable (dependent on time, stock level, selling price of the item, etc.). Here, we shall consider deterministic inventory models for known constant demand. Such models are usually referred to as economic lot-size models or economic order quantity (EOQ) models.

There are different types of models under this category, namely

(a) Purchasing inventory model with no shortage
(b) Manufacturing inventory model with no shortage
(c) Purchasing inventory model with shortages
(d) Manufacturing model with shortages
(e) Multi-item inventory model
(f) Price break inventory model.

(ii) ***Probabilistic inventory models***

These are the inventory models in which the demand is a random variable with a known probability distribution. Here, the future demand is determined by collecting data from past experience.

## 17.5 Purchasing Inventory Model with No Shortage (Model 1)

In this model, we want to derive the formula for the optimum order quantity per cycle of a single product so as to minimize the total average cost under the following assumptions and notation:

(i) Demand is deterministic and uniform at a rate $D$ units of quantity per unit time.
(ii) Production is instantaneous (i.e. production rate is infinite).
(iii) Shortages are not allowed.
(iv) The lead time is zero.
(v) The inventory planning horizon is infinite, and the inventory system involves only one item and one stocking point.
(vi) Only a single order is placed at the beginning of each cycle, and the entire lot is delivered in one batch.
(vii) The inventory carrying cost $C_1$ per unit quantity per unit time and the ordering cost $C_3$ per order are known and constant.
(viii) $T$ is the cycle length and $Q$ is the ordering quantity per cycle.

Let us assume that an enterprise purchases an amount of $Q$ units of an item at time $t = 0$. This amount will be depleted to meet customer demand. Ultimately, the stock level reaches zero at time $t = T$. The inventory situation is shown in Fig. 17.1.

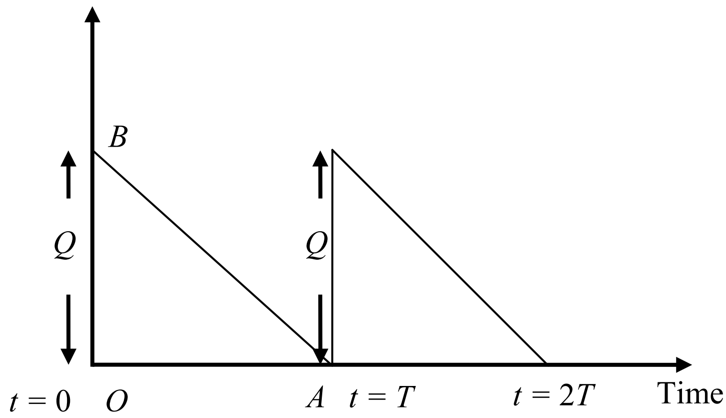



**Fig. 17.1** Pictorial representation of purchasing inventory model with no shortage

$$\text{Clearly,} \quad Q = DT \tag{17.1}$$

Now, the inventory carrying cost for the entire cycle $T$ is $C_1 \times$ (area of $\Delta$AOB) = $\frac{1}{2}C_1QT$, and the ordering cost for the said cycle $T$ is $C_3$.

Hence, the total cost for time $T$ is given by

$$X = C_3 + \frac{1}{2}C_1QT.$$

Therefore, the total average cost is given by

$$C(Q) = \frac{X}{T}$$

$$\begin{aligned} &\text{or} \quad C(Q) = \frac{C_3}{T} + \frac{1}{2}QC_1 \\ &\text{or} \quad C(Q) = \frac{C_3D}{Q} + \frac{1}{2}C_1Q \quad \left[\because\ Q = DT \ \therefore\ T = \frac{Q}{D}\right]. \end{aligned} \tag{17.2}$$

The optimum value of $Q$ which minimizes $C(Q)$ is obtained by equating the first derivative of $C(Q)$ with respect to $Q$ to zero,

$$\begin{aligned} &\text{i.e.,}\ \frac{\mathrm{d}C}{\mathrm{d}Q} = 0 \quad \text{or,}\ \frac{1}{2}C_1 - \frac{C_3D}{Q^2} = 0 \\ &\text{or}\ Q = \sqrt{\frac{2C_3D}{C_1}}. \end{aligned}$$

Now, $\frac{\mathrm{d}^2C(Q)}{\mathrm{d}Q^2} = \frac{2C_3D}{Q^3}$, which is positive for $Q = \sqrt{\frac{2C_3D}{C_1}}$.

Hence, $C(Q)$ is minimum for which the optimum value of $Q$ is

$$Q^* = \sqrt{\frac{2C_3D}{C_1}}. \tag{17.3}$$

This is known as either the economic lot-size formula or the EOQ formula. The corresponding optimum time interval is $T^* = \frac{Q^*}{D} = \sqrt{\frac{2C_3}{C_1D}}$, and the minimum cost per unit time is given by $C_{\min} = \frac{C_3D}{Q^*} + \frac{1}{2}C_1Q^* = \sqrt{2C_1C_3D}$.

This model was first developed by Ford Harris (1915) of the Westinghouse Corporation, USA, in the year 1915. He derived the well-known classical lot-size Formula (17.3). However, according to Erlenkotter (1989), the earliest model was formulated by Harris in the year 1913. This formula was also developed independently by R. H. Wilson after a few years; thus, it has been called the Harris-Wilson formula.

*Remark*

(i) The total inventory time units for the entire cycle $T$ is $\frac{1}{2}QT$, so the average inventory at any time is $\frac{1}{2}QT/T = \frac{1}{2}Q$.

(ii) Since $C_1 > 0$, from $f(Q) = \frac{1}{2}C_1Q$ it is obvious that the inventory carrying cost is a linear function of $Q$ with a positive slope; i.e. for a smaller average inventory, the inventory carrying costs are lower. In contrast, $g(Q) = \frac{C_3D}{Q}$; i.e. the ordering cost increases as $Q$ decreases.

(iii) In the preceding model, if we always maintain an inventory $B$ on hand as buffer stock, then the average inventory at any time is $\frac{1}{2}Q + B$. Therefore, the total cost per unit time is $C(Q) = \left(\frac{1}{2}Q + B\right)C_1 + C_3\frac{D}{Q}$.

As before, we obtain the optimal values of $Q$ and $T$ as follows:

$$Q = Q^* = \sqrt{\frac{2C_3D}{C_1}} \quad \text{and} \quad T = T^* = \sqrt{\frac{2C_3}{DC_1}}.$$

(iv) In the preceding model, if the ordering cost is taken as $C_3 + bQ$ (where $b$ is the purchase cost per unit quantity) instead of a fixed ordering cost, then there is no change in the optimum order quantity.

*Proof* In this case, the average cost is given by

$$C(Q) = \frac{1}{2}C_1Q + \frac{D}{Q}(C_3 + bQ). \tag{17.4}$$

As the necessary condition for the optimum of $C(Q)$ in (17.4), we have

$$C'(Q) = 0 \text{ implies } Q = \sqrt{\frac{2C_3D}{C_1}} \quad \text{and} \quad C''(Q) > 0.$$

Hence, $Q^* = \sqrt{\frac{2C_3D}{C_1}}$.

This shows that there is no change in $Q^*$ in spite of the change in the ordering cost.

*Example 1* An engineering factory consumes 5000 units of a component per year. The ordering, receiving and handling costs are Rs. 300 per order, the trucking cost is Rs. 1200 per order, the interest cost is Rs. 0.06 per unit per year, the deterioration and obsolescence costs are Rs. 0.004 per unit per year and the storage cost is Rs. 1000 per year for 5000 units. Calculate the economic order quantity and minimum average cost.

**Solution**

In the given problem, we have demand ($D$) = 5000 units.

$$\begin{aligned}
\text{Ordering cost/replenishment cost} &= \text{ordering, receiving, handling costs and trucking costs} \\
&= \text{Rs. } (300 + 1200) = \text{Rs.}1500 \text{ per order.} \\
\text{Inventory carrying cost} &= \text{interest costs} + \text{deterioration and obsolescence costs} \\
&\quad + \text{storage costs} \\
&= \left(0.06 + 0.004 + \frac{1000}{5000}\right) \text{rupees per unit per year} \\
&= \text{Rs. } 0.264 \text{ per unit per year.}
\end{aligned}$$

Hence, the economic order quantity is given by

$$Q^* = \sqrt{\frac{2C_3D}{C_1}} = \sqrt{\frac{2 \times 1500 \times 5000}{0.264}} = 753.8 \text{ (approx.)}.$$

Also, the minimum average cost is

$$\sqrt{2C_1C_3D} = \text{Rs. } \sqrt{2 \times 0.264 \times 1500 \times 5000} = \text{Rs. } 1989.97 \text{ (approx.)}.$$

## 17.6 Manufacturing Model with No Shortage (Model 2)

**(Economic lot-size model with finite rate of replenishment and without shortage)**

In this model, we shall derive the formula for the optimum production quantity per cycle of a single product so as to minimize the total average cost under the following assumptions and notation:

(i) Demand is deterministic and uniform at a rate $D$ units of quantity per unit time.
(ii) Shortages are not allowed.
(iii) The lead time is zero.
(iv) The production rate or replenishment rate is finite, say, $K$ units per unit time ($K > D$).
(v) The production-inventory planning horizon is infinite, and the production system involves only one item and one stocking point.
(vi) The inventory carrying cost $C_1$ per unit quantity per unit time and the setup cost $C_3$ per production cycle are known and constant.
(vii) $T$ is the cycle length and $Q$ is the economic lot size.

In this model, each production cycle time $T$ consists of two parts $t_1$ and $t_2$, where:

(i) $t_1$ is the period during which the stock is increasing at a constant rate $K - D$ units per unit time.
(ii) $t_2$ is the period during which there is no replenishment (or production), but inventory is decreasing at the rate of $D$ units per unit time.

Further, it is assumed that $S$ is the stock available at the end of time $t_1$ which is expected to be consumed during the remaining period $t_2$ at the consumption rate $D$. The pictorial representation of the model is shown in Fig. 17.2.

Therefore, $(K - D)t_1 = S$

$$\text{or} \quad t_1 = \frac{S}{K - D}. \tag{17.5}$$

Since the total quantity produced during the production period $t_1$ is $Q$,

$$\therefore Q = Kt_1$$

$$\text{or } Q = K\frac{S}{K - D}, \quad \text{which implies } S = \frac{K - D}{K}Q. \tag{17.6}$$

Again, $Q = DT$, i.e. $T = \frac{Q}{D}$.

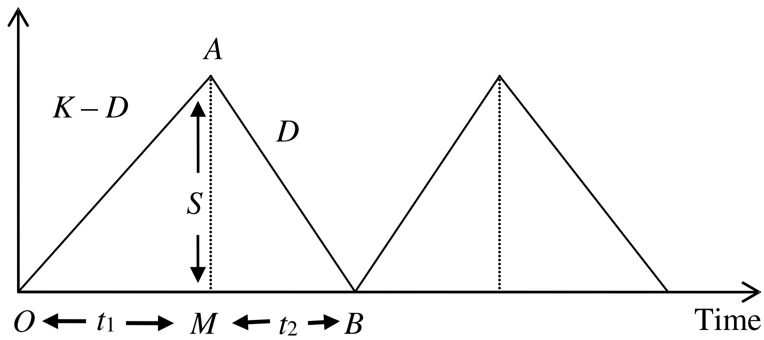



**Fig. 17.2** Pictorial representation of manufacturing inventory model with no shortage

Now the inventory carrying cost for the entire cycle $T$ is $(\Delta OAB)C_1 = \frac{1}{2}TSC_1$, and the setup cost for time period $T$ is $C_3$.

Therefore, the total cost for the entire cycle $T$ is given by $X = C_3 + \frac{1}{2}C_1ST$.

Therefore, the total average cost is given by $C(Q) = \frac{X}{T}$

$$\begin{aligned} \text{or } C(Q) &= \frac{C_3}{T} + \frac{1}{2}C_1S \\ \text{or } C(Q) &= \frac{C_3D}{Q} + \frac{1}{2}C_1\frac{K-D}{K}Q \quad \left[\because\ Q = DT \text{ and } S = \frac{K-D}{K}Q\right]. \end{aligned} \tag{17.7}$$

The optimum of $Q$ which minimizes $C(Q)$ is obtained by equating the first derivative of $C(Q)$ with respect to $Q$ to zero,

$$\begin{aligned} &\text{i.e. } \frac{\mathrm{d}C}{\mathrm{d}Q} = 0 \\ &\text{or } -\frac{C_3D}{Q^2} + \frac{1}{2}C_1\frac{K-D}{K} = 0 \\ &\text{or } Q = \sqrt{\frac{2C_3}{C_1}\cdot\frac{DK}{K-D}}. \end{aligned} \tag{17.8}$$

Again, $\frac{\mathrm{d}^2C}{\mathrm{d}Q^2} = \frac{2C_3D}{Q^3} =$ a positive quantity for $Q = \sqrt{\frac{2C_3}{C_1}\cdot\frac{DK}{K-D}}$.

Hence, $C(Q)$ is minimum for which the optimum value of $Q$ is

$$Q^* = \sqrt{\frac{2C_3}{C_1}\cdot\frac{DK}{K-D}}. \tag{17.9}$$

The corresponding time interval is

$$T^* = \frac{Q^*}{D} = \sqrt{\frac{2C_3K}{C_1D(K-D)}}, \tag{17.10}$$

and the minimum average cost is given by

$$C_{\min} = \frac{1}{2}\frac{K-D}{K}C_1Q^* + \frac{C_3D}{Q^*} = \sqrt{2C_1C_3D\frac{K-D}{K}}. \tag{17.11}$$

*Remark*

(i) For this model, $Q^*$, $T^*$ and $C_{\min}$ can be written in the following form:

$$Q^* = \sqrt{\frac{2C_3D}{C_1}\frac{1}{1-D/K}}, \quad T^* = \sqrt{\frac{2C_3}{DC_1}\frac{1}{1-D/K}}. \tag{17.12}$$

If $K \to \infty$, i.e. the production rate is infinite, this model reduces to Model 1. Therefore, when $K \to \infty$, then $Q^*$, $T^*$ and $C_{\min}$ reduce to the expressions for $Q^*$, $T^*$ and $C_{\min}$ of Model 1.

## 17.7 Purchasing Inventory Model with Shortage (Model 3)

In this model, we shall derive the optimal order level and the minimum average cost under the following assumptions and notation:

(i) Demand is deterministic and uniform at a rate $D$ units of quantity per unit time.
(ii) Production is instantaneous (i.e. production rate is infinite).
(iii) Shortages are allowed and fully backlogged.
(iv) The lead time is zero.
(v) The inventory planning horizon is infinite, and the inventory system involves only one item and one stocking point.
(vi) Only a single order will be placed at the beginning of each cycle, and the entire lot is delivered in one batch.
(vii) The inventory carrying cost $C_1$ per unit quantity per unit time, the shortage cost $C_2$ per unit quantity per unit time and the ordering cost $C_3$ per order are known and constant.
(viii) $Q$ is the lot size per cycle, $S_1$ is the initial inventory level after fulfilling the backlogged quantity of the previous cycle and $Q - S_1$ is the maximum shortage level.
(ix) $T$ be the cycle length or scheduling period whereas $t_1$ is the no shortage period.

The inventory situation is shown in Fig. 17.3. According to the assumptions of (viii) and (ix), we have $Q = DT$.
Regarding the cycle length or scheduling period of the inventory system, two cases may arise:

**Case 1**: The cycle length or scheduling period $T$ is constant.
**Case 2**: The cycle length or scheduling period $T$ is a variable.

For Case 1, $T$ is constant; i.e. the inventory is to be replenished after every time period $T$. As $t_1$ is the no shortage period, $S_1 = Dt_1$ or $t_1 = S_1/D$.

Now, the inventory carrying cost during the period 0 to $t_1$ is

$$C_1 \text{ (area of } \Delta\text{OAB}) = \frac{1}{2} C_1 S_1 t_1 = \frac{1}{2} C_1 S_1^2 / D.$$

Again the shortage cost during the interval ($t_1$, $T$) is

$$\begin{aligned} C_2(\text{area of } \Delta\text{ACD}) &= \frac{1}{2} C_2 (Q - S_1)(T - t_1) \\ &= \frac{1}{2} C_2 (Q - S_1)^2 / D \left[ \because\ T - t_1 = \frac{Q - S_1}{D} \right]. \end{aligned}$$

Hence the total average cost of the system is given by (see Fig. 17.3)

$$C = \left[ \frac{1}{2} C_1 \frac{S_1^2}{D} + \frac{1}{2} C_2 \frac{(Q - S_1)^2}{D} \right] / T. \tag{17.13}$$

Since the setup cost $C_3$ and time period $T$ are constant, the average setup cost $C_3/T$, also being constant, will not be considered in the cost expression.

Since $T$ is constant, $Q = DT$ is also constant. Hence, the preceding expression for the average cost is a function of a single variable $S_1$. Thus, we can easily minimize expression (13) with respect to $S_1$, as in Model 1.

$$\text{In this case, } S_1^* = \frac{C_2 Q}{C_1 + C_2} = \frac{C_2 DT}{C_1 + C_2} \text{ and } C_{\min} = \frac{C_1 C_2 Q}{C_1 + C_2} = \frac{C_1 C_2 DT}{C_1 + C_2}. \tag{17.14}$$

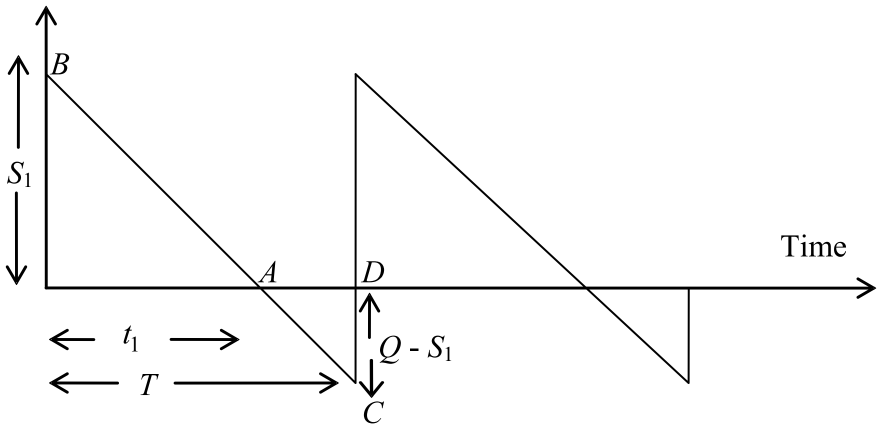



**Fig. 17.3** Pictorial representation of purchasing inventory model with shortage

For Case 2, the cycle length or scheduling period $T$ is a variable. As in Case 1, the average cost of the inventory system will be

$$C = \left[C_3 + \frac{1}{2}C_1\frac{S_1^2}{D} + \frac{1}{2}C_2\frac{(Q - S_1)^2}{D}\right]/T, \tag{17.15}$$

where $Q = DT$.

Here, the average cost $C$ is a function of two independent variables $T$ and $S_1$.

Now, for the optimal value of $C$, we have

$$\frac{\partial C}{\partial S_1} = 0 \quad \text{and} \quad \frac{\partial C}{\partial T} = 0.$$

Thus,

$$\frac{\partial C}{\partial S_1} = 0 \ \text{ gives } \ S_1 = C_2\frac{DT}{(C_1 + C_2)}. \tag{17.16}$$

Again,

$$\frac{\partial C}{\partial T} = 0 \ \text{ gives } \ -\frac{C_1}{2D}\frac{S_1^2}{T^2} + C_2\frac{DT - S_1}{T} - \frac{C_2}{2D}\frac{(DT - S_1)^2}{T^2} - \frac{C_3}{T^2} = 0. \tag{17.17}$$

Setting $S_1 = C_2\frac{DT}{(C_1 + C_2)}$ in this expression and simplifying, we have

$$T = T^* = \sqrt{\frac{2C_3(C_1 + C_2)}{C_1C_2D}}. \tag{17.18a}$$

Then,

$$S_1 = S_1^* = \sqrt{\frac{2C_2C_3D}{C_1(C_1 + C_2)}}. \tag{17.18b}$$

Obviously, for the values of $T$ and $S_1$ given by (17.18a) and (17.18b),

$$\frac{\partial^2 C}{\partial S_1^2} > 0, \frac{\partial^2 S}{\partial T^2} > 0 \quad \text{and} \quad \frac{\partial^2 C}{\partial S_1}\frac{\partial^2 C}{\partial T} - \left(\frac{\partial^2 C}{\partial S_1 \delta T}\right)^2 > 0.$$

Hence, $C$ is minimum for the values of $T$ and $S_1$ given by (17.18a) and (17.18b).

Therefore, the optimum order quantity for minimum cost is given by

$$Q^* = DT^* = D\sqrt{\frac{2C_3(C_1+C_2)}{C_1C_2D}} = \sqrt{\frac{2C_3(C_1+C_2)D}{C_1C_2}} \tag{17.19}$$

and

$$C_{\min} = C^* = \sqrt{\frac{2C_1C_2C_3D}{(C_1+C_2)}}. \tag{17.20}$$

*Remark*

(i) If $C_1 \to \infty$ and $C_2 > 0$, inventories are prohibited. In this case $S_1^* = 0$ and each lot size $Q^* = \sqrt{\frac{2C_3D}{C_2}}$ is used to fill the backorders.

(ii) If $C_2 \to \infty$ and $C_1 > 0$, then shortages are prohibited. In this case, $S_1^* = Q^* = \sqrt{\frac{2C_3D}{C_1}}$ and each batch $Q^*$ is used entirely for inventory.

(iii) If shortage costs are negligible, then $C_1 > 0$ and $C_2 \to 0$.
In this case, $S_1^* \to 0$ and $Q^* \to \infty$.

(iv) If the inventory carrying costs are negligible, then $C_1 \to 0$ and $C_2 > 0$. In this case, $Q^* \to \infty$ and $S_1^* \to \infty$; i.e. $S_1^* \to Q^*$. Thus, due to very small inventory carrying costs, a large lot size should be ordered and used to meet the future demand.

(v) When the inventory carrying costs and shortage costs are equal, i.e. when $C_1 = C_2$, $\frac{C_1}{C_1+C_2} = \frac{1}{2}$.

In this case, $Q^* = \sqrt{2}\sqrt{\frac{2C_3D}{C_1}}$, which shows that the lot size is $\sqrt{2}$ times the lot size of Model 1.

*Example 2* The demand for an item is 18,000 units per year. The inventory carrying cost is Rs. 1.20 per unit time, and the cost of shortages is Rs. 5.00. The ordering cost is Rs. 400.00. Assuming that the replenishment rate is instantaneous, determine the optimum order quantity, shortage quantity and cycle length.

**Solution**

For the problem, it is given that demand ($D$) = 1800 units per year, the carrying cost ($C_1$) = Rs. 1.20 per unit, the shortage cost ($C_2$) = Rs. 5.00 and the ordering cost ($C_3$) = Rs. 400 per order.

The optimum order quantity $Q^*$ is given by

$$Q^* = \sqrt{\frac{2C_3(C_1+C_2)D}{C_1C_2}} = \sqrt{\frac{2\times 400\times(1.2+5)\times 18000}{1.2\times 5}} = 3857 \text{ units.}$$

$$\text{The optimum shortage quantity } Q^* - S_1^* = 3857 - \sqrt{\frac{2C_2C_3D}{C_1(C_1+C_2)}}$$

$$= 3857 - \sqrt{\frac{2 \times 5 \times 400 \times 18,000}{1.2 \times (1.2+5)}} = 746 \text{ units (approx.)}$$

The optimal cycle length $T^* = \frac{Q^*}{D} = \frac{3857}{18000} = 0.214$ year (approx.)

*Example 3* The demand for an item is deterministic and constant over time and is equal to 600 units per year. The unit cost of the item is Rs. 50.00, while the cost of placing an order is Rs. 100.00. The inventory carrying cost is 20% of the unit cost of the item, and the shortage cost per month is Re. 1. Find the optimal ordering quantity. If shortages are not allowed, what would be the loss of the company?

**Solution**
It is given that $D$ = 600 units/year.

$C_1$ = 20% of Rs. 50.00 = Rs. 10.00
$C_2$ = Re 1.00 per month, i.e. Rs. 12.00 per year
$C_3$ = Rs. 100.00 per order.

When shortages are allowed, the optimal ordering quantity $Q^*$ is given by

$$Q^* = \sqrt{\frac{2C_3(C_1+C_2)D}{C_1C_2}} = 148\,\text{units}$$

and the minimum cost per year is $C(Q^*) = \sqrt{2C_1C_2C_3D/(C_1+C_2)}$ = Rs. 809.04.

If shortages are not allowed, then the optimal order quantity is

$$Q^* = \sqrt{\frac{2C_3D}{C_1}} = 109.5\,\text{units}$$

and the relevant average cost is given by $C(Q^*)$ = Rs. $\sqrt{2C_1C_3D}$ = Rs. 1095.44.

Therefore, if shortages are not allowed, the loss of the company will be Rs. (1095.44–809.04), i.e. Rs. 286.40.

## 17.8 Manufacturing Model with Shortage

**(Economic lot-size model with finite rate of replenishment and shortages)**

In this model, we shall derive the formula for the optimum production quantity, shortage quantity and cycle length of a single product by minimizing the average cost of the production system under the following assumptions and notation:

(i) The production rate or replenishment rate is finite, say $K$ units per unit time ($K > D$).
(ii) The production-inventory planning horizon is infinite, and the production system involves only one item and one stocking point.
(iii) Demand of the item is deterministic and uniform at a rate $D$ units of quantity per unit time.
(iv) Shortages are allowed.
(v) The lead time is zero.
(vi) The inventory carrying cost $C_1$ per unit quantity per unit time, the shortage cost $C_2$ per unit quantity per unit time and the setup cost $C_3$ per setup are known and constant.
(vii) $T$ is the cycle length of the system; i.e. $T$ is the interval between production cycles.
(viii) $Q$ is the economic lot size.

Let us assume that each production cycle of length $T$ consists of two parts $t_{12}$ and $t_{34}$ which are further subdivided into $t_1$ and $t_2$, $t_3$ and $t_4$, where (i) inventory is building up at a constant rate $K - D$ units per unit time during the interval $[0, t_1]$, (ii) at time $t = t_1$, the production is stopped and the stock level decreases due to meeting customer demand only up to the time $t_{12} = t_1 + t_2$, (iii) shortages are accumulated at a constant rate of $D$ units per unit time during the time $t_3$, i.e. during the interval $[t_{12}, t_{12} + t_3]$. (iv) Shortages are being filled immediately at a constant rate $K - D$ units per unit time during the time $t_4$, i.e., during the interval $[t_{12} + t_3, T]$. (v) The production cycle then repeats itself after the time $T = t_1 + t_2 + t_3 + t_4$.

Again, let the inventory level be $S_1$ at $t = t_1$ and at the end of time $t = t_1 + t_2$, the stock level reaches zero. Now shortages start, and we assume that shortages are built up of quantity $S_2$ at time $t = t_1 + t_2 + t_3$, and then these shortages are filled up to the time $t = t_1 + t_2 + t_3 + t_4$. The pictorial representation of the inventory situation is given in Fig. 17.4.

Now our objectives are to find the optimal value of $Q, S_1, S_2, t_1, t_2, t_3, t_4$ and $T$ with the minimum average total cost.

The inventory carrying cost over the time period $T$ is given by

$$C_h = C_1 \times \Delta OAC = C_1 \cdot \frac{1}{2} OC \cdot AB = \frac{1}{2} C_1 (t_1 + t_2) S_1,$$

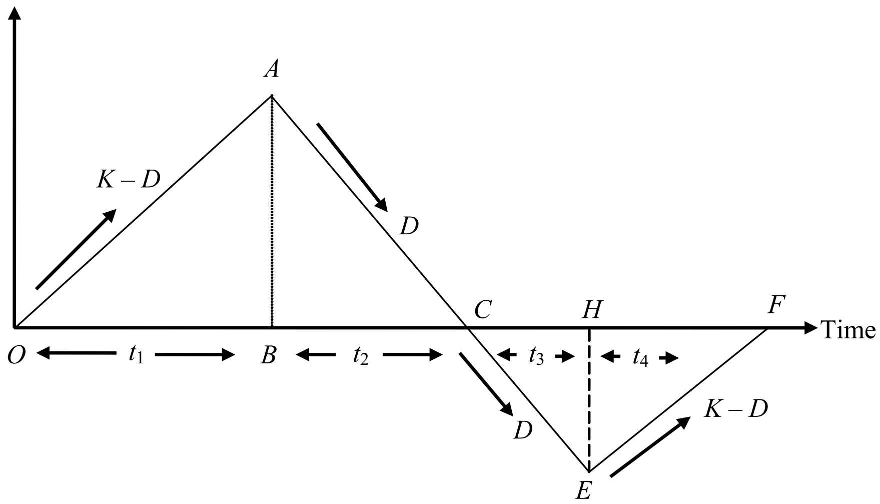



**Fig. 17.4** Pictorial representation of manufacturing inventory model with shortage

and the shortage cost over time $T$ is given by

$$C_s = C_2 \times \Delta CEF = C_2 \cdot \frac{1}{2} CF \cdot EH = \frac{1}{2} C_2 (t_3 + t_4) S_2.$$

Hence, the total average cost of the production system is given by

$$C = [C_3 + C_h + C_s]/T. \tag{17.21}$$

From Fig. 17.4, it is clear that

$$S_1 = (K - D)t_1 \quad \text{or,} \quad t_1 = \frac{S_1}{K - D}. \tag{17.22}$$

Again,

$$S_1 = Dt_2 \quad \text{or,} \quad t_2 = \frac{S_1}{D}. \tag{17.23}$$

Now, in a stock-out situation,

$$S_2 = Dt_3 \quad \text{or,} \quad t_3 = \frac{S_2}{D}$$

and

$$S_2 = (K - D)t_4 \quad \text{or,} \;\; t_4 = \frac{S_2}{K - D}. \tag{17.24}$$

Since the total quantity produced over the time period $T$ is $Q$,

$$Q = DT,$$

where $D$ is the demand rate

$$\begin{aligned} &\text{or} \;\; D(t_1 + t_2 + t_3 + t_4) = Q \\ &\text{or} \;\; D\left(\frac{S_1}{K - D} + \frac{S_1}{D} + \frac{S_2}{D} + \frac{S_2}{K - D}\right) = Q. \end{aligned} \tag{17.25}$$

After simplification, we have

$$S_1 + S_2 = \frac{K - D}{K} Q. \tag{17.26}$$

Again

$$t_1 + t_2 = \frac{K}{D(K - D)} S_1 \quad \text{and} \;\; t_3 + t_4 = \frac{K}{D(K - D)} S_2. \tag{17.27}$$

Now substituting the values of $t_1 + t_2$, $t_3 + t_4$ and $T = Q/D$ in (17.21), we have

$$C(Q, S_1, S_2) = \frac{1}{2Q}\frac{K}{K - D}\left(C_1 S_1 + C_2 S_2^2\right) + \frac{DC_3}{Q}. \tag{17.28}$$

Using (17.26), this equation reduces to

$$C(Q, S_2) = \frac{1}{2Q}\frac{K}{K - D}\left[C_1\left(\frac{K - D}{K}Q - S_2\right)^2 + C_2 S_2^2\right] + \frac{DC_3}{Q}. \tag{17.29}$$

Now, for the extreme values of $C(Q, S_2)$, we have

$$\begin{aligned} &\frac{\partial C}{\partial Q} = 0, \quad \frac{\partial C}{\partial S_2} = 0. \\ &\frac{\partial C}{\partial Q} = 0 \quad \text{implies} \;\; S_2 = C_1 \frac{K - D}{K}\frac{Q}{C_1 + C_2}. \end{aligned} \tag{17.30}$$

Again,

$$\frac{\partial C}{\partial S_2} = 0 \text{ gives } Q = \sqrt{\frac{2C_3(C_1 + C_2)}{C_1 C_2}} \cdot \sqrt{\frac{KD}{K - D}}. \tag{17.31}$$

For the values of $Q$ and $S_2$ given in (17.31) and (17.30), it can easily be verified that

$$\frac{\partial^2 C}{\partial Q^2} > 0, \frac{\partial^2 C}{\partial S_2^2} > 0 \text{ and } \frac{\partial^2 C}{\partial Q^2} \frac{\partial^2 C}{\partial S_2^2} - \left(\frac{\partial^2 C}{\partial Q \partial S_2}\right)^2 > 0.$$

Hence, $C(Q, S_2)$ is minimum and the optimal values of $Q$ and $S_2$ are given by

$$Q^* = \sqrt{\frac{2C_3(C_1 + C_2)}{C_1 C_2}} \sqrt{\frac{KD}{K - D}} \tag{17.32}$$

and

$$S_2^* = \sqrt{\frac{2C_1 C_3}{C_2(C_1 + C_2)}} \sqrt{\frac{D(K - D)}{K}} \tag{17.33}$$

$$T^* = \frac{Q^*}{D} = \sqrt{\frac{2C_3(C_1 + C_2)}{C_1 C_2}} \sqrt{\frac{K}{D(K - D)}} \tag{17.34}$$

$$S_1^* = \frac{K - D}{K} Q^* - S_2^* = \sqrt{\frac{2C_2 C_3}{C_1(C_1 + C_2)}} \sqrt{\frac{D(K - D)}{K}}. \tag{17.35}$$

Now

$$C_{\min} = C(Q^*, S_2^*) = \sqrt{\frac{2C_1 C_2 C_3}{C_1 + C_2}} \sqrt{\frac{D(K - D)}{K}}. \tag{17.36}$$

*Remarks*

(i) In this model, if we assume that the production rate is infinite, i.e. $K \to \infty$, then the optimal quantities obtained by taking $K \to \infty$ in (17.32), (17.34) and (17.36) are

$$Q^* = \sqrt{\frac{2C_3(C_1 + C_2)}{C_1 C_2} D},$$
$$T^* = \sqrt{\frac{2C_3(C_1 + C_2)}{C_1 C_2 D}} \text{ and } C_{\min} = \sqrt{\frac{2C_1 C_2 C_3 D}{C_1 + C_2}}.$$

This means that Model 4 reduces to Model 3 if $K \to \infty$.

(ii) If shortages are not allowed in Model 4, then it reduces to Model 3. In this case, taking $C_2 \to \infty$ in (17.32), (17.34) and (17.36), we obtain the required expressions of Model 3, which are as follows:

$$Q^* = \sqrt{\frac{2C_3KD}{C_1(K-D)}},$$
$$T^* = \sqrt{\frac{2C_3K}{C_1D(K-D)}} \text{ and } C_{\min} = \sqrt{\frac{2C_1C_3D(K-D)}{K}}.$$

*Example 4* The demand for an item in a company is 18,000 units per year. The company can produce the item at a rate of 3000 per month. The cost of one setup is Rs. 500, and the holding cost of one unit per month is Rs. 0.15. The shortage cost of one unit is Rs. 20 per month. Determine the optimum manufacturing quantity and the shortage quantity. Also determine the manufacturing time and the time between setups.

**Solution**

For this problem, it is given that

$C_1$ = Rs. 0.15 per month, $C_2$ = Rs. 20 per month, $C_3$ = Rs. 500.00 per setup, $K$ = 3000 per month and $D$ = 18,000 units per year, i.e. 1500 units per month.

The optimum manufacturing quantity $Q^*$ is given by

$$Q^* = \sqrt{\frac{2C_3(C_1+C_2)}{C_1C_2}}\sqrt{\frac{KD}{K-D}} = \sqrt{\frac{2\times 500\times(0.15+20)}{0.15\times 20}}\sqrt{\frac{3000\times 1500}{3000-1500}}$$
$$= 4489 \text{ units (approx.)}$$

The optimum shortage quantity is given by

$$S_2^* = C_1\frac{K-D}{K}\frac{Q^*}{(C_1+C_2)} = 17 \text{ units (approx.)}.$$

The manufacturing time = $\frac{Q^*}{K} = \frac{4489}{3000} = 1.5$ months, and the time between setups $\frac{Q^*}{D} = \frac{4489}{1500} = 3$ months.

## 17.9 Multi-item Inventory Model

In earlier, we have discussed inventory models for single items, or each item separately, but if there exists a relationship among the items under some limitations, then it is not possible to consider them separately. Thus, after constructing the average cost expression in these models, we shall use the method of Lagrange multipliers to minimize the average cost.

In all such problems, first of all we shall solve the problem ignoring the limitations and then consider the effect of limitations.

Now, we shall develop a multi-item inventory model under the following assumptions and notations:

(i) There are $n$ items with instantaneous production; i.e. the production rate of each item is infinite.
(ii) Shortages are not allowed.

For the $i$th ($i = 1, 2, \ldots, n$) item:

(iii) $D_i$ is the uniform demand rate.
(iv) The inventory carrying cost $C_{1i}$ per unit quantity per unit time and the ordering cost $C_{3i}$ per order are known and constant.
(v) $T_i$ is the cycle length.

Let $Q_i$ be the ordering quantity of the $i$th item.

$$\text{Then,}\quad Q_i = D_i T_i \ \text{ or, }\ T_i = Q_i/D_i.$$

Now, the total inventory time units for the $i$th item is $\frac{1}{2} Q_i T_i$.

Hence, the inventory carrying cost for the $i$th item over the inventory cycle is $\frac{1}{2} C_{1i} Q_i T_i$.

Therefore, the average cost for the $i$th item is

$$C_i = \left[C_{3i} + \frac{1}{2} C_{1i} Q_i T_i\right] / T_i$$

$$\text{or }\ C_i = C_{3i} D_i / Q_i + \frac{1}{2} C_{1i} Q_i.$$

Hence, the total average cost for $n$ items is given by

$$C = \sum_{i=1}^{n} C_i$$

$$\text{or }\ C = \sum_{i=1}^{n} \left[C_{3i} D_i / Q_i + \frac{1}{2} C_{1i} Q_i\right].$$

Here $C$ is a function of $Q_1, Q_2, \ldots, Q_n$.

For optimum values of $Q_i$ ($i$ = 1, 2, …, $n$), we must have

$$\partial C/\partial Q_i = 0,$$
$$\text{i.e.}\ \frac{1}{2}C_{1i} - C_{3i}D_i/Q_i^2 = 0$$
$$\text{or}\ Q_i = \sqrt{\frac{2C_{3i}D_i}{C_{1i}}}.$$

Hence, the optimum value of $Q_i$ is

$$Q_i^* = \sqrt{\frac{2C_{3i}D_i}{C_{1i}}} \tag{17.37}$$

### 17.9.1 Limitation on Inventories

If there is a limitation on inventories that requires that the average number of all units in inventory should not exceed $k$ units of all types, then the problem is to minimize the cost $C$ subject to the condition that:

$\frac{1}{2}\sum_{i=1}^{n} Q_i \leq k$ [Since the average number of inventory at any time for an item is $\frac{1}{2}Q_i$.]

$$\text{or}\ \sum_{i=1}^{n} Q_i - 2k \leq 0. \tag{17.38}$$

Here two cases may arise:

**Case I** When $\frac{1}{2}\sum\limits_{i=1}^{n} Q_i^* \leq k$

In this case, the optimal values $Q_i^*$ ($i$ = 1, 2,…,$n$) given by

$$Q_i^* = \sqrt{\frac{2C_{3i}D_i}{C_{1i}}}$$

satisfy the constraint directly.

**Case II** When $\frac{1}{2}\sum\limits_{i=1}^{n} Q_i^* > k$

In this case, we have to solve the following problem:

$$\text{Minimize} \quad C = \sum_{i=1}^{n} \left[ \frac{1}{2} C_{1i} Q_i + C_{3i} D_i / Q_i \right]$$

subject to the constraint (17.38).

To solve this problem, we shall use the Lagrange multiplier method. The corresponding Lagrangian function is

$$L = \sum_{i=1}^{n} \left[ \frac{1}{2} C_{1i} Q_i + C_{3i} D_i / Q_i \right] + \lambda \left[ \sum_{i=1}^{n} Q_i - 2k \right],$$

where $\lambda\,( > 0)$ is the Lagrange multiplier.

The necessary conditions for $L$ to be minimum are

$$\frac{\partial L}{\partial Q_i} = 0, \; i = 1, 2, \ldots, n \; \text{ and } \; \frac{\partial L}{\partial \lambda} = 0.$$

Now, from $\frac{\partial L}{\partial Q_i} = 0$ we have

$$\frac{1}{2} C_{1i} - \frac{C_{3i} D_i}{Q_i^2} + \lambda = 0, \quad i = 1, 2, \ldots, n$$

$$\text{or} \quad Q_i = \left[ \frac{2 C_{3i} D_i}{C_{1i} + 2\lambda} \right]^{\frac{1}{2}}.$$

Again, from $\frac{\partial L}{\partial \lambda} = 0$ we have

$$\sum_{i=1}^{n} Q_i - 2k = 0 \quad \text{or} \quad \sum_{i=1}^{n} Q_i = 2k.$$

Hence, the optimum value of $Q_i$ is

$$Q_i^* = \left[ \frac{2 C_{3i} D_i}{C_{1i} + 2\lambda^*} \right]^{\frac{1}{2}} \tag{17.39}$$

$$\text{and} \quad \sum_{i=1}^{n} Q_i = 2k. \tag{17.40}$$

To obtain the values of $Q_i^*$ from (17.39), we find the optimal value of $\lambda^*$ of $\lambda$ by the successive trial and error method or the linear interpolation method, subject to the condition given by (17.40). Equation (17.40) implies that $Q_i^*$ must satisfy the inventory constraint in an equality sense.

### 17.9.2 Limitation of Floor Space (or Warehouse Capacity)

Here we shall discuss the multi-item inventory model with the limitation of warehouse floor space. Let $A$ be the maximum storage area available for $n$ different items, $a_i$ be the storage area required per unit of the $i$th item and $Q_i$ be the amount ordered for the $i$th item.

Thus, the storage requirement constraint becomes

$$\sum_{i=1}^{n} a_i Q_i \le A, \quad Q_i > 0$$

$$\text{or} \quad \sum_{i=1}^{n} a_i Q_i - A \le 0. \tag{17.41}$$

Now two possibilities may arise:

**Case 1** When $\sum_{i=1}^{n} a_i Q_i^* \le A$

In this case, the optimal values $Q_i^*$ ($i$ = 1, 2, …, $n$) given by $Q_i^* = \sqrt{\frac{2C_{3i}D_i}{C_{1i}}}$ satisfy the constraint (17.41) directly. Hence, these optimal values $Q_i$ are the required values.

**Case 2** When $\sum_{i=1}^{n} a_i Q_i^* > A$

In this case, we have to solve the problem as follows:

Minimize $C = \sum_{i=1}^{n} \left[\frac{1}{2} C_{1i} Q_i + C_{3i} D_i / Q_i\right]$

subject to the constraint (17.41).

To solve it, we shall use the Lagrange multiplier method, and the corresponding Lagrangian function is

$$L = \sum_{i=1}^{n} \left[\frac{1}{2} C_{1i} Q_i + C_{3i} D_i / Q_i\right] + \lambda \left[\sum_{i=1}^{n} a_i Q_i - A\right],$$

where $\lambda$ (>0) is the Lagrange multiplier.

The necessary conditions for $L$ to be minimum are

$$\frac{\partial L}{\partial Q_i} = 0, \quad i = 1, 2, \ldots, n$$

$$\text{and} \quad \frac{\partial L}{\partial \lambda} = 0.$$

Now from $\frac{\partial L}{\partial Q_i} = 0$, we have

$$\frac{1}{2} C_{1i} - \frac{C_{3i} D_i}{Q_i^2} + \lambda\, a_i = 0, \quad i = 1, 2, \ldots, n. \tag{17.42a}$$

Again, from $\frac{\partial L}{\partial Q_i} = 0$, we have

$$\sum_{i=1}^{n} a_i Q_i - A = 0 \tag{17.42b}$$

Solving (17.42a) and (17.42b), we have the optimal values of $Q_i$ as

$$Q_i^* = \left[ \frac{2C_{3i} D_i}{C_{1i} + 2\,\lambda^* a_i} \right]^{\frac{1}{2}}, \quad i = 1, 2, \ldots, n \tag{17.43}$$

and

$$\sum_{i=1}^{n} a_i Q_i^* = A. \tag{17.44}$$

To obtain the values of $Q_i^*$ from (17.43), we find the optimal value $\lambda^*$ from $\lambda$ by the successive trial and error method or the linear interpolation method subject to the condition given by (17.44). Equation (17.44) implies that $Q_i^*$ must satisfy the inventory constraint in the equality sense.

### 17.9.3 Limitation on Investment

In this case, there is an upper limit $M$ on the amount to be invested on inventory. Let $C_{4i}$ be the unit price of the $i$th item. Then

$$\sum_{i=1}^{n} C_{4i} Q_i \le M. \tag{17.45}$$

Now two possibilities may arise:

***Case I*** When $\sum_{i=1}^{n} C_{4i} Q_i^* \le M$

In this case, the constraint is satisfied by $Q_i^*$ automatically. Hence, the optimal values of $Q_i^*$ ($i$ = 1, 2, …, $n$) are given by

$$Q_i^* = \sqrt{\frac{2C_{3i} D_i}{C_{1i}}}.$$

***Case II*** When $\sum_{i=1}^{n} C_{4i}Q_i^* > M$

In this case, our problem is as follows:

$$\text{Minimize} \quad C = \sum_{i=1}^{n} \left[\frac{1}{2}C_{1i}Q_i + C_{3i}D_i/Q_i\right]$$

subject to the constraint (17.45).

To solve it, we shall use the Lagrange multiplier method, and the corresponding Lagrangian function is

$$L = \sum_{i=1}^{n} \left[\frac{1}{2}C_{1i}Q_i + C_{3i}D_i/Q_i\right] + \lambda\left[\sum_{i=1}^{n} C_{4i}Q_i - M\right],$$

where $\lambda$ (>0) is the Lagrange multiplier.

The necessary conditions for $L$ to be minimum are

$$\frac{\partial L}{\partial Q_i} = 0, \; i = 1, 2, \ldots, n$$
$$\text{and} \quad \frac{\partial L}{\partial \lambda} = 0.$$

Now from $\frac{\partial L}{\partial Q_i} = 0$, we have

$$Q_i = \left[\frac{2C_{3i}D_i}{C_{1i} + 2\lambda C_{4i}}\right]^{\frac{1}{2}}, \quad i = 1, 2, \ldots, n.$$

Again, from $\frac{\partial L}{\partial \lambda} = 0$, we have

$$\sum_{i=1}^{n} C_{4i}Q_i - M = 0 \quad \text{or} \quad \sum_{i=1}^{n} C_{4i}Q_i = M.$$

Hence, the optimum value of $Q_i$ is given by

$$Q_i^* = \left[\frac{2C_{3i}D_i}{C_{1i} + 2\lambda^* C_{4i}}\right]^{\frac{1}{2}} \tag{17.46}$$

$$\text{and} \quad \sum_{i=1}^{n} C_{4i}Q_i^* = M. \tag{17.47}$$

Thus, the values of $Q_i^*$ are obtained from (17.46) subject to the condition given by (17.47) where the optimal value $\lambda^*$ of $\lambda$ can be found by either the successive trial and error method or the linear interpolation method.

*Example 5* A workshop produces three machine parts $A$, $B$, $C$, and the total storage space available is 640 m$^2$. Obtain the optimal lot size for each item from the following data:

| | Items | | |
|---|---|---|---|
| | $A$ | $B$ | $C$ |
| Cost per unit (Rs.) | 10 | 15 | 5 |
| Storage space required (m$^2$/unit) | 0.60 | 0.80 | 0.45 |
| Ordering cost (Rs.) ($C_3$) | 100 | 200 | 75 |
| No. of units required/year | 5000 | 2000 | 10,000 |

The carrying charge on each item is 20% of unit cost.

**Solution**
Considering one year as one unit of time, we have:

carrying charge of $A$ ($C_{11}$) = Rs. (20% of 10) = Rs. 2.00
carrying charge of $B$ ($C_{12}$) = Rs. (20% of 15) = Rs. 3.00
carrying charge of $C$ ($C_{13}$) = Rs. (20% of 5) = Re. 1.00.

Now, without considering the effect of restriction on storage space availability, the optimal value $Q_i^*$ of the $i$th item is given by $Q_i^* = \sqrt{\frac{2C_{3i}D_i}{C_{1i}}},\ \ i = 1, 2, 3.$

$$\therefore\ Q_1^* = \sqrt{\frac{2D_iC_{3i}}{C_{11}}} = \sqrt{\frac{2 \times 5000 \times 100}{2}} = 707$$

$$Q_2^* = \sqrt{\frac{2 \times 2.000 \times 200}{3}} = 516$$

$$Q_3^* = \sqrt{2 \times 10,000 \times 75} = 1225.$$

Then the total storage space required for these values of $Q_i^*$ ($i$ = 1, 2, 3) is

$$\sum_{i=1}^{3} a_iQ_i^* = a_1Q_1^* + a_2Q_2^* + a_3Q_3^*$$

$$= (0.60707 + 0.80516 + 0.451225)\ \text{m}^2$$

$$= 1388.25\ \text{m}^2.$$

This storage space is greater then the available storage space of 640 m$^2$. Therefore, we shall try to find a suitable value of $\lambda^*$ by the trial and error method for computing $Q_i^*$ by using

$$Q_i^* = \left[\frac{2C_{3i}D_i}{C_{1i} + 2\lambda^* a_i}\right]^{\frac{1}{2}}$$

and $\sum_{i=1}^{3} a_i Q_i^* = 640.$

If we take $\lambda^*$ = 5,

$$\begin{aligned} Q_1^* &= \sqrt{\frac{2C_{3i}D_i}{C_{1i} + 2 \times 5 \times a_1}} = \sqrt{\frac{2 \times 5000 \times 100}{2 + 2 \times 5 \times 0.60}} = 354 \\ Q_2^* &= \sqrt{\frac{2 \times 2000 \times 200}{3 + 2 \times 5 \times 0.80}} = 270 \\ \text{and} \quad Q_3^* &= \sqrt{\frac{2 \times 10000 \times 75}{1 + 2 \times 5 \times 0.45}} = 522. \end{aligned}$$

Hence, the corresponding storage space is 0.60354 + 0.80270 + 0.45552 = 663.30 m$^2$. This storage space is greater than the available storage space of 640 m$^2$.

Again, if we take $\lambda^*$ = 6, then

$$\begin{aligned} Q_1^* &= \sqrt{\frac{2 \times 5000 \times 100}{2 + 2 \times 6 \times 0.60}} = 330 \\ Q_2^* &= \sqrt{\frac{2 \times 2000 \times 200}{3 + 2 \times 6 \times 0.80}} = 252 \\ \text{and} \quad Q_3^* &= \sqrt{\frac{2 \times 10000 \times 75}{1 + 2 \times 6 \times 0.45}} = 484. \end{aligned}$$

Hence, the corresponding storage space is 0.60330 + 0.80352 + 0.45484 = 617.40 m$^2$, which is less then the available storage space of 640 m$^2$.

Hence, it is clear that the most suitable value of $\lambda^*$ lies between 5 and 6.

Let us assume that the required storage space will be 640 m$^2$ for $\lambda^*$ = $x$.

Now considering the linear relationship between the value of $\lambda^*$ and the required storage space, we have

$$\begin{aligned} &\tfrac{x-6}{640-617.4} = \tfrac{5-6}{663.3-617.4} \\ \text{or} \quad &x - 6 = \tfrac{-22.6}{45.9} \quad \text{or} \;\; x = 5.5 \text{ (approx.)} \\ &\therefore\; \lambda^* = 5.5. \end{aligned}$$

For this value of $\lambda^*$,

$$Q_1^* = \sqrt{\frac{2 \times 5000 \times 100}{2 + 2 \times 5.5 \times 0.60}} = 341$$

$$Q_2^* = \sqrt{\frac{2 \times 2000 \times 200}{2 + 2 \times 5.5 \times 0.80}} = 272$$

$$\text{And} \quad Q_3^* = \sqrt{\frac{2 \times 10000 \times 75}{1 + 2 \times 5.5 \times 0.45}} = 502.$$

Hence, the optimal lot sizes of the three machine parts $A$, $B$, $C$ are $Q_1^* = 341$ units, $Q_2^* = 272$ units and $Q_3^* = 502$ units.

*Example 6* A company producing three items has a limited inventory of averagely 750 items of all types. Determine the optimal production quantities for each item separately, when the following information is given:

| Product | 1 | 2 | 3 |
|---|---|---|---|
| Holding cost (Rs.) | 0.05 | 0.02 | 0.04 |
| Ordering cost (Rs.) | 50 | 40 | 60 |
| Demand | 100 | 120 | 75 |

**Solution**

Neglecting the restriction of the total value of the inventory level, we get the optimal values $Q_i^*$ for the $i$th item, which is given by $Q_i^* = \sqrt{\frac{2C_{3i}D_i}{C_{1i}}}$, $i = 1, 2, 3.$

$$\text{Hence,} \quad Q_1^* = \sqrt{\frac{2 \times 50 \times 100}{0.05}} = 447$$

$$Q_2^* = \sqrt{\frac{2 \times 40 \times 120}{0.02}} = 693$$

$$\text{And} \quad Q_3^* = \sqrt{\frac{2 \times 60 \times 75}{0.04}} = 474.$$

Therefore, the total average inventory is (447 + 693 + 474)/2 units = 807 units.

But the average inventory is 750 units. Therefore, we have to determine the value of parameter $\lambda^*$ by the trial and error method for computing $Q_i^*$ by using

$$Q_i^* = \sqrt{\frac{2C_{3i}D}{C_{1i} + 2\lambda^*}} \quad \text{and} \quad \frac{1}{2}\sum Q_i^* = 750.$$

Now, for $\lambda^* = 0.005$,

$$Q_1^* = \sqrt{\frac{2 \times 50 \times 100}{0.05 + 2 \times 0.005}} = 408, \quad Q_2^* = 566 \text{ and } Q_3^* = 424.$$

Therefore, the total average inventory is (408 + 566 + 424)/2 = 699 units, which is less than the given average inventory of items.
Again, for $\lambda^* = 0.003$,

$$Q_1^* = 423, \quad Q_2^* = 608, \quad Q_3^* = 442,$$

and the average inventory = (423 + 608 + 442) = 737, which is less than 750 units.

Again, for $\lambda^* = 0.002$, we have $Q_1^* = 430$, $Q_2^* = 632$, $Q_3^* = 452$, and the average inventory = (430 + 632 + 452) = 757, which is greater than 750 units.
Therefore, the most suitable value of $\lambda^*$ lies between 0.002 and 0.003.
Let us assume that for $\lambda^* = x$ the average inventory will be 750.
Now, considering the linear relationship between $\lambda^*$ and the average inventory, we have

$$\frac{x - 0.003}{750 - 737} = \frac{0.003 - 0.002}{737 - 757} \quad \text{or} \quad x - 0.003 = \frac{13 \times 0.001}{-20}$$
$$\text{or} \quad x = 0.00235 \quad \text{or} \quad \lambda^* = 0.00235.$$

For $\lambda^* = 0.00235$, $Q_1^* = 428, Q_2^* = 623, Q_3^* = 449$.

Hence, the optimal production quantities for products 1, 2 and 3 are 428, 623 and 449 units respectively.

## 17.10 Inventory Models with Price Breaks

In the earlier discussion, we assumed that the unit production cost or unit purchase cost is constant, so, we did not need to consider this cost in the analysis. However, in reality, it is not always true that the unit cost of an item is independent of the quantity procured or produced. Again, discounts are offered by the supplier or wholesaler or manufacturer for the purchase of large quantities. Such discounts are referred to as quantity discounts or price breaks.

In this section, we shall consider a class of inventory in which cost is a variable. When items are purchased in bulk, some discount price is usually offered by the supplier.

Let us assume that the unit purchase cost of an item is $p_j$ when the purchased quantity lies between $b_{j-1}$ and $b_j$ ($j = 1, 2, \ldots, m$). Explicitly, we have:

| Range of quantity to be purchased | unit purchase cost |
|---|---|
| $b_0 < Q < b_1$ | $p_1$ |
| $b_1 \le Q < b_2$ | $p_2$ |
| $b_2 \le Q < b_3$ | $p_3$ |
| ... | ... |
| $b_{j-1} \le Q < b_j$ | $p_j$ |
| ... | ... |
| $b_{m-1} \le Q < b_m$ | $p_m$ |

In general, $b_0 = 0$ and $b_m = \infty$ and $p_1 > p_2 > \cdots > p_j > \cdots > p_m$. The values $b_1$, $b_2$, $b_3$, …, $b_{m-1}$ are termed as price breaks, as the unit price lies in the intervals between these values.

The problem is to determine an economic order quantity $Q$ which minimizes the total cost.

In this model, the assumptions are as follows:

(i) The demand rate is known and uniform.
(ii) Shortages are not permitted.
(iii) Production for the supply of commodities is instantaneous.
(iv) The lead time is zero.

### 17.10.1 Purchasing Inventory Model with Single Price Break

Let $D$ be the demand rate, $C_1$ be the holding cost per unit quantity per unit time and $C_3$ the fixed ordering cost per order. Also, let $p_1$ be the purchasing cost per unit quantity if the ordered quantity is less than $b$ and $p_2$ ($p_2 < p_1$) be the purchasing cost per unit quantity if the ordered quantity is greater than or equal to $b$ quantities, i.e.

| Range of quantity to be purchased | Unit purchase cost |
|---|---|
| $0 < Q < b$ | $p_1$ |
| $b \le Q$ | $p_2$ |

Hence, the total average cost $C(Q)$ is given by

$$C(Q) = \langle \text{ordering cost} \rangle + \langle \text{purchasing cost} \rangle + \langle \text{holding cost} \rangle,$$

i.e.

$$C(Q) = \begin{cases} C'(Q) & \text{for} \quad 0 < Q < b \\ C''(Q) & \text{for} \quad Q > b \end{cases},$$

where $C'(Q) = C_3 \frac{D}{Q} + p_1 D + \frac{1}{2} C_1 Q$ and $C''(Q) = C_3 \frac{D}{Q} + p_2 D + \frac{1}{2} C_1 Q$.

Thus, $C(Q)$ has a discontinuity at $Q = b$, and it may be shown that the minimum value of $C(Q)$ occurs either at the point where $\frac{\mathrm{d}C(Q)}{\mathrm{d}Q} = 0$ or at the point of discontinuity.

We have $\frac{\mathrm{d}C(Q)}{\mathrm{d}Q} = -\frac{C_3 D}{Q^2} + \frac{1}{2} C_1$ except at $Q = b$ where it is not defined. Thus, the optimal value of $Q$ is given by

$$\left.\frac{\mathrm{d}C(Q)}{\mathrm{d}q}\right|_{Q=Q^*} = 0,$$

$$\text{which implies } Q^* = \sqrt{\frac{2C_3 D}{C_1}}. \tag{17.48}$$

Now we consider the case in which $Q^* \geq b$ and $Q^* < b$.

(i) If $Q^*$ [given by (17.48)] $> b$, then the optimal lot size $Q^*$ is obtained by (17.48), and in this case, the minimum total average cost is given by

$$\begin{aligned} C_{\min}(Q^*) &= \frac{C_3 D}{\sqrt{\frac{2C_3 D}{C_1}}} + p_2 D + \frac{1}{2} C_1 \sqrt{\frac{2C_3 D}{C_1}} \\ &= p_2 D + \sqrt{2C_1 C_3 D}. \end{aligned}$$

(ii) If $Q^* < b$, then there may arise two cases as follows:

**Case 1**: $C''(b) < C'(Q^*)$ for $Q^* < b$
**Case 2**: $C''(b) > C'(Q^*)$ for $Q^* < b$.

Now, $C'(Q^*) = p_1 D + \sqrt{2C_1 C_3 D}$

and $C''(b) = p_2 D + \frac{C_3 D}{b} + \frac{1}{2} C_1 b$.

Hence, if $C''(b) > C'(Q^*)$, then $Q^*$ given by (17.48) is the optimum order quantity. Otherwise, if $C''(b) < C'(Q^*)$, then $b$ is the optimum order quantity.

**Working rule**:

Step I. Compute $Q^*$ by the formula $Q^* = \sqrt{\frac{2C_3 D}{C_1}}$ for the case $Q > b$ and then compare this $Q^*$ with the value of $b$.

(i) If $Q^* > b$, then the optimum lot size is $Q^*$.
(ii) If $Q^* < b$, then go to Step 2.

Step 2. Evaluate $Q^*$ by the formula $Q^* = \sqrt{\frac{2C_3D}{C_1}}$ for the case $Q < b$ and evaluate $C'(Q^*)$ and $C''(b)$.

(i) If $C'(Q^*) < C''(b)$, then $Q^*$ is the optimum lot size.

(ii) Otherwise, $b$ is the optimal lot size.

### *17.10.2 Purchase Inventory Model with Two Price Breaks*

| Range of quantity to be purchased | Unit purchase cost |
|---|---|
| $0 < Q < b_1$ | $p_1$ |
| $b_1 \leq Q < b_2$ | $p_2$ |
| $b_2 \leq Q$ | $p_3$ |

The procedure used involving one price break is extended to the case with two price breaks.

***Working rule***:

Step 1. Compute $Q^*$ for $Q > b_2$ (say $Q_3^*$). If $Q_3^* > b_2$, then the optimal lot size is $Q_3^*$; otherwise, go to Step 2.

Step 2. Compute $Q^*$ for $b_1 \leq Q^* < b_2$ (say $Q_2^*$). Since $Q_3^* < b_2$, then $Q_2^*$ is also less than $b_2$. In this case there are two possibilities: either $Q_2^* \geq b_1$ or $Q_2^* < b_1$. If $Q_2^* \geq b_1$, then compare the cost $C(Q_2^*)$ and $C(b_2)$ to obtain the optimum lot size. The quantity with lower cost will naturally be the optimum one. If $Q_2^* < b_1$, then go to Step 3.

Step 3. If $Q_2^* < b_1$, then compute $Q_1^*$ for the case $0 < Q < b_1$ and compare the cost $C(Q_1^*)$, $C(b_1)$ and $C(b_2)$ to determine the optimal lot size. The quantity with lower cost will naturally be the optimum one.

*Example 7* Find the optimum order quantity for a product for which the price breaks are as follows:

| Range of quantity to be purchased | Purchase cost per unit |
|---|---|
| $0 < Q < 100$ | Rs. 20.00 |
| $100 \leq Q < 200$ | Rs. 18.00 |
| $200 \leq Q$ | Rs. 16.00 |

The monthly demand for the product is 400 units. The storage cost is 20% of the unit cost of the product and the cost of ordering is Rs. 25.00 per order.

**Solution**

Here $D$ = 400 units/month, $C_3$ = Rs. 25.00 per order

$C_1$ = 20% of purchase cost per unit = 0.2 time of purchase cost per unit

$$\text{Let } Q_3^* = \sqrt{\frac{2C_3D}{C_1}} \text{ for } \; Q > 200$$
$$= \sqrt{\frac{2 \times 25 \times 400}{0.2 \times 16}} = 79.$$

Since $Q_3^* < 200$, $Q_3^*$ is not the optimum order quantity. Therefore, we have to proceed to calculate $Q_2^*$.

$$\text{Now } Q_2^* = \sqrt{\frac{2C_3D}{C_1}} \text{ for } \; 100 \le Q < 200$$
$$= \sqrt{\frac{2 \times 25 \times 400}{0.2 \times 18}} = 75.$$

Again, since $Q_2^* < 100$, $Q_2^*$ is not the optimum order quantity.

Now we have to proceed to calculate $Q_1^*$.

$$Q_1^* = \sqrt{\frac{2C_3D}{C_1}} \text{ for } 0 < Q < 100$$
$$= \sqrt{\frac{2 \times 25 \times 400}{0.2 \times 20}} = 71.$$

Since $0 < Q_1^* < 100$, then we have to compute $C(Q_1^*)$, $C(100)$, $C(200)$.

Now

$$C(Q_1^*) = C'(Q_1^*) = p_1D + \frac{1}{2}C_1Q_1^* + \frac{C_3D}{Q_1^*}$$
$$= 20 \times 400 + \frac{1}{2} \times 0.2 \times 20 \times 71 + \frac{25 \times 400}{71}$$
$$= \text{Rs. } 8282.85$$

$$\mathrm{C}(100) = C''(100) = p_2D + \frac{1}{2}C_1 \times 100 + \frac{C_3D}{100}$$
$$= 18 \times 400 + \frac{1}{2} \times 0.2 \times 18 \times 100 + \frac{25 \times 400}{100}$$
$$= \text{Rs. } 7480.00$$

and

$$\begin{aligned} C(200) = C'''(200) &= p_3 D + \frac{1}{2} C_1 \times 200 + \frac{C_3 D}{200} \\ &= 16 \times 400 + \frac{1}{2} \times 0.2 \times 16 \times 200 + \frac{25 \times 400}{200} \\ &= \text{Rs.}\, 6770.00. \end{aligned}$$

Since $C(200) < C(100) < C(Q_1^*)$, the optimal order quantity is 200 units, i.e. $Q^* = 200$.

## 17.11 Probabilistic Inventory Model

Here we consider the situations when the demand is not known exactly, but the probability distribution of the demand is somehow known. The control variable in such cases is assumed to be either the scheduling period or the order level or both. The optimum order levels will thus be derived by minimizing the total expected cost rather than the actual cost involved.

### *17.11.1 Single Period Inventory Model with Continuous Demand (No Replenishment Cost Model)*

In this model, we have to find the optimum order quantity so as to minimize the total expected cost with the following assumptions:

(i) The scheduling period $T$ is fixed and known. Hence, we do not include the setup cost in our derivation, as it is a prescribed constant.
(ii) Production is instantaneous.
(iii) The lead time is zero.
(iv) The demand is uniformly distributed over the period.
(v) Shortages are allowed and fully backlogged.
(vi) The holding cost $C_1$ per unit quantity per unit time and the shortage cost $C_2$ per unit quantity per unit time are known and constant.

**Model Formulation**:
Let $x$ be the amount on hand before an order is placed.
Also, let $Q$ be the level of inventory at the beginning of each period, and $r$ units is the demand per time period. Depending on the amount $r$, two cases may arise:

**Case *I***: $r \leq Q$
**Case *II***: $r > Q$

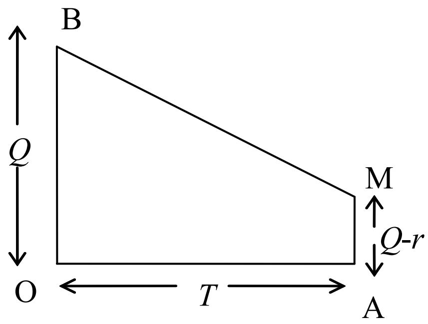



**Fig. 17.5** Pictorial representation of single period model with continuous demand (Case I)

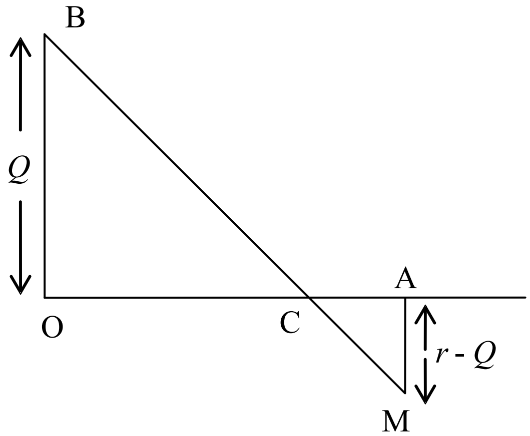



**Fig. 17.6** Pictorial representation of single period model with continuous demand (Case II)

The inventory situation for the case of $r \leq Q$ is shown in Fig. 17.5 and the situation for the case of $r > Q$ is shown in Fig. 17.6.

In the first case, $r \leq Q$ as shown in Fig. 17.5, no shortage occurs. In the second case, $r > Q$ as shown in Fig. 17.6, both the costs are involved.

### Discrete case: when *r* is a discrete random variable

Let the demand for $D = r$ units be estimated at a discontinuous rate with probability $p(r)$, $r = 1, 2, \ldots, n, \ldots$ That is, we may expect the demand for one unit with probability $p(1)$, 2 units with probability $p(2)$ and so on. Since all the possibilities are to be considered, we must have $\sum_{r=1}^{\infty} p(r) = 1$ and $p(r) \geq 0$.

We also assume that $r$ will only be non-negative integers.

***Case I:*** In this case, $r \leq Q$. Thus, there is no shortage, and the total inventory is represented by the total area OAMB $= \frac{1}{2}(Q + Q - r)T = (Q - r/2)T$ (see Fig. 17.5).

Hence, the holding cost for the time period $T$ is $C_1$ $(Q - r/2)$ $T$. This is the holding cost when $r$ $(\leq Q)$ units is the demand rate in one period. But, the probability of the demand of $r$ units is $p(r)$. Hence, the expected value of the cost is $C_1$ $(Q - r/2)$ $T$ $p(r)$.

Now, $r$ can have only values less than $Q$. Hence, the total expected cost when $r < Q$ is equal to

$$\sum_{r=0}^{Q} C_1\left(Q-\frac{r}{2}\right)Tp(r).$$

***Case II:*** In this case, $r > Q$, both the holding and shortage costs are involved.

Here, the holding cost is $\frac{1}{2}C_1Qt_1$ and the shortage cost is $\frac{1}{2}C_2(r-Q)t_2$ where $t_1$ and $t_2$ represent the no-shortage and shortage cases and $t_1 + t_2 = T$.

Now, from the similar triangles OBC and ACM in Fig. 17.6, we have

$$\begin{aligned}
&\frac{t_1}{Q}=\frac{t_2}{r-Q}\\
\text{or}\quad &\frac{t_1}{Q}=\frac{t_2}{r-Q}=\frac{t_1+t_2}{r}\quad \text{or,}\quad \frac{t_1}{Q}=\frac{t_2}{r-Q}=\frac{T}{r}\\
\text{or}\quad &t_1=\frac{QT}{r}\quad\text{and}\quad t_2=\frac{(r-Q)T}{r}.
\end{aligned}$$

Hence the expected cost is

$$\begin{aligned}
&\sum_{r=Q+1}^{\propto}\left[\frac{1}{2}C_1Qt_1p(r)+\frac{1}{2}(r-Q)C_2t_2p(r)\right]\\
&=\sum_{r=Q+1}^{\propto}\left[C_1\frac{Q^2}{2r}Tp(r)+C_2\frac{(r-Q)^2}{2r}Tp(r)\right].
\end{aligned}$$

Therefore, the average expected cost is given by

$$\mathrm{TEC}(Q)=C_1\sum_{r=0}^{Q}\left(Q-\frac{r}{2}\right)p(r)+C_1\sum_{r=Q+1}^{\infty}\frac{Q^2}{2r}p(r)+C_2\sum_{r=Q+1}^{\infty}\frac{(r-Q)^2}{2r}p(r). \tag{17.49}$$

The problem is now to find $Q$, so as to minimize TEC($Q$). Let an amount $Q + 1$ instead of $Q$ be produced. Then the average total expected cost is

$$\mathrm{TEC}(Q+1)=C_1\sum_{r=0}^{Q+1}\left(Q+1-\frac{r}{2}\right)p(r)+C_1\sum_{r=Q+2}^{\propto}\frac{(Q+1)^2}{2r}p(r)+C_2\sum_{r=Q+1}^{\propto}\frac{(r-Q-1)^2}{2r}p(r).$$

But

$$
\begin{aligned}
C_1 \sum_{r=o}^{Q+1} \left(Q+1-\frac{r}{2}\right)p(r) &= C_1 \sum_{r=0}^{Q} \left(Q+1-\frac{r}{2}\right)p(r) \\
&\quad + C_1\left(Q+1-\frac{Q+1}{2}\right)p(Q+1) \\
&= C_1 \sum_{r=0}^{Q} \left(Q-\frac{r}{2}\right)p(r) + C_1 \sum_{r=0}^{Q} p(r) + C_1 \frac{Q+1}{2} p(Q+1).
\end{aligned}
$$

Again,

$$
\begin{aligned}
C_1 \sum_{r=Q+2}^{\infty} \frac{(Q+1)^2}{2r} p(r) &= C_1 \sum_{r=Q+1}^{\infty} \frac{(Q+1)^2}{2r} p(r) - C_1 \frac{(Q+1)^2}{2(Q+1)} p(Q+1) \\
&= C_1 \sum_{r=Q+1}^{\infty} \frac{Q^2+2Q+1}{2r} p(r) - C_1 \frac{Q+1}{2} p(Q+1) \\
&= C_1 \sum_{r=Q+1}^{\infty} \frac{Q^2}{2r} p(r) + C_1 \sum_{r=Q+1}^{\infty} \frac{Q}{r} p(r) + \frac{C_1}{2} \sum_{r=Q+1}^{\infty} \frac{p(r)}{r} \\
&\quad - \frac{C_1}{2}(Q+1)p(Q+1).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&C_2 \sum_{r=Q+2}^{\infty} \frac{(r-Q-1)^2}{2r} p(r) \\
&= C_2 \sum_{r=Q+1}^{\infty} \frac{(r-Q-1)^2}{2r} p(r) - 0 \\
&= C_2 \sum_{r=Q+1}^{\infty} \frac{(r-Q)^2}{2r} p(r) - C_2 \sum_{r=Q+1}^{\infty} p(r) + C_2 \sum_{r=Q+1}^{\infty} \frac{Q}{r} p(r) + \frac{C_2}{2} \sum_{r=Q+1}^{\infty} \frac{p(r)}{r}.
\end{aligned}
$$

Substituting these values in TEC($Q$ + 1) and then simplifying, we get

$$
\mathrm{TEC}(Q+1) = \mathrm{TEC}(Q) + (C_1+C_2)\left[\sum_{r=0}^{Q} p(r) + \left(Q+\frac{1}{2}\right) \sum_{r=Q+1}^{\infty} \frac{p(r)}{r}\right] - C_2.
$$

If we set $\displaystyle \sum_{r=0}^{Q} p(r) + \left(Q+\frac{1}{2}\right) \sum_{r=Q+1}^{\infty} \frac{p(r)}{r} = L(Q),$ (17.50)

$$\text{then } TEC(Q+1) = TEC(Q) + (C_1 + C_2)L(Q) - C_2 \tag{17.51}$$

Similarly, substituting $Q - 1$ in place of $Q$ in (17.49), we have

$$TEC(Q-1) = TEC(Q) - (C_1 + C_2)L(Q-1) + C_2. \tag{17.52}$$

For optimal $Q$, we must have

$$TEC(Q+1) - TEC(Q) > 0$$

i.e. $(C_1 + C_2)L(Q^*) - C_2 > 0$ [From (17.51)]

$$\text{or} \quad L(Q*) > \frac{C_2}{C_1 + C_2}. \tag{17.53}$$

Again, for optimal $Q$, we have

$$\begin{aligned} &\mathrm{TEC}(Q-1) - TEC(Q) > 0 \\ &\text{or} - (C_1 + C_2)L(Q^* - 1) + C_2 > 0 \\ &\text{or} \;\; L(Q^* - 1) < \frac{C_2}{C_1 + C_2}. \end{aligned} \tag{17.54}$$

Combining (17.53) and (17.54) for the optimal value of $Q^*$, we have

$$L(Q^* - 1) < \frac{C_2}{C_1 + C_2} < L(Q^*), \tag{17.55}$$

where $L(Q) = \sum_{r=0}^{Q} p(r) + \left(Q + \frac{1}{2}\right) \sum_{r=Q+1}^{\infty} \frac{p(r)}{r}$.

Using the relation (17.55), we find the range of optimum values of $Q$. In these cases, $Q^*$ need not be unique. If $\frac{C_2}{C_1 + C_2} = L(Q^*)$, then both $Q^*$ and $Q^* + 1$ are the optimal values. Similarly, if $\frac{C_2}{C_1 + C_2} = L(Q^* - 1)$, then both $Q^*$ and $Q^* - 1$ are the optimal values.

**Continuous case**: *when r is a continuous random variable*

When the uncertain demand is estimated as a continuous random variable, the cost expressions of inventory, holding and shortage costs involve integrals instead of summation signs.

Let *f(r)* be the known probability density function for demand *r*. The discrete point probabilities *p(r)* are replaced by the probability differential *f(r)*d*r* for a small interval, say, $\left(r - \frac{\mathrm{d}r}{2}, r + \frac{\mathrm{d}r}{2}\right)$. In this case, we have

$$\int\limits_0^\infty f(r)\mathrm{d}r = 1 \;\text{ and }\; f(r) \geq 0.$$

**Case 1** When $r \leq Q$

Proceeding as before for $r \leq Q$, the holding cost is $C_1\left(Q - \frac{r}{2}\right)t$ and there is no shortage cost.

**Case 2** When $r > Q$

Proceeding as before for $r > Q$, the holding cost is $\frac{C_1 Q^2 t}{2r}$ and the shortage cost is $C_2\frac{(r-Q)^2 t}{2r}$.

$$\mathrm{TEC}(Q) = \int\limits_0^Q C_1\left(Q - \frac{r}{2}\right) f(r)\mathrm{d}r + \int\limits_Q^\infty \left[C_1\frac{Q^2}{2r} + C_2\frac{(r-Q)^2}{2r}\right] f(r)\mathrm{d}r$$

$$\therefore\;\; \frac{\mathrm{dTEC}(Q)}{\mathrm{d}Q} = C_1\int\limits_0^Q f(r)\mathrm{d}r + C_1[\left(Q - \frac{r}{2}\right) f(r)\frac{\mathrm{d}r}{\mathrm{d}Q}\bigg|_0^Q$$

$$+ \int\limits_Q^\infty \left[\frac{C_1}{2r}2Q - \frac{C_2}{2r}2(r-Q)\right] f(r)\mathrm{d}r + \left\{\frac{c_1 Q^2}{2r} + \frac{C_2(r-Q)^2}{2r}\right\} f(r)\frac{\mathrm{d}r}{\mathrm{d}Q}\bigg|_Q^\infty$$

After simplification, we have

$$\frac{\mathrm{dTEC}(Q)}{\mathrm{d}Q} = (C_1 + C_2)\int\limits_0^Q f(r)\mathrm{d}r + (C_1 + C_2)\int\limits_Q^\infty \frac{Qf(r)}{r}\mathrm{d}r - C_2. \tag{17.56}$$

The necessary condition for *TEC*(*Q*) to be optimum is $\frac{\mathrm{dTEC}(Q)}{\mathrm{d}Q} = 0 \quad \text{for} \;\; Q = Q^*$,

i.e. $(C_1 + C_2)\int_0^{Q^*} f(r)\mathrm{d}r + (C_1 + C_2)\int_{Q^*}^\infty \frac{Q^* f(r)}{r} - C_2 = 0$

$$\text{or} \quad \int\limits_0^{Q^*} f(r)\mathrm{d}r + \int\limits_{Q^*}^\infty \frac{Q^* f(r)}{r} = \frac{C_2}{C_1 + C_2}. \tag{17.57}$$

Again from (17.56),

$$\frac{d^2\mathrm{TEC}(Q)}{dQ^2} = (C_1 + C_2)f(Q) + (C_1 + C_2)\int_Q^\infty \frac{f(r)}{r}\,dr - (C_1 + C_2)f(Q).$$

After simplification, we have

$$\frac{d^2 TEC(Q)}{dQ^2} = (C_1 + C_2)\int_Q^\infty \frac{f(r)}{r}\,dr = \text{a positive quantity.}$$

Hence, Eq. (17.57) gives the optimum value of $Q$ for the minimum expected average cost.

*Example 8* A contractor of second-hand motor trucks maintains a stock of trucks every month. The demand of the trucks occurs at a relatively constant rate but not with a constant amount. The demand follows the following probability distributions:

| Demand ($r$) | 0 | 1 | 2 | 3 | 4 | 5 | 6 or more |
|---|---|---|---|---|---|---|---|
| Probability [$p(r)$] | 0.40 | 0.24 | 0.20 | 0.10 | 0.05 | 0.01 | 0.0 |

The holding cost of an old truck in stock for one month is Rs. 100.00, and the penalty for a truck not being supplied on demand is Rs. 1000.00. Determine the optimal size of the stock for the contractor.

**Solution**

| $Q$ | $r$ | $p(r)$ | $\frac{p(r)}{r}$ | $\sum_{r=Q+1}^{\infty}\frac{p(r)}{r}$ | Q + $\frac{1}{2}$ | $(Q+\frac{1}{2})\sum_{r=Q+1}^{Q+\frac{1}{2}}\frac{p(r)}{r}$ | $\sum_{r=0}^{Q}p(r)$ | $L(Q) = (\mathbf{Q}+\frac{1}{2})\sum_{r=Q+1}^{\infty}\frac{p(r)}{r} + \sum_{r=0}^{Q}p(r)$ |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0.40 | $\infty$ | 0.3878 | 0.5 | 0.1939 | 0.40 | 0.5937 |
| 1 | 1 | 0.24 | 0.2400 | 0.1478 | 1.5 | 0.2217 | 0.64 | 0.8617 |
| 2 | 2 | 0.20 | 0.1000 | 0.0478 | 2.5 | 0.1195 | 0.84 | 0.9595 |
| 3 | 3 | 0.10 | 0.0333 | 0.0145 | 3.5 | 0.0575 | 0.94 | 0.9907 |
| 4 | 4 | 0.05 | 0.0125 | 0.0020 | 4.5 | 0.0090 | 0.99 | 0.9990 |
| 5 | 5 | 0.01 | 0.0020 | 0.0000 | 5.5 | 0.0000 | 1.0 | 1.0000 |
| 6 or more | 6 or more | 0.00 | 0.0000 | 0.0000 | 6.5 | 0.0000 | 1.0 | 1.0000 |

Here $C_1$ = Rs. 100.00 and $C_2$ = Rs. 1000.00.

$$\therefore \quad \frac{C_2}{C_1 + C_2} = \frac{1000}{100 + 1000} = \frac{10}{11} = 0.9090.$$

Now for the optimal value of $Q$ (say, $Q^*$), we must have

$$L(Q^* - 1) < \frac{C_2}{C_1 + C_2} < L(Q^*),$$

where $L(Q) = \sum_{r=0}^{Q} p(r) + \left(Q + \frac{1}{2}\right) \sum_{r=Q+1}^{\infty} \frac{p(r)}{r}$.

From the table, it is clear that for $Q = 2$, the preceding inequality is satisfied;

i.e. $L(1) < 0.9090 < L(2)$ or $L(2 - 1) < 0.9090 < L(2)$.

Hence, $Q^* = 2$; i.e. the optimum stock level of trucks is 2.

### *17.11.2 Single Period Inventory Model with Instantaneous Demand (no Replenishment Cost Model)*

In this model, we have to find the optimum order quantity which minimizes the total expected cost under the following assumptions:

(i) $T$ is the constant interval between orders ($T$ may also be considered as unity, e.g. daily, weekly, monthly, etc.).
(ii) $Q$ is the stock level at the beginning of each period $T$.
(iii) The lead time is zero.
(iv) The holding cost $C_1$ per unit quantity per unit time and the shortage cost $C_2$ per unit quantity per unit time are known and constant.
(v) $r$ is the demand at each interval $T$.

**Model Formulation**:

In this model it is assumed that the total demand is filled at the beginning of the period.

Thus, depending on the amount $r$ demanded, the inventory position just after the demand occurs may be either positive (surplus) or negative (shortage); i.e. there are two cases:

**Case 1**: $r \leq Q$
**Case 2**: $r > Q$

The corresponding inventory situations are shown in Figs. 17.7 and 17.8.

**Discrete case:** ***when r is discrete***

Let $r$ be the estimated demand at an instantaneous rate with probabilities $p(r)$. Then there is only a holding cost and no shortage cost.

Here the holding cost is $C_1(Q - r)$.

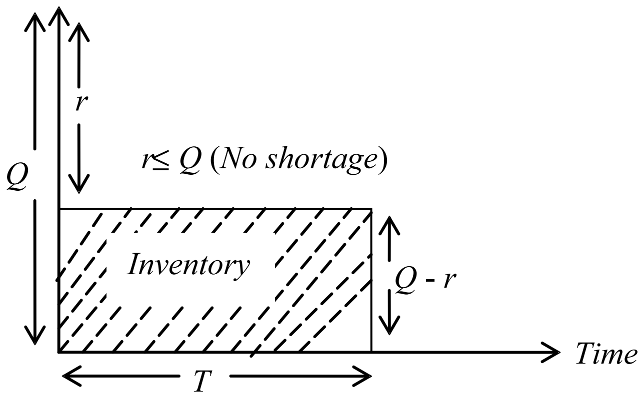



**Fig. 17.7** Pictorial representation of single period model with no shortage

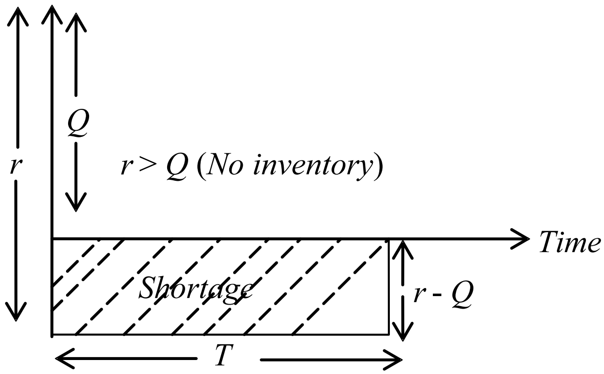



**Fig. 17.8** Pictorial representation of single period model with no inventory

In the second case, the demand $r$ is filled at the beginning of the period. There is only a shortage cost, no holding cost. Therefore, the shortage cost in this case is $C_2$ $(r - Q)$.

Therefore, the total expected cost for this model is

$$\text{TEC}(Q) = C_1 \sum_{r=0}^{Q} (Q - r)p(r) + C_2 \sum_{r=Q+1}^{\propto} (r - Q)p(r). \tag{17.58}$$

Our problem is now to find $Q$ so that $TEC(Q)$ is minimum. Let an amount $Q + 1$ instead of $Q$ be ordered. Than the total expected cost given in (17.58) reduces to

$$\mathrm{TEC}(Q+1) = C_1 \sum_{r=0}^{Q} (Q+1-r)p(r) + C_2 \sum_{r=Q+2}^{\propto} (r-Q-1)p(r).$$

On simplification, we have

$$\begin{aligned}\mathrm{TEC}(Q+1) &= C_1 \sum_{r=0}^{Q} (Q-r)p(r) + C_2 \sum_{r=Q+1}^{\propto} (r-Q)p(r) + C_1 \sum_{r=0}^{Q} p(r) - C_2 \sum_{r=Q+1}^{\propto} p(r) \\ &= \mathrm{TEC}(Q) + (C_1+C_2)\sum_{r=0}^{Q} p(r) - C_2 \left[\text{Since } \sum_{r=Q+1}^{\propto} p(r) = 1 - \sum_{r=0}^{Q} p(r)\right].\end{aligned}$$

$$\therefore\ TEC(Q+1) - TEC(Q) = (C_1+C_2)\sum_{r=0}^{Q} p(r) - C_2. \tag{17.59}$$

Similarly, when an amount $Q-1$ instead of $Q$ is ordered, then we have

$$\begin{aligned}&\mathrm{TEC}(Q-1) = \mathrm{TEC}(Q) - (C_1+C_2)\sum_{r=0}^{Q-1} p(r) + C_2 \\ &\text{or } \mathrm{TEC}(Q-1) - \mathrm{TEC}(Q) = -(C_1+C_2)\sum_{r=0}^{Q-1} p(r) + C_2.\end{aligned} \tag{17.60}$$

For optimal $Q^*$, we must have

$$\mathrm{TEC}(Q^*+1) - \mathrm{TEC}(Q^*) > 0.$$

From (17.59), we have

$$\begin{aligned}&(C_1+C_2)\sum_{r=0}^{Q^*} p(r) - C_2 > 0 \\ &\text{or } \sum_{r=0}^{Q^*} p(r) > \frac{C_2}{C_1+C_2}.\end{aligned} \tag{17.61}$$

Similarly, for optimal $Q^*$, $\mathrm{TEC}(Q^* - 1) - \mathrm{TEC}(Q^*) > 0$.

From (17.60), we have

$$\begin{aligned} &-(C_1+C_2)\sum_{r=0}^{Q^*-1} p(r)+C_2>0 \\ &\text{or}\ \sum_{r=0}^{Q^*-1} p(r)<\frac{C_2}{C_1+C_2}. \end{aligned} \tag{17.62}$$

Thus, combining (17.61) and (17.62), we have

$$\begin{aligned} &\sum_{r=0}^{Q^*-1} p(r)<\frac{C_2}{C_1+C_2}<\sum_{r=0}^{Q^*} p(r) \\ &\text{or}\ \ p(r<Q^*-1)<\frac{C_2}{C_1+C_2}<p(r\le Q^*), \end{aligned} \tag{17.63}$$

where $p(r\le Q^*)$ represents the probability for $r<Q^*$.

**Continuous case: *where r is a continuous variable***

Let the demand $r$ be a continuous variable with the probability density function *f(r)*. Then proceeding as before, the total expected cost for this model is

$$\mathrm{TEC}(Q)=C_1\int_0^Q (Q-r)f(r)\mathrm{d}r+C_2\int_Q^\propto (r-Q)f(r)\mathrm{d}r.$$

$$\begin{aligned} \therefore \frac{\mathrm{d}TEC(Q)}{\mathrm{d}Q} &= C_1\int_0^Q f(r)\mathrm{d}r - C_2\int_Q^\propto f(r)\mathrm{d}r \\ &\qquad\text{[Using Leibniz's rule for differentiation under the integral sign]} \\ &= C_1\int_0^Q f(r)\mathrm{d}r - C_2\left[\int_0^\propto f(r)\mathrm{d}r - \int_0^Q f(r)\mathrm{d}r\right] \\ &= -C_2+(C_1+C_2)\int_0^Q f(r)\mathrm{d}r. \end{aligned}$$

For the optimum value of $Q$, we must have

$$\frac{\mathrm{dTEC}(Q)}{\mathrm{d}Q}=0,$$

i.e. $(C_1 + C_2)\int_0^Q f(r)\mathrm{d}r = C_2$

$$\text{or} \quad \int_0^Q f(r)\mathrm{d}r = \frac{C_2}{C_1 + C_2}. \tag{17.64}$$

Moreover, it can be proved that

$$\frac{\mathrm{d}^2\mathrm{TEC}(Q)}{\mathrm{d}Q^2} = (C_1 + C_2)f(Q) > 0.$$

Therefore, the optimum value of $Q$, i.e. $Q^*$, is given by (17.64).

*Example 9* A newspaper boy buys papers for Rs. 1.40 each and sells them for 2.00. He cannot return the unsold newspapers. Daily demand has the following distribution.

| No. of customers | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 |
|---|---|---|---|---|---|---|---|---|---|---|
| Probability | 0.01 | 0.03 | 0.06 | 0.10 | 0.20 | 0.25 | 0.15 | 0.10 | 0.05 | 0.05 |

If each day's demand is independent of the previous day's demand, how many papers should he ordered each day?

**Solution**

Let $Q$ be the number of newspapers ordered per day and $r$ be the demand for them, i.e. the number that are actually sold per day.

In this problem, the holding cost per paper per day is $C_1$ = Rs. 1.40, and the shortage cost per paper per day is $C_2$ = Rs. (2.00–1.40) = Rs. 0.60.

Now to obtain the optimal solution, we shall find the cumulative probability of daily demand of newspapers.

Calculations are given in the following table.

| $r$ | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 |
|---|---|---|---|---|---|---|---|---|---|---|
| $p(r)$ | 0.01 | 0.03 | 0.06 | 0.10 | 0.20 | 0.25 | 0.15 | 0.10 | 0.05 | 0.05 |
| $\sum_{r=23}^{Q} p(r)$ | 0.01 | 0.04 | 0.10 | 0.20 | 0.40 | 0.65 | 0.80 | 0.90 | 0.95 | 1.00 |

The required optimum value for $Q^*$ is determined by

$$\sum_{r=23}^{Q^*-1} p(r) < \frac{C_2}{C_1 + C_2} < \sum_{r=23}^{Q^*} p(r)$$

$$\text{or} \quad \sum_{r=23}^{Q^*-1} p(r) < 0.3 < \sum_{r=23}^{Q^*} p(r) \quad \left[\because \ \frac{C_2}{C_1 + C_2} = \frac{0.60}{1.40 + 0.60} = \frac{3}{10} = 0.3\right].$$

From the table, we have

$$\sum_{r=23}^{27} p(r) = 0.40 > 0.3 \text{ and } \sum_{r=23}^{26} p(r) = 0.20 < 0.3.$$

$$\therefore \sum_{r=23}^{26} p(r) < 0.3 < \sum_{r=23}^{27} p(r).$$

Hence, $Q^* = 27$; i.e. the optimum number of newspapers to be ordered is 27.

*Example 10* The demand for a certain product has a rectangular distribution between 4000 and 5000. Find the optimal order quantity if the storage cost is Re. 1.00 per unit and the shortage cost is Rs. 7.00 per unit.

**Solution**
Here $C_1$ = Re. 1.00, $C_2$ = Rs. 7.00, $C_3$ = 0.

The required optimum value for $Q^*$ is determined by

$$\int_0^{Q^*} f(r)\mathrm{d}r = \frac{C_2}{C_1 + C_2}. \tag{17.65}$$

Since the distribution is rectangular, the probability density function is given by $f(r) = \frac{1}{b-a}$, $a \le r \le b$. Here $a$ = 4000 and $b$ = 5000.

Hence, from (17.65), we have

$$\int_{4000}^{Q^*} \frac{1}{5000 - 4000}\mathrm{d}r = \frac{7}{1+7} \text{ or } \int_{4000}^{Q^*} \frac{1}{1000}\mathrm{d}r = \frac{7}{8}$$

$$\text{or } \frac{Q - 4000}{1000} = \frac{7}{8} \text{ or } Q^* = 4875.$$

Hence, the optimal order quantity is $Q^* = 4875$ units.

*Example 11* An ice cream company sells one type of its ice cream by weight. If the product is not sold on the day it is prepared, it can be sold for a loss of Rs. 0.50 per pound. But there is an unlimited market for one-day-old ice cream. On the other hand, the company makes a profit of Rs. 3.20 on every pound of ice cream sold on the day it is prepared. If daily orders form a distribution with $f(x) = 0.02 - 0.0002x$, $0 \le x \le 100$, how many pounds of ice cream should the company prepare every day?

**Solution**

It is given that $C_1$ = Rs. 0.50 and $C_2$ = Rs. 3.20.

Let $Q$ be the amount of ice cream prepared every day. The required optimum for $Q^*$ is determined by

$$\int_0^{Q^*} f(x)\mathrm{d}x = \frac{C_2}{C_1 + C_2}$$

$$\text{or} \quad \int_0^{Q^*} (0.02 - 0.0002x)\mathrm{d}x = \frac{32}{0.5 + 3.2}$$

$$\text{or} \quad 0.0002Q^{*2} + 0.04Q^* + 1.730 = 0.$$

On solving, $Q^*$ = 136.7 and 63.5.

But $Q^*$ = 136.7 is not possible, as $0 \leq x \leq 100$.

Hence, $Q^*$ = 63.5 lb is the optimum quantity; i.e. the company will prepare 63.5 lb of ice cream per day.

## 17.12 Fuzzy Inventory Model

In this section, we shall discuss the solution procedure of a fuzzy non-linear programming problem.

A deterministic or crisp non-linear programming problem may be defined as follows:

$$\begin{aligned}
&\text{Minimize } g_0(x, C_0) \\
&\text{subject to} \\
&\qquad g_i(x, C_i) \leq b_i, \quad i = 1, 2, \ldots, m \\
&\text{and } \ x \geq 0,
\end{aligned} \tag{17.66}$$

where $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ is a variable vector, $f$, $g_i$'s are algebraic expressions in $x$ with coefficients $C_0 = (C_{01}, C_{02}, \ldots, C_{0l_0})$ and $C_i = (C_{i1}, C_{i2}, \ldots, C_{il_i})$ respectively.

Introducing fuzziness in crisp parameters, the Problem (17.66) reduces to

$$\begin{aligned}
&\text{Minimize } g_0\left(x, \widetilde{C_0}\right) \\
&\text{subject to} \\
&\qquad g_i\left(x, \widetilde{C_i}\right) \leq \widetilde{b}_i, \quad i = 1, 2, \ldots, m \\
&\text{and } \ x \geq 0,
\end{aligned} \tag{17.67}$$

where the wavy bar ($\sim$) represents the fuzziness of the parameters.

According to fuzzy set theory, the fuzzy objective, coefficients and constraints are defined by their membership functions, which may be either linear or non-linear. According to Bellman and Zadeh (1970) and following the techniques of Carlsson and Korhonen (1986) and Trappy et al. (1988), the problem (17.67) is transformed into a new optimization problem as follows:

$$\begin{aligned}
&\text{Maximize } \alpha \\
&\text{subject to} \\
&\qquad g_i\left(x, \mu_{C_i}^{-1}(\alpha)\right) \le \mu_{b_i}^{-1}(\alpha), \quad i = 0, 1, 2, \ldots, m \\
&\text{and} \quad x \ge 0,
\end{aligned} \tag{17.68}$$

where $\mu_{C_i}(x) = \left\{\mu_{C_{i1}}(x), \mu_{C_{i2}}(x), \ldots, \mu_{C_{il_i}}(x)\right\}$ are the membership functions of the fuzzy coefficients, and $\mu_{b_i}(x), \quad i = 1, 2, \ldots, m$ are the membership functions of the fuzzy objective and fuzzy constraints.

Here the additional variable $\alpha$ is known as the aspiration level.

Therefore, the Lagrangian function $L(\alpha, x, \lambda)$ is given by

$$L(\alpha, x, \lambda) = \alpha - \sum_{i=0}^{m} \lambda_i \left\{ g_i\left(x, \mu_{C_i}^{-1}(\alpha)\right) - \mu_{b_i}^{-1}(\alpha) \right\},$$

where $\lambda = (\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_m)^{\mathrm{T}}$ is the Lagrange multiplier vector.

Now the Kuhn-Tucker necessary conditions are

$$\begin{aligned}
&\frac{\partial L}{\partial x_j} = 0, \quad j = 1, 2, \ldots, n \\
&\frac{\partial L}{\partial \alpha} = 0 \\
&\lambda_i \left\{ g_i\left(x, \mu_{C_i}^{-1}(\alpha)\right) - \mu_{b_i}^{-1}(\alpha) \right\} = 0 \\
&g_i\left(x, \mu_{C_i}^{-1}(\alpha)\right) - \mu_{b_i}^{-1}(\alpha) \le 0, \quad i = 0, 1, 2, \ldots, m \\
&\lambda_i \le 0.
\end{aligned} \tag{17.69}$$

Solving the problem (17.69), the corresponding optimal solution can be obtained.

In a deterministic EOQ model, the problem is to determine the order level $Q$ (>0) which minimizes the average total cost $C(Q)$, i.e.

$$\begin{aligned}
&\text{Min}\, C(Q) = C_1 Q/2 + C_3 D/Q \\
&\text{subject to} \\
&\qquad AQ \leq B, Q > 0,
\end{aligned} \tag{17.70}$$

where

$C_3$ = ordering cost per order,
$C_1$ = holding cost per unit quantity per unit time,
$D$ = demand per unit time,
$B$ = maximum available warehouse space (in square feet),
$A$ = the space required by each unit (in square units).

When the preceding objective goal, costs and available storage area become fuzzy, the problem (17.70) is transformed to:

$$\begin{aligned}
&\text{Min} \quad \widetilde{C(Q)} = \widetilde{C_1} Q/2 + \widetilde{C_3} D/Q \\
&\text{subject to} \\
&\qquad AQ \leq \widetilde{B}, \quad Q > 0.
\end{aligned} \tag{17.71}$$

So the corresponding fuzzy non-linear programming problem is

$$\begin{aligned}
&\text{Max } \alpha \\
&\text{subject to} \\
&\qquad \mu_{C_1}^{-1}(\alpha) Q/2 + \mu_{C_3}^{-1}(\alpha) D/Q \leq \mu_{C_0}^{-1}(\alpha) \\
&\qquad AQ \leq \mu_B^{-1}(\alpha), \quad Q > 0, \quad \alpha \in [0, 1],
\end{aligned}$$

where $\mu_{C_3}(x), \mu_{C_1}(x), \mu_{C_0}(x)$ and $\mu_B(x)$ are the membership functions for the fuzzy ordering cost, holding cost, objective goal and storage area respectively.

Here the Lagrangian function is

$$\begin{aligned}
L(\alpha, Q, \lambda_1, \lambda_2) = \alpha - \lambda_1 &\left\{ \mu_{C_1}^{-1}(\alpha) Q/2 + \mu_{C_3}^{-1}(\alpha) D/Q - \mu_{C_0}^{-1}(\alpha) \right\} \\
&- \lambda_2 \left\{ AQ - \mu_B^{-1}(\alpha) \right\},
\end{aligned} \tag{17.72}$$

where $\lambda_1$ and $\lambda_2$ are Lagrange multipliers.

Now, we consider three combinations of different types of membership functions to represent the fuzzy goal, costs and warehouse space.

**Fuzzy goal, costs and storage area represented by linear membership functions**

In this case, $\mu_{C_i}(i = 1, 3)$ is given by

$$
\begin{aligned}
\mu_{C_i}(u) &= 1 && \text{for } u > C_i \\
&= 1 - (C_i - u)/P_i && \text{for } C_i - P_i \le u \le C_i (i = 1, 3) \\
&= 0 && \text{for } u < C_i - P_i.
\end{aligned}
$$

Graphically, this function is represented in Fig. 17.9.

$$
\begin{aligned}
\text{Again, } \mu_{C_0}(C(Q)) &= 1 && \text{for } C(Q) < C_0 \\
&= 1 - (C(Q) - C_0)/P_0 && \text{for } C_0 \le C(Q) \le C_0 + P_0 \\
&= 0 && \text{for } C(Q) > C_0 + P_0.
\end{aligned}
$$

Its pictorial representation is given in Fig. 17.10.

The linear membership function for the storage area is given by

$$
\begin{aligned}
\mu_B(AQ) &= 1 && \text{for } AQ < B \\
&= 1 - (AQ - B)/P && \text{for } B \le AQ \le B + P \\
&= 0 && \text{for } AQ > B + P.
\end{aligned}
$$

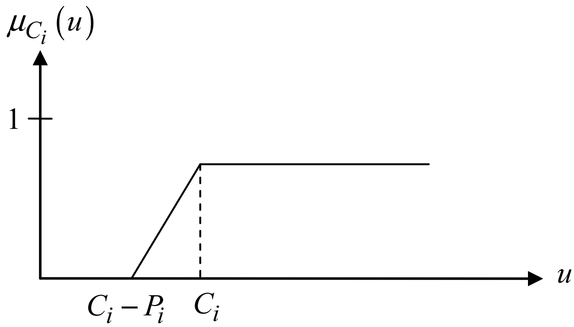



**Fig. 17.9** Linear membership function of $C_i$

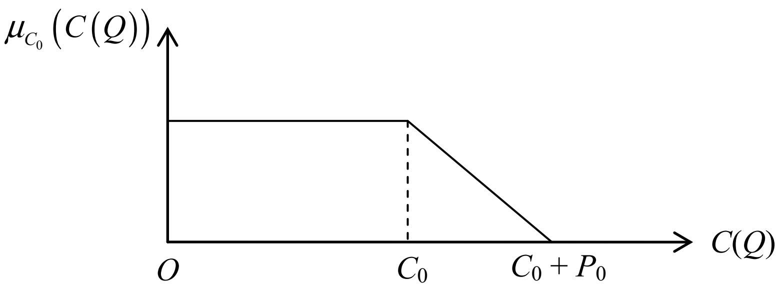



**Fig. 17.10** Linear membership function of $\widetilde{C}(Q)$

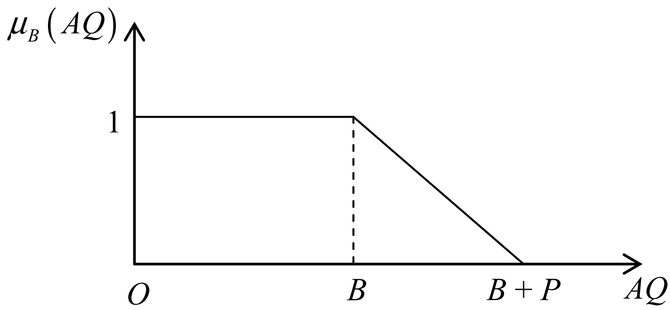



**Fig. 17.11** Linear membership function of $AQ$

This membership function is depicted in Fig. 17.11.

Here $P_0$, $P$ and the $P_i$'s ($i$ =1, 3) are the maximally acceptable violations of the aspiration levels $C_0$, $B$ and the $C_i$'s ($i$ = 1, 3). Considering the nature of these parameters, we assume the membership function to be non-decreasing for fuzzy inventory costs and non-increasing for the fuzzy goal and storage area.

Hence,

$$\begin{aligned}
&\mu_{C_i}^{-1}(\alpha) = C_i - (1-\alpha)P_i, \quad (i = 1, 3) \\
&\mu_{C_0}^{-1}(\alpha) = C_0 + (1-\alpha)P_0 \\
&\mu_B^{-1}(\alpha) = B + (1-\alpha)P.
\end{aligned}$$

From (17.72), we have

$$\begin{aligned}
L(\alpha, Q, \lambda_1, \lambda_2) = \alpha - \lambda_1[&\{C_1 - (1-\alpha)P_1\}Q/2 + \{C_3 - (1-\alpha)P_3\}D/Q - C_0 - (1-\alpha)P_0] \\
&- \lambda_2\{AQ - B - (1-\alpha)P\}.
\end{aligned}$$

Now, the Kuhn-Tucker necessary conditions are

$$\begin{aligned}
&\frac{\partial L}{\partial \alpha} = 0 \\
&\frac{\partial L}{\partial Q} = 0 \\
&\{C_1 - (1-\alpha)P_1\}Q/2 + \{C_3 - (1-\alpha)P_3\}D/Q - C_0 - (1-\alpha)P_0 = 0 \\
&AQ - B - (1-\alpha)P = 0.
\end{aligned}$$

Solving these equations, the expression for the optimal order quantity is

$$Q^* = \{B + (1 - \alpha^*)P\}/A,$$

where $\alpha^*$ is a root of

$$\begin{aligned} &P_1P^2(1-\alpha)^3 - \left(C_1P^2 - 2BPP_1 - 2AP_0P\right)(1-\alpha)^2 \\ &\quad - \left(2BC_1P - P_1B^2 - 2DA^2P_3 - 2APC_0 - 2AP_0B\right)(1-\alpha) \\ &\quad - \left(C_1B^2 + 2DA^2C_3 - 2AC_0B\right) = 0. \end{aligned}$$

**Fuzzy goal and storage area represented by parabolic concave membership functions and costs by linear membership functions**

In this case, the membership functions $\mu_{C_i}(u), i = 1, 3$ for the fuzzy inventory cost will be the same as defined in the preceding section.

The membership function $\mu_{C_0}(C(Q))$ for the fuzzy goal is defined as follows:

$$\begin{aligned} \mu_{C_0}(C(Q)) &= 1 && \text{for } C(Q) < C_0 \\ &= 1 - [\{C(Q) - C_0\}/P_0]^2 && \text{for } C_0 \le C(Q) \le C_0 + P_0 \\ &= 0 && \text{for } C(Q) > C_0 + P_0. \end{aligned}$$

It is graphically represented in Fig. 17.12.

Again, the membership function for the storage area is defined as

$$\begin{aligned} \mu_B(AQ) &= 1 && \text{for } AQ < B \\ &= 1 - [(AQ - B)/P]^2 && \text{for } B \le AQ \le B + P \\ &= 0 && \text{for } AQ > B + P. \end{aligned}$$

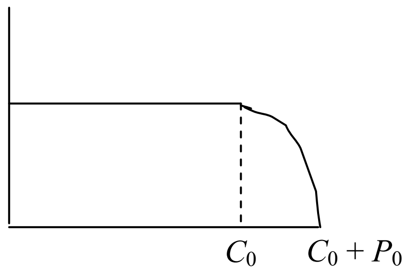



**Fig. 17.12** Parabolic membership function of $\widetilde{C}(Q)$

Proceeding exactly as before, the optimal order quantity is given by

$$Q^* = \left\{B + \sqrt{(1-\alpha^*)}P\right\}/A,$$

where $\alpha^*$ is a root of

$$P_1P^2(1-\alpha)^2 + 2BPP_1(1-\alpha)^{\frac{3}{2}} - \left(C_1P^2 - P_1B^2 - 2A^2DP_3 - 2AP^2\right)(1-\alpha)$$
$$- (2BPC_1 - 2AC_0P - 2APB)(1-\alpha)^{\frac{1}{2}} - \left(C_1B^2 + 2A^2DC_3 - 2AC_0B\right) = 0.$$

## 17.13 Exercises

1. The demand rate of a particular item is 12,000 units per year. The ordering cost per order is Rs. 350.00, and the holding cost is Rs. 0.20 per unit per month. If no shortage is allowed and the replenishment is instantaneous, determine (i) the optimal lot size, (ii) the optimum scheduling period, (iii) the minimum total average cost.
2. The annual requirement for a product is 3000 units. The ordering cost is Rs. 100.00 per order. The cost per unit is Rs. 10.00. The carrying cost per unit per year is 30% of the unit cost.

   (a) Find the EOQ.
   (b) If a new EOQ is found by using the ordering cost as Rs. 80.00, what would be the further savings in cost?

3. The demand rate for an item of a company is 18,000 units per year. The company can produce at the rate of 3000 units per month. The setup cost is Rs. 500.00 per order, and the holding cost is 0.15 per units per month. Calculate: (i) Optimum manufacturing quantity, (ii) the time of manufacture, (iii) cycle length, (iv) the optimum annual cost if the cost of an item is Rs. 2.00 per unit.
4. A contractor has to supply 10,000 bearings per day to an automobile manufacturer. He finds that when he starts a production run he can produce 25,000 bearings per day. The holding cost of a bearing in stock is Rs. 0.02 per year. The setup cost of a production is Rs. 18.00. How frequently should production runs be made?
5. The demand for an item is 18,000 units per year. The holding cost is Rs. 1.20 per unit per unit item, and the shortage cost is Rs.5.00. The ordering cost is Rs. 400.00. Assuming the replenishment rate as instantaneous, determine the optimum order quantity along with the total minimum cost of the system.
6. Find the optimum order level which minimizes the total expected cost under the following assumptions:

   (i) $t$ is the constant interval between orders.

(ii) $Q$ is the stock (in discrete units) at the beginning.
(iii) $d$ is the estimated random instantaneous demand at a discontinuous rate.
(iv) $C_1$ and $C_2$ are the holding and shortage costs per item per $t$ time unit.
(v) The level time is zero.

7. Let $F(x)$ be the probability density function of the demand $x$ in a prescribed scheduling period of $T$ weeks. The demand is assumed to occur with a uniform pattern, and the probability distribution is continuous. The unit carrying cost and the shortage cost are respectively $C_1$ money units and $C_2$ money units, both per unit per week. There is no setup cost for the system. Obtain the expression for optimal control.
8. Let the probability density of demand of a certain item during a week be

$$f(x) = \begin{cases} 0.1, & 0 \leq x \leq 10 \\ 0, & \text{Otherwise} \end{cases}.$$

This demand is assumed to occur with a uniform pattern over the week. Let the unit carrying cost of the item in inventory be Rs. 2.00 per week and the unit shortage cost be Rs. 8.00 per week. Determine the optimal order level of the inventory and the total minimum cost of the inventory system.
9. The annual demand for a product is 500 units. The cost of storage per unit per year is 10% of the unit cost. The ordering cost is Rs. 180 for each order. The unit cost depends upon the amount ordered. The range of amount ordered and the unit price are as follows:

| Range of amount ordered | Unit cost (Rs.) |
|---|---|
| $1 \leq q < 500$ | 25.00 |
| $500 \leq q < 1500$ | 24.80 |
| $1500 \leq q < 3000$ | 24.60 |
| $3000 \leq q$ | 24.40 |

Find the optimal order quantity.
10. Show that when determining the optimal inventory level which minimizes the total expected cost in case of continuous (non-discrete) quantities, the condition to be satisfied is

$$F(S_0) = \frac{C_2}{C_1 + C_2},$$

where

$$F(S_0) = \int_0^{S_0} f(r)\mathrm{d}r,$$

$f(r)$ = the probability density function of the requirement of $r$ quantity
$C_2$ = shortage cost per unit per unit time

$C_1$ = holding cost per unit per unit time.

11. An ice cream company sells one of its types of ice cream by weight. If the product is not sold on the day it is prepared, it can be sold at a loss of 50 paise per pound. There is, however, an unlimited market for one-day-old ice cream. On the other hand, the company makes a profit of Rs. 3.20 on every pound of ice cream sold on the day it is prepared. Past daily orders form a distribution with

$$f(x) = 0.02 - 0.0002x, \quad 0 \le x \le 100.$$

How many pounds of ice cream should the company prepare every day?

12. The following data describe three inventory items. Determine the economic order quantity for each of the three items to be accommodated within a total available storage area of 25 square feet.

| Item | Setup cost (Rs.) | Demand (units per day) | Unit holding cost per unit time (Rs.) | Storage area required per unit ($ft^2$) |
|---|---|---|---|---|
| 1 | 10 | 2 | 0.3 | 1 |
| 2 | 5 | 4 | 0.1 | 1 |
| 3 | 15 | 4 | 0.2 | 1 |

Find the economic lot size for the inventory model with finite replenishment rate, shortages allowed but fully backlogged, uniform finite demand and zero lead time so that the total average cost is minimum.

13. A shop produces three items in lots. The demand rate for each item is constant and can be assumed to be deterministic. No backorders are allowed. The pertinent data for the items are given in the following table:

| Item | I | II | III |
|---|---|---|---|
| Carrying cost (Rs.) | 20 | 20 | 20 |
| Setup cost (Rs.) | 50 | 40 | 60 |
| Cost per unit (Rs.) | 6 | 7 | 5 |
| Yearly demand rate (units) | 10,000 | 12,000 | 7500 |

Determine approximately the economic order quantity for the three items subject to the condition that the total value of average inventory levels of these items does not exceed Rs. 1000.00.

14. Find the optimum order quantity for a product for which the price breaks are as follows:

| Quantity | Purchasing cost (per unit) |
|---|---|
| $0 \le Q_1 < 100$ | Rs. 20 |
| $100 \le Q_2 < 200$ | Rs. 18 |
| $200 \le Q_3$ | Rs. 16 |

The monthly demand for the product is 400 units. The storage cost is 20% of the unit cost of the product, and the cost of ordering is Rs. 25.00 per month.

15. A newspaper boy buys papers for Rs. 1.70 each and sells them for Rs. 2.00 each. He cannot return unsold newspapers. The daily demand has the following distribution:

| No. of customers | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 |
|---|---|---|---|---|---|---|---|---|---|---|
| Probability | 0.01 | 0.03 | 0.06 | 0.10 | 0.20 | 0.25 | 0.15 | 0.10 | 0.05 | 0.05 |

If each day's demand is independent of the previous day's demand, how many papers should he order each day?

16. A small shop produces three machine parts 1, 2, 3 in lots. The shop has only 700 $m^2$ of storage space. The appropriate data for the three items are presented in the following table:

| Item | 1 | 2 | 3 |
|---|---|---|---|
| Demand (unit per year) | 2000 | 5000 | 10000 |
| Setup cost (Rs.) | 100 | 200 | 75 |
| Cost per unit (Rs.) | 10 | 20 | 5 |
| Floor space required (sq.mt/unit) | 0.50 | 0.60 | 0.30 |

The shop uses an inventory carrying charge of 20% of the average inventory valuation per annum. If no stock-outs are allowed, determine the optimal lot size for each item.

# Chapter 18
# Mathematical Preliminaries

## 18.1 Objective

The objective of this chapter is to discuss some mathematical background of several topics, like matrices, determinants, vectors, probability and Markov processes, which will be very helpful for studying the different subject matters of this book.

## 18.2 Matrices

A rectangular array of $mn$ elements ($m$, $n$ being positive integers), arranged in $m$ rows and $n$ columns as shown

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{or } A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

is called matrix of order ($m \times n$) or an $m$ by $n$ matrix.

The numbers $a_{ij}$ are called the elements of the matrix.

A useful, easy way of writing the matrix shown below:

$$A = \left[a_{ij}\right]_{m \times n} \text{or} \left(a_{ij}\right)_{m \times n}, \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n.$$

In general, matrices are denoted by boldface uppercase letters $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, ….

© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2_18

**Square matrix:** If $m = n$, the matrix $A$ is called a square matrix of order $n$ (or of order $m$).

**Row matrix:** If $m = 1$, i.e. the matrix $A$ has only one row, then $A$ is called a row matrix (or sometimes called an $n$-dimensional row vector).

**Column matrix:** If $n = 1$, i.e. the matrix $A$ has only one column, then $A$ is called a column matrix (or sometimes called an $m$-dimensional column vector).

**Equality of matrices:** Two matrices $A$ and $B$ are said to be equal if $A$ and $B$ are of the same order and each element of $A$ is the same as the corresponding element of $B$; i.e. if $A = \left[a_{ij}\right]_{m\times n}$ and $B = \left[b_{ij}\right]_{m\times n}$, then $A = B$ only when $a_{ij} = b_{ij}$ for each $i = 1,\ 2,\ \ldots,\ m$ and each $j = 1,\ 2,\ \ldots,\ n$.

It is to be noted that the question of equality between two matrices of different orders does not arise.

**Null/zero or void matrix:** A matrix $A = \left[a_{ij}\right]_{m\times n}$ is called a null matrix or a void matrix of order $m \times n$, if every element of $A$ is 0 (zero), i.e. $a_{ij} = 0$ for each $i = 1,\ 2,\ \ldots,\ m$ and $j = 1,\ 2,\ \ldots,\ n$. The null matrix is denoted by $O_{m\times n}$ or simply $O$.

**Diagonal element of a matrix $A$**

Let $A = \left[a_{ij}\right]_{m\times n}$ be a square matrix of order $n$. The diagonal through the element at the left-hand top corner of $A$ is called the principal (or leading) diagonal or simply the diagonal of $A$, and the elements of this diagonal are called the diagonal elements of $A$.

$$\text{Thus, if } A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

then the diagonal elements of $A$ are $a_{11},\ a_{22},\ \ldots,\ a_{nn}$.

**Diagonal matrix:** If all the elements of $A$ other than the diagonal elements are zero, i.e. $a_{ij} = 0$ for all $i, j$ with $i \neq j$, then the matrix $A$ is called a diagonal matrix and is denoted by $(a_{11},\ a_{22},\ a_{33},\ \ldots,\ a_{nn})$.

**Scalar matrix:** If $A$ is a diagonal matrix with all the diagonal elements are same, i.e. $a_{11} = a_{22} = \cdots = a_{nn}$ and $a_{ij} = 0$ for all $i, j$ with $i \neq j$, then $A$ is said to be a scalar matrix.

**Identity matrix or unit matrix:** If $A$ be the scalar matrix with each diagonal element as unity, i.e. $a_{11} = a_{22} = \cdots = a_{nn} = 1$ and $a_{ij} = 0$ for all $i, j$ with $i \neq j$, then $A$ is said to be a unit matrix and is denoted by $I_n$.

**Upper triangular matrix:** If all the elements below the diagonal elements are zero, i.e. $a_{ij} = 0$ for $i > j$, then the matrix is called an upper triangular matrix.

**Lower triangular matrix:** If all the elements above the diagonal elements are 0, i.e. $a_{ij} = 0$ for $(i < j)$, then the matrix is called a lower triangular matrix.

**Triangular matrix:** A matrix $A$ is called a triangular matrix if it is either an upper triangular or a lower triangular matrix.

Here it is mentioned that a matrix is a diagonal matrix if and only if it is both upper triangular and lower triangular.

**Addition of matrices**

Two matrices $A$ and $B$ are said to be conformable for addition of $A$ and $B$ with same order (i.e. addition of a matrix $A$ with a matrix $B$ is defined only when they are of the same order). If $A = \left[a_{ij}\right]_{m\times n}$ and $B = \left[b_{ij}\right]_{m\times n}$ are two matrices, then $C = A + B$ is a matrix of order $m \times n$, whose elements $c_{ij}$ are given by

$$c_{ij} = a_{ij} + b_{ij} \text{ for all } i \text{ and } j.$$

Thus, if $A = \begin{bmatrix} 2 & 1 & -2 \\ 5 & 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 4 & 7 \\ -1 & 2 & 4 \end{bmatrix}$,

then $C = A + B = \begin{bmatrix} 2 & 1 & -2 \\ 5 & 3 & 4 \end{bmatrix} + \begin{bmatrix} 3 & 4 & 7 \\ -1 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 5 & 5 \\ 4 & 5 & 8 \end{bmatrix}$.

If $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 5 & 3 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & -3 \\ 1 & 4 & -1 \end{bmatrix}$,

then these two matrices are not conformable for addition.

**Negative of a matrix:** Let $A = \left[a_{ij}\right]_{m\times n}$ be a matrix. Then the matrix $B$ of order $m \times n$ is called the negative of $A$ if $A + B = O_{m\times n}$. Clearly, $B$ is the matrix $\left[-a_{ij}\right]_{m\times n}$ and we denote the negative of $A$ by $-A$, i.e. $-A = \left[-a_{ij}\right]_{m\times n}$.

If $A$ and $B$ are two matrices of the same order, say $A = \left[a_{ij}\right]_{m\times n}$, $B = \left[b_{ij}\right]_{m\times n}$, then we define the difference between $A$ and $B$ to be a matrix of order $m \times n$ denoted by $A - B$ and given by $A - B = A + (-B)$.

$$\therefore A - B = \left(a_{ij}\right)_{m\times n} + \left(-b_{ij}\right)_{m\times n} = \left(a_{ij} - b_{ij}\right)_{m\times n}.$$

**Multiplication by a scalar:**

Let $A = \left[a_{ij}\right]_{m\times n}$ be a matrix of order $(m \times n)$ and $c$ be a scalar. Then the scalar multiple $c\,A$ is defined to be a matrix of order $(m \times n)$, each element of which is $c$ times the corresponding element of $A$,

$$\text{i.e.} \quad cA = [ca_{ij}]_{m\times n}.$$

If $c = 5$, $A = \begin{bmatrix} 3 & 6 \\ 9 & -1 \\ 2 & 0 \end{bmatrix}$, then $cA = \begin{bmatrix} 15 & 30 \\ 45 & -5 \\ 10 & 0 \end{bmatrix}$.

***Some algebric properties of matrices***

(i) $A + B = B + A$ (matrix addition is commutative)
(ii) $A + (B + C) = (A + B) + C$ (matrix addition is associative)
(iii) $(c_1 + c_2)A = c_1A + c_2A$, where $c_1$ and $c_2$ are scalars
(iv) $c(A + B) = cA + cB$, where $c$ is a scalar
(v) $(-1)A = -A$
(vi) $A - B = A + (-B)$
(vii) $c(-A) = (-c)A$, where $c$ is a scalar
(viii) $-(-A) = A = 1 \cdot A$
(ix) $c(A - B) = cA + c(-B) = cA + (-c)B = cA - cB$
(x) $(c_1 - c_2)A = c_1A - c_2A$.

**Multiplication of matrices**

Let $A$ and $B$ be two matrices. Then the product of $A$ and $B$ denoted by $AB$ is defined if and only if the number of columns of $A$ is equal to the number of rows of $B$.

Thus, if $A = [a_{ij}]_{m\times p}$ and $B = [b_{ij}]_{p\times n}$ matrix, then the product

$C = AB = [c_{ij}]$ will exist and will be an $m \times n$ matrix where the elements $c_{ij}$ of $C$ are given by

$$c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}, \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n.$$

Let $A = \begin{pmatrix} 10 & -1 \\ 21 & 3 \end{pmatrix}_{2\times 3}$ and $B = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 3 \\ 2 & 1 & -2 \end{pmatrix}_{3\times 3}$.

$$\therefore \quad AB = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 3 \\ 2 & 1 & -2 \end{pmatrix}$$
$$= \begin{pmatrix} 1\times 1 + 0\times(-1) + (-1)\times 2 & 1\times 0 + 0\times 2 + (-1)\times 1 & 1\times 0 + 0\times 3 + (-1)\times(-2) \\ 2\times 1 + 1\times(-1) + 3\times 2 & 2\times 0 + 1\times 2 + 3\times 1 & 2\times 0 + 1\times 3 + 3\times(-2) \end{pmatrix}$$
$$= \begin{pmatrix} -1 & -1 & 2 \\ 7 & 5 & -3 \end{pmatrix}.$$

**Note**

(i) Matrix multiplication satisfies the associative and the distributive laws. For the matrices $A$, $B$, $C$,

$$(AB)C = A(BC),$$
$$(A+B)C = AC+BC \quad \text{and} \quad A(B+C) = AB+AC$$

provided that multiplication and/or addition are defined.

(ii) For any two matrices $A$, $B$:

- Neither $AB$ nor $BA$ may be defined.
- $AB$ may be defined but $BA$ is not defined and vice versa.
- $AB$ and $BA$ are both defined if the number of rows of $A$ = number of columns of $B$ and number of columns of $A$ = number of rows of $B$.

(iii) Matrix multiplication is not, in general, commutative; i.e. for two matrices $A$, $B$, the equality $AB = BA$ need not be true (even if $AB$ and $BA$ are defined).

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

$AB = \begin{pmatrix} 5+2 & 6+4 \\ 15+4 & 18+8 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 19 & 26 \end{pmatrix}$ and

$BA = \begin{pmatrix} 5+18 & 10+24 \\ 1+6 & 2+8 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 7 & 10 \end{pmatrix}$.

Thus, $AB \neq BA$.

(iv) If $A$ is a square matrix of order $(n \times n)$ or $n$ (say), then $A$ commutes *with* $I_n$, the identity matrix. If $O_n$ be the null matrix and $A$ be any scalar matrix then

- $A I_n = I_n A = A$
- $O_n A = A O_n = O_n$

**Transpose of a matrix**

Let $A$ be a matrix of order $m \times n$ (say). Then the matrix of order $n \times m$ obtained by interchanging the rows and columns of $A$ is called the transpose of $A$, to be denoted by $A^{\mathrm{T}}$ or $A'$.

So for $A = [a_{ij}]_{m\times n}$, $A^{\mathrm{T}} = [a_{ji}]_{n\times m}$.

In this process, the $i$th row of $A$ becomes the $i$th column of $A^{\mathrm{T}}$.

Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}_{2\times 3}$, $\quad \therefore A^{\mathrm{T}} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}_{3\times 2}$.

**Some properties:**

(i) $(A^{\mathrm{T}})^{\mathrm{T}} = A$
(ii) $(cA)^{\mathrm{T}} = cA^{\mathrm{T}}$, where $c$ is a scalar
(iii) $(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$
(iv) $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$

**Symmetric matrix:** A square matrix $A = [a_{ij}]_{n\times n}$ of order $n$ is said to be symmetric if $A = A^{\mathrm{T}}$, i.e. $a_{ij} = a_{ji}, \ \forall i, j$.

**Skew-symmetric matrix:** A square matrix $A = [a_{ij}]_{n\times n}$ of order $n$ is said to be a skew-symmetric matrix if $A = -A^{\mathrm{T}}$, i.e. $a_{ij} = -a_{ji} \ \forall i, j$.

**Some properties:**

(i) Every diagonal matrix is symmetric.
(ii) For any matrix $A$, $AA^{\mathrm{T}}$ and $A^{\mathrm{T}}A$ are symmetric matrices.
(iii) For a square matrix $A$, $A + A^{\mathrm{T}}$ is symmetric, while $A - A^{\mathrm{T}}$ is a skew-symmetric matrix.
(iv) The diagonal elements of a skew-symmetric matrix are all zero.

**Orthogonal matrix:** A square matrix $A$ is said to be orthogonal if $A^{\mathrm{T}}A = AA^{\mathrm{T}} = I$ (identity matrix).

**Submatrix:** A matrix formed by omitting some rows and columns of a matrix is known as a submatrix of the original matrix.

**Partitioned matrices:** Let $A$ be a matrix given by

$$A = \begin{bmatrix} a_{11} & a_{12} & \vdots & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & \vdots & a_{23} & a_{24} & a_{25} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{31} & a_{32} & \vdots & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & \vdots & a_{43} & a_{44} & a_{45} \end{bmatrix}$$

Now, if we write

$$A_{11} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad A_{12} = \begin{bmatrix} a_{13} & a_{14} & a_{15} \\ a_{23} & a_{24} & a_{25} \end{bmatrix}$$
$$A_{21} = \begin{bmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \quad A_{22} = \begin{bmatrix} a_{33} & a_{34} & a_{35} \\ a_{43} & a_{44} & a_{45} \end{bmatrix}$$

then $A$ can be written as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

Therefore, the matrix $A$ has been partitioned into four submatrices, $A_{ij},\ \ i = 1,\ 2;$ $j = 1,\ 2.$

## 18.3 Determinant

There is a number associated with every square matrix $A$ which is called the determinant of $A$ and is usually denoted by $|A|$ or det $A$.

The determinant of $A$ of order $n$ is defined as

$$|A| = \sum (\pm 1) a_{1j}\, a_{2j} \cdots a_{np},$$

the sum being taken over all the permutations of the second subscripts.

**Note** If $A$, $B$ are square matrices of the same order (say, $n$), then

(i) $|A| = |A^{\mathrm{T}}|$
(ii) $|A\,B| = |A|\ |B|$
(iii) $|k\,A| = k^n |A|$, for any scalar $k$.

**Minor:** The minor $M_{ij}$ of the element $a_{ij}$ of a matrix $A$ is the determinant obtained by striking out the $i$th row and $j$th column of matrix $A$.

Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$.

Then the minor of $a_{11}$ is $M_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$

and the minor of $a_{23}$ is $M_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$.

**Cofactor:** The cofactor of the element $a_{ij}$ of the matrix $A$, denoted by $A_{ij}$, is the coefficient of $a_{ij}$ in the expansion of $A$,

$$\text{i.e.}\quad A_{ij} = (-1)^{i+j} M_{ij}.$$

**Note** The determinant of a square matrix $A$ of order $n$ can be expressed as

$|A| = \sum\limits_{j=1}^{n} a_{ij} A_{ij}, \quad i = 1, 2, \ldots, n,$

where $A_{ij}$ is the cofactor of $a_{ij}$.

Obviously, $|A| = \sum_{i=1}^{n} a_{ij} A_{ij}, \quad j = 1,\ 2,\ \ldots,\ n.$

**Properties of determinants:**

(i) The value of a determinant remains unaltered if its rows are changed into columns and its columns into rows.
(ii) If any two rows (or columns) of a determinant are interchanged, the determinant retains its absolute value but changes its sign.
(iii) If two rows or columns of a determinant are identical, then the determinant vanishes.
(iv) If all the elements of any row (or of any column) are multiplied by the same quantity, the determinant is multiplied by the same quantity.

**Adjoint or adjugate of a square matrix:**

Let $A = \left[a_{ij}\right]_{n\times n}$ be a square matrix, and let $A_{ij}$ denote the cofactor of $a_{ij}$ in $|A|$ (for $i,\ j = 1,\ 2,\ \ldots,\ n$). Let $B = \left[A_{ij}\right]_{n\times n}$. Then the transpose $B^{\mathrm{T}}$ of $B$ is called the adjoint or adjugate of $A$ and denoted by $\mathrm{Adj}(A)$ or $\mathrm{Adj}\,A$.

**Properties:**

(i) $A \cdot (\mathrm{Adj}\,A) = (\mathrm{Adj}\,A) \cdot A = |A| \cdot I_n$
(ii) $\mathrm{Adj}(A^{\mathrm{T}}) = (\mathrm{Adj}\,A)^{\mathrm{T}}$
(iii) $\mathrm{Adj}(kA) = k^{n-1}\mathrm{Adj}(A)$, where $k$ is a scalar
(iv) $|\mathrm{Adj}\,A| = |A|^{n-1}, \quad$ if $|A| \neq 0$

**Inverse of a matrix:** Let $A$ be a square matrix of order $n$. A square matrix $B$ (if it exists) is said to be an inverse of $A$ if $AB = BA = I_n$. If such an inverse $B$ of $A$ exists, then $A$ is said to be invertible.

**Singular matrices:** A square matrix $A$ is said to be singular if $|A| = 0$.

**Non-singular matrix:** A square matrix $A$ is said to be non-singular if $|A| \neq 0$.

**Note**

(i) A square matrix is invertible if it is non-singular.
(ii) An invertible matrix has a unique inverse.
(iii) $A^{-1} = \frac{\mathrm{Adj}\,A}{|A|}, \quad$ if $|A| \neq 0$.

**Some properties of the inverse of a matrix**

If $A$, $B$ are two non-singular matrices of the same order, then

(i) $\left(A^{-1}\right)^{-1} = A$
(ii) $(AB)^{-1} = B^{-1}A^{-1}$

(iii) $(AT)^{-1} = (A^{-1})^{T}$
(iv) $(kA)^{-1} = k^{-1}A^{-1}$, for any scalar $k \neq 0$.

**Note** If $A_1, A_2, \ldots, A_m$ are $m$ invertible matrices, then each $A_1, A_2, \ldots, A_m$ is also invertible, and $(A_1 A_2 \ldots, A_m)^{-1} = A_m^{-1} A_{m-1}^{-1} \ldots, A_2^{-1} A_1^{-1}$.

**Integral powers of a square matrix**

Let $A$ be a square matrix of order $n$. Then the integral powers of $A$ are defined as follows:

(i) If $A$ is non-singular, then $A^0 = I_n$
(ii) $A^1 = A$
(iii) $A^m = A^{m-1}A$
(iv) If $A$ is non-singular, $A^{-m} = (A^{-1})^m = (A^m)^{-1}$, where $m$ is a positive integer.

## 18.4 Rank of a Matrix

Let $A$ be an $(m \times n)$ non-null matrix. The rank of matrix $A$ will be $k$ ($\leq$ min ($m$, $n$)), if the determinant of every square submatrix of $A$ of order ($k$ + 1) vanishes while there exists at least one square submatrix of order $k$ whose determinant does not vanish. That is, the rank of $A$ is the order of the largest square submatrix in $A$ whose determinant does not vanish. The rank of $A$ is denoted by $r(A)$.

$$\text{If } A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ -3 & -6 & -9 \end{pmatrix}, \text{ then } r(A) = 1, \text{ since } |A| = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ -3 & -6 & -9 \end{vmatrix} = 0$$

and all square submatrices of order 2 also vanish. Hence, $r(A) = 1$.

$$\text{If } A = \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}, \text{ then } r(A) = 2, \text{ since } |A| \neq 0.$$

**Properties:**

(i) The rank of matrix $A$ = the rank of $A^T$.
(ii) The rank of a non-singular matrix of order $n$ is $n$.
(iii) The rank of a singular square matrix of order $n$ is less than $n$.
(iv) The rank of a null matrix is defined to be zero or infinite.

## 18.5 Solution of a System of Linear Equations

The system of linear equations

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{array}\right\} \tag{18.1}$$

can be written in matrix form as

$$Ax = b$$

where $A = [a_{ij}]_{n\times n}$ is the coefficient matrix, $x = (x_1,\ x_2,\ \ldots,\ x_n)^{\mathrm{T}}$ and $b = (b_1,\ b_2,\ \ldots,\ b_n)^{\mathrm{T}}$. The system has a unique solution $x = A^{-1}b$, provided $|A| \neq 0$, i.e. the coefficient matrix $A$ is non-singular.

**Note** Let us consider a matrix $\overline{A} = [A\ b] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{bmatrix}$.

The matrix $\overline{A}$ is called the augmented matrix of the system of Eq. (18.1).

It is obvious that $r(\overline{A}) = n$, since $|A| \neq 0$ and there does not exist a square sub-matrix of $\overline{A}$ of order $(n + 1)$.

Thus, it follows that the system (18.1) has a unique solution if $r(A) = r(\overline{A}) = n$.

If $r(A) < n$, i.e. if $A$ is singular, then two cases arise:

(i) If $r(A) = r(\overline{A}) < n$, then the system (18.1) has an infinite number of solutions.
(ii) If $r(A) < r(\overline{A}) \leq n$, then the system (18.1) does not have any solutions. In this case, the system (18.1) is inconsistent.

The system (18.1) is consistent if and only if $r(A) = r(\overline{A})$, in which the necessary and sufficient condition for the unique solution of the system (18.1) is $r(A) = r(\overline{A}) = n$.

## 18.6 Vectors

**Row vector:** A row matrix with $n$ elements $a_1$, $a_2$, …, $a_n$ is known as an $n$-tuple row vector. It is usually denoted by $a = (a_1, a_2, \ldots, a_n)$.

**Column vector:** A column matrix with elements $b_1, b_2, \ldots, b_n$ is known as an $n$-tuple column vector $b$ and $b = (b_1, b_2, \ldots, b_n)^{\mathrm{T}}$.

Analytically, both the row and column vectors may be considered as points in the $n$-dimensional space, which is known as a vector space.

**Unit vector:** A vector is said to be a unit vector if all its components are zero except one with unit value. $e_i$ is unit vector in the $n$-dimensional vector space whose $i$th component is 1. Clearly, there are $n$ $n$-component unit vectors as follows:

$$e_1 = (1, 0, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, 0, \ldots, 1).$$

**Null vector:** A vector is said to be a null vector if all the components of the vector are equal to zero. The null vector is usually denoted by O. An $n$-component null vector is written as O = (0, 0, …, 0) with $n$ zeros.

**Linear combination of vectors:** The linear combination of a set of $k$ $n$-component vectors $a_1$, $a_2$, …, $a_k$ is a vector $a$ with $n$-components given by the relation

$$a = \lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_k a_k = \sum_{i=1}^{k} \lambda_i a_i,$$

where $\lambda_i$ are all scalar quantities.

**Linear dependence of vectors:** A set of $k$ $n$-component vectors $a_1, a_2, \ldots, a_k$ is said to be linearly dependent if there exist scalars $\lambda_i (i = 1, 2, \ldots, k)$ with at least one $\lambda_i (\neq 0)$ such that

$$\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_k a_k = O$$

is satisfied, where O is an $n$-component null vector.

**Linearly independent vectors:** A set of $k$ $n$-component vectors $(k \leq n) a_1, a_2, \ldots, a_k$ is said to be linearly independent if the linear combination of all the vectors is equal to an $n$-component null vector only, when all the scalars are equal to zero. Mathematically, the set of vectors $a_1, a_2, \ldots, a_k$ is said to be linearly independent if the condition

$$\sum_{i=1}^{k} \lambda_i a_i = O$$

is satisfied only when all the scalar quantities $\lambda_i$ are zero.

**Spanning set:** A set of $k$ $n$-component vectors $a_1, a_2, \ldots, a_k$ in a vector space is said to be a spanning set or a generating set if every vector in the space can be expressed as a linear combination of the vectors $a_1, a_2, \ldots, a_k$.

**Basis set:** A spanning set of vectors, all of which are linearly independent, is known as a basis set. The set of all $n$ $n$-component unit vectors $e_1$, $e_2$, …, $e_n$ is said to form a basis set in $n$-dimensional Euclidean space, and this basis set is known as a standard basis set.

**Properties:**

(i) A set of vectors is either linearly independent or dependent.
(ii) A set of vectors with a null vector is always linearly dependent.
(iii) A non-empty subset of a linearly independent set of vectors is always linearly independent, and a superset of a linearly dependent set of vectors is always linearly dependent.

**Replacement theorem:**

If $a_1, a_2, \ldots, a_n$ is any basis set in $R^n$ and $b$ is any often non-null vector in $R^n$ such that

$$b = \lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_r a_r, \text{ where } \lambda_i \neq 0 \text{ for some } i,$$

then the vector $a_k$ (for which $\lambda_k \neq 0$) can be replaced by the vector $b$. The new set of $n$ vectors $a_1, a_2, \ldots, a_{k-1}, b, a_{k+1}, \ldots, a_n$ is also a basis set in $R^n$.

*Example 1* Show that the vectors (4, 5), (6, 2), (8, 10) are linearly dependent.

**Solution** Let $a_1$ = (4, 5), $a_2$ = (6, 2), $a_3$ = (8, 10).

Let $\lambda_1, \lambda_2, \lambda_3$ be three scalars such that

$$\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 = 0. \tag{18.2}$$

This implies that $\lambda_1(4,5) + \lambda_2(6,2) + \lambda_3(8,10) = (0,0)$.

Equating the corresponding components, we have

$$4\lambda_1 + 6\lambda_2 + 8\lambda_3 = 0$$
$$5\lambda_1 + 2\lambda_2 + 10\lambda_3 = 0$$

From the above, we have

$$\frac{\lambda_1}{60-16} = \frac{\lambda_2}{40-40} = \frac{\lambda_3}{8-30} = k = 1\,(\text{say}).$$

Therefore, $\lambda_1 = 44,\ \lambda_2 = 0,\ \lambda_3 = 22,$

i.e. Eq. (18.2) is satisfied if $\lambda_1 = 44,\ \lambda_2 = 0$ and $\lambda_3 = -22$, which are not all zero.

Hence, $44a_1 - 22a_3 = 0$, i.e. the given vectors are linearly dependent.

*Example 2* Show that the vectors (4, 3, 2), (2, 1, 4), (2, 3, −8) are linearly dependent.

**Solution** The matrix formed with the components of the given three vectors is given by

$$A = \begin{bmatrix} 4 & 3 & 2 \\ 2 & 1 & 4 \\ 2 & 3 & -8 \end{bmatrix}.\ \text{Now, } |A| = \begin{vmatrix} 4 & 3 & 2 \\ 2 & 1 & 4 \\ 2 & 3 & -8 \end{vmatrix} = 0.$$

Hence, the rank of the matrix $A$ is less than 3. Therefore, the given three vectors are linearly dependent.

**Alternative method:** Let $a_1 = (4,3,2),\ a_2 = (2,1,4),\ a_3 = (2,3,-8).$

Let $\lambda_1, \lambda_2, \lambda_3$ be three scalars such that

$$\begin{aligned} &\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 = 0 \\ \text{or}\quad &\lambda_1(4,3,2) + \lambda_2(2,1,4) + \lambda_3(2,3,-8) = (0,0,0). \end{aligned} \tag{18.3}$$

Now equating the corresponding components, we have

$$4\lambda_1 + 2\lambda_2 + 2\lambda_3 = 0 \tag{18.4}$$

$$3\lambda_1 + \lambda_2 + 3\lambda_3 = 0 \tag{18.5}$$

$$2\lambda_1 + 4\lambda_2 - 8\lambda_3 = 0. \tag{18.6}$$

From (18.4) and (18.5), we have

$$\begin{aligned} &\frac{\lambda_1}{6-2} = \frac{\lambda_2}{6-12} = \frac{\lambda_3}{4-6} \\ \text{or} \quad &\frac{\lambda_1}{4} = \frac{\lambda_2}{-6} = \frac{\lambda_3}{-2} = k = 1\,(\text{say})\,. \\ \text{or} \quad &\lambda_1 = 4,\ \lambda_2 = -6,\ \lambda_3 = -2 \end{aligned}$$

Substituting these values of $\lambda_1, \lambda_2, \lambda_3$ in (18.3), we have

$$4a_1 - 6a_2 - 2a_3 = 0,$$

which implies that the given three vectors are linearly dependent.

*Example 3* Show that the vectors (1, 2, 3), (2, 4, 1) and (3, 2, 9) are linearly independent.

**Solution** The matrix formed with the components of the given three vectors is given by

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \\ 3 & 2 & 9 \end{bmatrix}.\ \text{Now, } |A| = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \\ 3 & 2 & 9 \end{vmatrix} = -20 \neq 0.$$

Hence, the rank of the square matrix $A$ is 3. Therefore, the given three vectors are linearly independent.

**Alternative method:**

Let us assume that the given three vectors $a_1 = (1,2,3),\ a_2 = (2,4,1),\ a_3 = (3,2,9)$ are linearly dependent.

Then there exist three quantities $\lambda_1, \lambda_2, \lambda_3$, not all zero, such that

$$\begin{aligned} &\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 = 0, \\ \text{i.e.} \quad &\lambda_1(1,2,3) + \lambda_2(2,4,1) + \lambda_3(3,2,9) = (0,0,0). \end{aligned} \tag{18.7}$$

Now equating the corresponding components, we have

$$\lambda_1 + 2\lambda_2 + 3\lambda_3 = 0 \tag{18.8}$$

$$2\lambda_1 + 4\lambda_2 + 2\lambda_3 = 0 \tag{18.9}$$

$$3\lambda_1 + \lambda_2 + 9\lambda_3 = 0 \tag{18.10}$$

From (18.8) and (18.9), we have

$$\frac{\lambda_1}{4-12} = \frac{\lambda_2}{6-2} = \frac{\lambda_3}{4-4} = k = 1 \text{ (say)},$$
$$\text{i.e.}\quad \lambda_1 = -8,\ \lambda_2 = 4,\ \lambda_3 = 0.$$

With these values of $\lambda_1, \lambda_2, \lambda_3$, Eq. (18.10) will not be satisfied. Then Eq. (18.7) will not be satisfied; therefore, Eq. (18.7) will not be satisfied with at least one of $\lambda_1, \lambda_2, \lambda_3$ as non-zero. Equation (18.7) will be satisfied only for $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Hence, the given three vectors are linearly independent.

*Example 4* Show that the vectors (2, −1, 0), (3, 5, 1), (1, 1, 2) form a basis for $E^3$.

**Solution** The matrix formed with the components of these three vectors is given by

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 3 & 5 & 1 \\ 1 & 1 & 2 \end{bmatrix}. \text{ Now, } |A| = \begin{vmatrix} 2 & -1 & 0 \\ 3 & 5 & 1 \\ 1 & 1 & 2 \end{vmatrix} = 23 \neq 0.$$

Thus, the matrix formed by the elements of the given vectors is non-singular and its rank is 3.

Hence, the given vectors are linearly independent.

If the given vectors span $E^3$, then for any vector $(a_1, a_2, a_3)$ in $E^3$, there exist $\lambda_1, \lambda_2, \lambda_3$ such that

$$(a_1, a_2, a_3) = \lambda_1(2, -1, 0) + \lambda_2(3, 5, 1) + \lambda_3(1, 1, 2). \tag{18.11}$$

Now equating the corresponding components, we have

$$\left.\begin{aligned} 2\lambda_1 + 3\lambda_2 + \lambda_3 &= a_1 \\ -\lambda_1 + 5\lambda_2 + \lambda_3 &= a_2 \\ \lambda_2 + 2\lambda_3 &= a_3 \end{aligned}\right\}. \tag{18.12}$$

The coefficient matrix of these equations is given by

$$\begin{bmatrix} 2 & 3 & 1 \\ -1 & 5 & 1 \\ 0 & 1 & 2 \end{bmatrix} = A^{\mathrm{T}}. \text{ Now, } \left|A^{\mathrm{T}}\right| = |A| = 23 \neq 0.$$

Hence, there exists a unique solution for $\lambda_1, \lambda_2, \lambda_3$ whatever may be the value of $a_1$, $a_2$, $a_3$, and the vector $(a_1, a_2, a_3)$ can be expressed as the linear combination of the vectors (2, −1, 0), (3, 5, 1) and (1, 1, 2). Hence, the given three vectors form a basis for $E^3$.

*Example 5* Find a basis for $E^3$ that contains the vectors (1, 2, 2) and (2, 0, 1).

**Solution** We know that the standard basis for $E^3$ is $\{e_1, e_2, e_3\}$, where $e_1$ = (1, 0, 0), $e_2$ = (0, 1, 0) and $e_3$ = (0, 0, 1).

Let $a_1 = (1,2,2)$ and $a_2 = (2,0,1)$.

Now $a_1 = (1,2,2) = e_1 + 2e_2 + 2e_3$.

Hence, $a_1$ can replace any one vector from $\{e_1, e_2, e_3\}$ to form a new basis. Let it be $\{a_1, e_2, e_3\}$.

$$\begin{aligned} &\text{Again, let}\, a_2 = \lambda_1 a_1 + \lambda_2 e_2 + \lambda_3 e_3, \\ &\quad \text{i.e.} \quad (2,0,1) = \lambda_1(1,2,2) + \lambda_2(0,1,0) + \lambda_3(0,0,1). \end{aligned} \tag{18.13}$$

Equating the corresponding components, we have

$$\lambda_1 = 2,\ 2\lambda_1 + \lambda_2 = 0,\ 2\lambda_1 + \lambda_3 = 1.$$

Solving, we have $\lambda_1 = 2,\ \lambda_2 = -4,\ \lambda_3 = -3.$

Hence, from (18.1), we have

$$a_2 = 2a_1 - 4e_2 - 3e_3.$$

Therefore, $a_2$ can replace $e_2$ or $e_3$ from $\{a_1, e_2, e_3\}$ to form a new basis. Hence, the new basis is $\{a_1,\ a_2,\ e_3\}$.

## 18.7 Probability

**Random experiment:** An experiment performed without a definite prediction about the result is known as a random experiment.

**Example:** Whenever a coin is tossed or a die is thrown, it cannot be predicted whether a 'head' or a 'tail' will appear or which face of the die will come up on top. These are examples of random experiments.

**Trial and event:** If multiple experiments are performed under essentially with the same physical condition, then each experiment is called a trial, and the results (outcomes) are known as events.

**Example:** If a coin is tossed several times, then each tossing is a trial and the results ('heads' or 'tails') are the events.

Events are classified as (i) elementary events and (ii) composite events.

Elementary events are those which are single in nature. 'head' and 'tail' are examples of elementary events in the case of tossing a coin. Similarly, in the case of throwing a die, 'face 3' and 'face 4' are examples of elementary events.

On the other hand, composite events are those which are formed through the combination of two or more elementary events. The event 'even-numbered face' is an example of a composite event in the case of throwing a die.

**Mutually exclusive cases (or events):** If two or more events do not happen simultaneously then these events are said to be mutually exclusive.

**Equally likely cases (or events):** Two events are said to be equally likely if any one of them cannot be predicted over the other during a trial.

**Example:** In the case of tossing an unbiased coin, both the events 'head' and 'tail' are equally likely.

**Exhaustive cases (or events):** The number of all possible cases that may happen during a trial are all together said to form an exhaustive case.

**Example:** In the case of throwing a die, there are six possibilities: 'face one', 'face two', 'face three', 'face four', 'face five', 'face six'. So there are six exhaustive cases.

**Favourable cases (or events):** The cases which favour the occurrence of a particular event are said to be favourable to that event.

**Example:** In the case of throwing a die, the event 'multiple of 3' covers the two cases 'face 3' and 'face 6'. So these cases are favourable to the event.

**Classical definition of probability:**

If there are $n$ exhaustive, pairwise mutually exclusive, and of these equally likely outcomes of a random experiment, and if $m$ out of these are favourable to a particular event '$A$', then the probability of the happening of $A$ is denoted by $P(A)$ and is defined by

$$P(A) = \frac{m}{n}.$$

**Properties:**

(i) An event which is bound to happen is known as a certain event. The probability of a certain event is unity, i.e. $P(S) = 1$, where $S$ is the certain event.

(ii) An event which will never happen is known as an impossible event. The probability of an impossible event is zero, i.e. $P(X) = 0$, where $X$ is the impossible event.
(iii) For any event $A$, $P(A) \geq 0$.
(iv) The probability of an event lies between 0 and 1, i.e. $0 \leq P(A) \leq 1$.
(v) $P(\overline{A}) = 1 - P(A)$, where $\overline{A}$ is the complementary event of $A$.

**Sample space or event space:**
The set of all possible outcomes of a random experiment is called the sample space associated with the experiment. The possible outcomes (elements of the sample space) are called sample points.

*Example* For the random experiment of tossing a coin, there are two outcomes: (i) head, (ii) tail.

If we denote the occurrence of a head by $H$, the occurrence of a tail by $T$ and the sample space by $S$, then

$$S = \{H, T\}.$$

If the random experiment consists of tossing a coin twice, the associated sample space is

$$S = \{HH, HT, TH, TT\},$$

where $HT$ denotes the occurrence of $H$ from the first toss and the occurrence of $T$ from the second one.

**Discrete and continuous sample spaces:**
A sample space that consists of a finite (or an infinite but countable) number of sample points is called a discrete sample space. A sample space which is not discrete is called a continuous sample space.

**Axiomatic definition of probability**
Let $E$ be a random experiment described by the event space $S$, and let $A$ be any event connected with $E$. The probability of $A$ is a number associated with $A$ to be denoted by $P(A)$ such that the following axioms are satisfied:

(i) $P(A) \geq 0$.
(ii) The probability of a certain event $P(S) = 1$.
(iii) For any number of pairwise mutually exclusive events $A_1$, $A_2$, $A_3$, …
$P(A_1 + A_2 + A_3 + \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$

**Conditional probability**
For two events $A$ and $B$, if $A$ occurs when $B$ has already occurred, then the probability of $A$ with respect to $B$ is called the conditional probability of $A$ with respect to $B$. It is denoted by $P(A/B)$ and defined by

$$P\left(\frac{A}{B}\right) = \frac{P(AB)}{P(B)}, \text{ provided } P(B) \neq 0.$$

**Properties:**

(i) $P(AB) = P(A)\ P(B/A) = P(B)\ P(A/B)$
(ii) $P(ABC) = P(A)\ P(B/A)\ P(C/AB)$ for the events $A$, $B$, $C$.

**Independent events**

Two events $A$ and $B$ are said to be independent if the occurrence of one remains unaffected by the occurrence of the other.

$$\text{Hence, } P(A/B) = P(A), \quad \text{i.e. } \tfrac{P(AB)}{P(B)} = P(A),$$
$$\text{i.e. } P(AB) = P(A)P(B).$$

This result can be generalized in the following form:

$$P(A_1A_2 \cdots A_n) = P(A_1)P(A_2) \cdots P(A_n).$$

**Random variable**

If $S$ is the sample space associated with a random experiment, then a function $X : S \to R$ which assigns one and only one real number for each element $x \in S$ is called a random variable.

*Example* The sample space associated with the random experiment of tossing a coin twice is

$$S = \{HH, HT, TH, TT\}.$$

If a random variable is defined by

$$X = \text{number of heads obtained},$$

then we have

$$X(HH) = 2,\ X(HT) = 1,\ X(TH) = 1,\ X(TT) = 0.$$

**Discrete random variable:**

A random variable $X$ is said to be discrete if $X$ possesses values which are either finite or countably infinite.

**Discrete distribution**

If the random variable $X$ takes a discrete set of values $\ldots x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots,$ where $\ldots x_{-2} < x_{-1} < x_0 < x_1 < x_2, \ldots$ with probabilities

$$P(X = x_i) = f_i \quad (i = 0, \pm 1, \pm 2, \ldots),$$

then the distribution function is obtained as follows.

In $x_i \leq x < x_{i+1}$

$$\begin{aligned} F(x) &= P(-\infty \leq X \leq x) = P\left\{\sum_{k=-\infty}^{i} (X = x_k)\right\} \\ &= \sum_{k=-\infty}^{i} P(X = x_k) = \sum_{k=-\infty}^{i} f_k (i = 0, \pm 1, \pm 2, \ldots). \end{aligned}$$

Here $f_i$ is called the probability mass function of $X$.

Properties of the probability mass function are:

(i) $f_i \geq 0 \ \forall i$

(ii) $\sum_{i=-\infty}^{\infty} f_i = 1$

(iii) $P(a < X \leq b) = \sum_{a < x_i \leq b} f_i.$

**Some important discrete distributions**

(i) **Binomial distribution**

A discrete random variable $X$ having the set $\{0, 1, 2, \ldots, n\}$ as the spectrum is said to have a binomial distribution with parameters $n, p$ if the probability mass function of $X$ is given by

$$f_i = {}^n c_i p^i q^{n-i}, \text{for } i = 0, 1, 2, \ldots, n,$$

where $n$ is a positive integer and $0 < p < 1, q = 1 - p$.

(ii) **Poisson distribution**

A discrete random variable $X$ having the set $\{0, 1, 2, \ldots\}$ as the spectrum is said to have a Poisson distribution with parameter $\mu(>0)$ if the probability mass function of $X$ is given by

$$f_i = \mathrm{e}^{-\mu} \frac{\mu^i}{i!}, \text{ for } i = 0, 1, 2, \ldots$$

**Continuous random variable**

A random variable $X$ is said to be continuous if $X$ possesses values which are uncountably infinite.

**Continuous distribution**

The distribution of a random variable $X$ is said to be continuous if the distribution function $F(x)$ is continuous and its derivative $F'(x)$ is piecewise continuous everywhere.

The derivative of the distribution function of $X$, i.e. $F'(x)$, is called the probability density function (PDF) of the random variable $X$ and is denoted by $f(x)$.

The necessary conditions for a probability density function are as follows:

(i) $f(x) \geq 0$ everywhere

(ii) $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1.$

**Properties:**

(i) $P(X = a) = F(a) - F(a - 0) = 0$ since $F(x)$ is continuous at any point $a$ ($a$ being constant)

(ii) $P(a < X \leq b) = F(b) - F(a) = \int_a^b f(x)\mathrm{d}x$

(iii) $F(x) = \int_{-\infty}^{x} f(u)\mathrm{d}u$

(iv) $$\begin{aligned} P(x < X \leq x + \mathrm{d}x) &= F(x + \mathrm{d}x) - F(x) \\ &= \mathrm{d}F(x) = F'(x)\mathrm{d}x = f(x)\mathrm{d}x. \end{aligned}$$

**Density curve:**

The curve $y = f(x)$ is called the density curve. It gives a useful graphical representation in the continuous case.

**Some important continuous distributions:**

(i) **Uniform or rectangular distribution**

A continuous random variable $X$ having the interval $(a, b)$ as the spectrum is said to have a uniform distribution if the probability density function of $X$ is given by

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise} \end{cases},$$

where $a$ and $b$ are two parameters of the distribution.

(ii) **Normal $(\mathbf{m}, \sigma)$ distribution**

A continuous random variable $X$ having the interval $(-\infty, \infty)$ as the spectrum is said to have a normal distribution if the probability density function of $X$ is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

It is a probability distribution with two parameters $m, \sigma$ and is denoted by $N(m, \sigma)$.

If $m = 0$ and $\sigma = 1$, then $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty.$

The corresponding distribution is called a standard normal distribution and is denoted by $N(0, 1)$.

(iii) **Gamma distribution**

A continuous random variable $X$ having $(0, \infty)$ as its spectrum is said to have a gamma distribution if the probability density function of $X$ is given by

$$f(x) = \begin{cases} \dfrac{e^{-x} x^{l-1}}{\Gamma(l)}, & 0 < x < \infty,\ l > 0 \\ 0, & \text{otherwise} \end{cases}$$

Here $l$ is the only parameter of the distribution.

(iv) **Beta distribution of the first kind**

A continuous random variable $X$ having $(0, 1)$ as its spectrum is said to have a beta distribution of the first kind if the probability density function of $X$ is given by

$$f(x) = \begin{cases} \dfrac{x^{m-1}(1-x)^{n-1}}{\beta(m, n)}, & 0 < x < 1, \quad m, n > 0 \\ 0, & \text{otherwise} \end{cases}.$$

(v) **Beta distribution of the second kind**

A continuous random variable $X$ having $(0, \infty)$ as its spectrum is said to have a beta distribution of the second kind if the probability density function of $X$ is given by

$$f(x) = \begin{cases} \dfrac{x^{m-1}}{\beta(m, n)(1+x)^{m+n}}, & 0 < x < \infty, \quad m, n > 0 \\ 0, & \text{otherwise} \end{cases}.$$

(vi) **Cauchy distribution**

A continuous random variable $X$ having $(-\infty, \infty)$ as its spectrum is said to have a Cauchy distribution if its probability density function is given by

$$f(x) = \begin{cases} \dfrac{1}{\pi} \dfrac{1}{\lambda^2 + (x-\mu)^2} \end{cases}, -\infty < x < \infty \text{ and } \lambda, \mu > 0.$$

(vii) **Exponential distribution**

A continuous random variable $X$ having $(0, \infty)$ as its spectrum is said to have an exponential distribution if the probability density function of $X$ is given by

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, \; x, & \theta > 0 \\ 0, & \text{otherwise} \end{cases},$$

where $\theta(>0)$ is the only parameter of the distribution.

**Transformation of continuous random variables**

Let $X$ be a continuous random variable, and let $f_x(x)$ be the corresponding probability density function. Also, let $y = g(x)$ be a continuous differentiable function for all $x$. If $f_y(y)$ is the probability density function of the random variable $Y$ given by $Y = g(X)$, and if $\frac{\mathrm{dy}}{\mathrm{dx}}$ is either positive or negative for all $x$, then

$$f_y(y) = f_x(x) \left| \frac{\mathrm{d}x}{\mathrm{d}y} \right|.$$

**Two-dimensional distribution function**

Let $X$ and $Y$ be two random variables defined on the same sample space $S$. The joint distribution function $F(x, y)$ of $X$ and $Y$ is defined by

$$F(x, y) = P(-\infty < X \leq x, \; -\infty < Y \leq y),$$

where the event $(-\infty < X \leq x, \; -\infty < Y \leq y)$ means the joint occurrence of two events $-\infty < X \leq x$ and $-\infty < Y \leq y$, i.e.

$$(-\infty < X \leq x, \quad -\infty < Y \leq y) = (-\infty < X \leq x)(-\infty < Y \leq y).$$

**Properties:**

(i) $P(a < X \leq b, \; -\infty < Y \leq c) = F(b, c) - F(a, c)$

(ii) $P(a < X \leq b, \; c < Y \leq d) = F(b, d) + F(a, c) - F(a, d) - F(b, c)$

(iii) $F(-\infty, y) = 0, \;\; F(x, -\infty) = 0, \;\; F(\infty, \infty) = 1.$

**Marginal distribution**

The individual distributions $F_x(x)$ and $F_y(y)$ of the random variables $X$ and $Y$ respectively are called the marginal distributions. They can easily be calculated from the joint distribution function $F(x, y)$ of $X$ and $Y$:

$$F_x(x) = F(x, \infty), \; F_y(y) = F(\infty, y).$$

**Independent random variables:**

If the random variables $X$ and $Y$ are independent, then $F(x, y) = F_x(x)F_y(y)$.

**Discrete distribution**

The two-dimensional distribution of random variables $X$, $Y$ will be called discrete if the distribution function $F(x, y)$ is a step function in two dimensions having steps of height $f_{ij}(>0)$ at the point $(x_i, y_i)(i,j = 0, \pm 1, \pm 2, \ldots)$, i.e.

$$F(x, y) = \sum_{l=-\infty}^{i} \sum_{k=-\infty}^{j} f_{kl} \text{ for } x_i \leq x < x_{i+1}, \; y_j \leq y < y_{j+1} (i,j = 0, \pm 1, \pm 2, \ldots)$$

This function will satisfy all the necessary conditions for a distribution function

$$\sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} f_{ij} = 1.$$

The marginal distribution of $X$ is given by

$$F_x(x) = \sum_{l=-\infty}^{\infty} \sum_{k=-\infty}^{i} f_{kl} = \sum_{k=-\infty}^{i} f_{k.},$$

where $f_{i.} = \sum_{j=-\infty}^{\infty} f_{ij}$.

Similarly, the marginal distribution of $Y$ is given by

$$F_y(y) = \sum_{l=-\infty}^{j} f_{.l},$$

where $f_{.j} = \sum_{i=-\infty}^{\infty} f_{ij}$.

**Continuous distribution:**

The joint distribution of two random variables $X$ and $Y$ is defined to be continuous if their joint distribution function $F(x, y)$ is continuous everywhere, and its first- and second-order partial derivatives are piecewise continuous everywhere.

The joint probability density function of $X$ and $Y$ is denoted by $f(x, y)$ and is defined by

$$P(a<X\leq b,\ c<X\leq b) = \int_c^d \int_a^b f(x,y)\mathrm{d}x\mathrm{d}y,$$

where $f(x,y) = \frac{\partial^2 F}{\partial x \partial y}$.

The necessary conditions for $f(x, y)$ to be a possible joint probability density function are given as follows:

(i) $f(x,y) \geq 0$ for all $x$, $y$

(ii) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\mathrm{d}x\mathrm{d}y = 1.$

The marginal distribution of $X$ is given by

$$F_x(x) = F(x,\infty) = \int_{-\infty}^{\infty} \int_{-\infty}^{x} f(x,y)\mathrm{d}x\mathrm{d}y$$

with the marginal density function as

$$f_x(x) = F_x'(x) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}y.$$

Similarly, the marginal distribution function of $Y$ is given by

$$F_y(y) = F(\infty,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{y} f(x,y)\mathrm{d}x\mathrm{d}y$$

with the marginal density function as

$$f_y(y) = F_y'(y) = \int_{-\infty}^{\infty} F(x,y)\mathrm{d}x.$$

The necessary and sufficient condition for the independence of two continuous random variables $X$ and $Y$ is

$$f(x, y) = f_x(x) f_y(y).$$

**Mathematical expectation**

The mathematical expectation or the mean value of the function $g(X)$ of the random variable $X$ is denoted by $E\{g(X)\}$ and is defined by

$$\begin{aligned} E\{g(X)\} &= \sum_{i=-\infty}^{\infty} g(x_i) f_i \quad \text{for a discrete distribution} \\ &= \int_{-\infty}^{\infty} g(x) f(x) \mathrm{d}x \quad \text{for a continuous distribution,} \end{aligned}$$

provided the series or the improper integral converges absolutely.

**Properties:**

(i) $E(a) = a$, $a$ being a constant
(ii) $E\{ag(X)\} = aE\{g(X)\}$, $a$ being a constant
(iii) $E\{g_1(X) + g_2(X) + \cdots + g_n(X)\}$
$= E\{g_1(X)\} + E\{g_2(X)\} + \cdots + E\{g_n(X)\}$

**Mean**

The mean of $X$ or that of the corresponding distribution is defined to be $E(X)$ and is denoted by the symbol $m(X)$ or simply $m$, i.e. $m = E(X)$.

**Moments**

Let $k$ be a non-negative integer. The moment of order $k$ of the $k$th moment of $X$ about a fixed point $a$ is defined to be the mean value $E\left\{(X-a)^k\right\}$.

The $k$th moment about the origin is denoted by $\alpha_k$, i.e. $\alpha_k = E\left(X^k\right)$.

The $k$th moment about the mean is called the $k$th central moment $(\mu_k)$ and is defined by

$$\mu_k = E\left\{(X-m)^k\right\}.$$

**Variance**

The second central moment $\mu_2$ is called the variance, written as $\mathrm{Var}(X)$, i.e.

$$\mathrm{Var}(X) = \mu_2 = E\left\{(X-m)^2\right\}.$$

**Moment-generating Function:**

The moment-generating function of a random variable $X$ is a function of a real variable $t$ denoted by $\psi_x(t)$ or $\psi(t)$ and defined by

$$\psi(t) = E\{e^{tX}\}.$$

**Characteristic function**

The characteristic function of $X$ is denoted by $X(t)$ and is defined by

$$X(t) = E\left(e^{itX}\right) = E\{\cos(tX) + i\,\sin(tX)\}.$$

**Median**

The median is an important measure of locating which value denotes the point dividing the probability distribution into two halves. Mathematically, the median $\mu$ is defined by the equation $F(\mu) = \frac{1}{2}$.

**Mode**

In a continuous distribution, any point for which the density function $f(x)$ is maximum is called a mode of the distribution. If $f(x)$ is maximum at a single point, two points or many points, then the distribution is called unimodal, bimodal or multimodal respectively.

**Expectation for a bivariate distribution**

The expectation or the mean value of a function $g(X, Y)$ of $X$, $Y$ is defined by

$$\begin{aligned} E\{g(X,Y)\} &= \sum_{j=-\infty}^{\infty}\sum_{i=-\infty}^{\infty} g(x_i, y_i) f_{ij} \text{ for the discrete case} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y \text{ for the continuous case,} \end{aligned}$$

provided the series or integral is absolutely convergent.

For two independent random variables $X$ and $Y$,

$$E(X+Y) = E(X) + E(Y) \text{ and } E(XY) = E(X)E(Y).$$

### Moments

The moments (about the origin) of the joint distribution of $X$ and $Y$ are defined by

$$\alpha_{kl} = E\left(X^k Y^l\right),$$

where $k$ and $l$ are non-negative integers.

The central moments are given by

$$\mu_{kl} = E\left\{(X - m_x)^k \left(Y - m_y\right)^l\right\}.$$

The central moment $\mu_{11}$ is known as the covariance of $X$ and $Y$ and is denoted by Cov($X$, $Y$).

So $\mathrm{Cov}(X, Y) = E\left\{(X - m_x)\left(Y - m_y\right)\right\}$.

### $\chi^2$-distribution

The probability density function is defined by

$$\begin{aligned} f(\chi^2) &= \frac{e^{-\frac{1}{2}\chi^2}\left(\frac{1}{2}\chi^2\right)^{n/2-1}}{2\Gamma\left(\frac{1}{2}n\right)}, \quad && \chi^2 \geq 0 \\ &= 0, && \chi^2 < 0, \end{aligned}$$

where $n$, the only parameter of the distribution, is a positive integer and is called the number of degrees of freedom of the distribution. A $\chi^2$-distribution with $n$ degrees of freedom is referred to as a $\chi^2(n)$-distribution.

### *t*-distribution:

The probability density function $t$-distribution or Student's distribution is given by

$$f(t) = \frac{1}{\sqrt{n}B\left(\frac{1}{2}, \frac{1}{2}n\right)\left(1 + \frac{t^2}{n}\right)^{(n+1)/2}}, \quad -\infty < t < \infty,$$

where the parameter $n$ is a positive integer, called the number of degrees of freedom of the distribution.

### *F*-distribution:

The probability density function of the *F*-distribution is given by

$$\begin{aligned} f(F) &= \frac{m^{m/2} n^{n/2} F^{\frac{m}{2}-1}}{B\left(\frac{1}{2}m, \frac{1}{2}n\right)(mF+n)^{(m+n)/2}}, \quad && F > 0 \\ &= 0, && F < 0, \end{aligned}$$

where *m*, *n*, both positive integers, are the two parameters of the distribution. The random variable is called the *F*(*m*, *n*) variate.

### Weibull distribution

The probability density function of the *three-parameter Weibull distribution* is given by

$$f(t) = \alpha\beta(t-\gamma)^{\beta-1}\exp\{-\alpha(t-\gamma)^{\beta}\},$$

where $\alpha \geq 0, \beta \geq 0, \gamma \geq 0$.

For this distribution, $\alpha$ = scale parameter, $\beta$ = shape parameter, $\gamma$ = location parameter.

The probability density function of the *two-parameter Weibull distribution* is obtained by setting $\gamma = 0$ and is given by

$$f(t) = \alpha\beta t^{\beta-1}\exp(-\alpha t^{\beta}),$$

where $\alpha \geq 0, \beta \geq 0$.

The probability density function of the *one-parameter Weibull distribution* is obtained by setting $\gamma = 0$ and assuming $\beta = C =$ constant and is given by

$$f(t) = \alpha C t^{C-1}\exp(-\alpha t^{C}),$$

where $\alpha \geq 0$.

### Log-normal distribution

A log-normal distribution is a probability distribution of a random variable whose logarithm is normally distributed; i.e. if *X* is log-normally distributed, $Y = \log(X)$ is normally distributed. Hence, if Y is a random variable with a normal distribution, then $X = \exp(Y)$ is log-normally distributed.

The probability density function of a log-normal distribution is

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \ x > 0,$$

where $\mu$ and $\sigma$ are the mean and standard deviation of the random variable of the natural logarithm.

If $X$ is a log-normally distributed variable, its mean, variance and standard deviation are

$$\begin{aligned} E(X) &= e^{\mu + \frac{1}{2}\sigma^2}, \mathrm{Var}(X) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}, \mathrm{S.D.}(X) = \sqrt{\mathrm{Var}(X)} \\ &= e^{\mu + \frac{1}{2}\sigma^2}\sqrt{(e^{\sigma^2} - 1)}. \end{aligned}$$

If the values of the mean and variance are known, then the parameters $\mu$ and $\sigma$ can be obtained as follows:

$$\mu = \ln(E(X)) - \frac{1}{2}\ln\left\{1 + \frac{\mathrm{Var}(X)}{\{E(X)\}^2}\right\}$$

$$\text{and} \quad \sigma^2 = \ln\left\{1 + \frac{\mathrm{Var}(X)}{\{E(X)\}^2}\right\}.$$

**Erlang distribution**

The probability density function of an Erlang distribution is

$$f(x) = \frac{(k\mu)^k x^{k-1} e^{-k\mu x}}{(k-1)!}, \ x \geq 0,$$

where $\mu$ and $k(> 0)$ are the parameters of the distribution.

## 18.8 Markov Process

A Markov process (chain) is a stochastic process used to analyse decision-making problems in which the occurrence of a specific event is dependent on the occurrence of the event immediately prior to the current event. Basically, this process is used to identify:

(i) A specific state of the system being studied
(ii) The state transition relationship.

The occurrence of an event at a specified point in time (say, in the $n$th period) puts the system in a given state, say $S_n$. If, after the passage of one time unit, another event occurs (during the $(n+1)$th time period), the system has moved from state $S_n$ to state $S_{n+1}$. For example, $S_n$ may represent the number of persons in the queue at a railway window at time $t = n$. As the time changes to $t = n + 1$, the state of the queue also changes to $S_{n+1}$.

The probability of moving from one state to another or remaining in the same state in a single time period is called the transition probability. Also, since the probability of moving from one state to another depends on the probability of the preceding state, the transition probability is a conditional probability. The transition probability $p_{ij}$ is the probability that the system presently in state $E_i$ will be in state $E_j$ at some later step.

A finite Markov process must satisfy the following conditions:

(i) The process consists of a finite number of states.
(ii) The probability of moving from one state to another is dependent only on the immediately preceding state.
(iii) Transition probabilities are stationary.
(iv) The process has a set of initial probabilities which may be given or determined.

**State transition matrix:**

A state transition matrix is a rectangular array whose elements represent the transition probabilities for a Markov process.

Let $S_i$ represent the $i$th ($i = 1, 2, \ldots, m$) state of a stochastic process and $p_{ij}$ the transition probability of moving from state $S_i$ to state $S_j$ in one step.

Then a one-stage state transition matrix $P$ can be described as follows:

$$P = \begin{array}{c} \\ S_1 \\ S_2 \\ \vdots \\ S_m \end{array} \begin{pmatrix} S_1 & S_2 & \cdots & S_m \\ p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}.$$

In the transition matrix of the Markov process, $p_{ij} = 0$ when no transition occurs from state $i$ to state $j$, and $p_{ij} = 1$ when the system is moving from the $i$th state to the $j$th state only, at the next transition.

Each row of the transition matrix represents a one-step transition probability distribution over all states. This means that

$$p_{i1} + p_{i2} + \cdots + p_{im} = 1 \text{ for all } i$$

and $0 \le p_{ij} \le 1$.

**Transition diagram:**

In a transition diagram, the transition probabilities which can occur in any situation are shown (see Fig. 18.1).

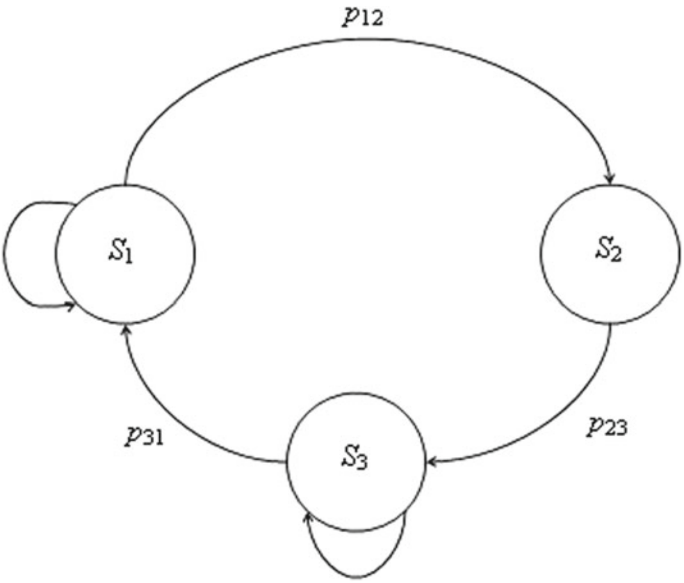



**Fig. 18.1** Transition diagram

The arrows from each state indicate the possible state to which a process can move from the given state. The transition matrix corresponding to Fig. 18.1 is as follows:

$$P = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{pmatrix} p_{11} & p_{12} & 0 \\ 0 & 0 & p_{23} \\ p_{31} & 0 & p_{33} \end{pmatrix} \end{array}.$$

***n*-step transition probability:**

Here we shall discuss the probabilistic behaviour of a system over a period of time. Let the system occupy state $S_i$ initially (i.e. at time 0). Then the probability that the system moves to state $S_j$ at time $n$ (i.e. after $n$ steps) is an $n$-step transition probability. It is denoted $p_{ij}^{(n)}$. We note that $p_{ij}^{(1)} = p_{ij}$ for all $i$ and $j$.

Let $p_i^{(n)}$ denote the probability that the system is in state $S_n$ at time $n$. Then the probability $p_j^{(n+1)}$ that the system will be in state $S_j$ at time $n + 1$ is related to $p_i^{(n)}$ by the following equation:

$$p_j^{(n+1)} = \sum_{i=1}^{n} p_i^{(n)} p_{ij} \quad n = 0, 1, 2, \ldots$$

where $p_{ij}$ is the transition probability of transitioning from $S_i$ to $S_j$ in one step.

For $j$ = 1, 2, …, $m$, the preceding relation gives a system of $m$ equations that can be written in matrix form as follows:

$$\begin{pmatrix} p_1^{(n+1)} \\ p_2^{(n+1)} \\ \vdots \\ p_m^{(n+1)} \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix} \begin{pmatrix} p_1^{(n)} \\ p_2^{(n)} \\ \vdots \\ p_m^{(n)} \end{pmatrix}$$

$$\text{or} \quad p_{(n+1)} = Ap_{(n)},$$

where $p_{(n+1)},\ p_{(n)}$ are the state probability vectors at time $n$ + 1 and $n$ respectively, and $A$ is a one-stage state transition matrix.

Therefore, if we know the initial state probabilities ($n$ = 0), we can easily compute the probabilities at any time successively as follows:

$$\begin{aligned} &p_{(1)} = Ap_{(0)} \\ &p_{(2)} = Ap_{(1)} = A^2p_{(0)} \\ &\cdots \\ &p_{(n)} = Ap_{(n-1)} = A^np_{(0)}. \end{aligned}$$

From these results, we can easily calculate the probability that the system occupies each of its states after $n$ steps by premultiplying the initial state probability vector by the $n$th power of the transition matrix.

**Steady state system:**

In any system, when it is observed that there is no change in the transition probabilities, the system would reach a point of equilibrium; i.e. no further changes would occur in the state probabilities. There is a limiting probability that the system in a finite (but large) number of transitions will reach steady state equilibrium. This means that when $n$ becomes very large, each $p_{ij}$ tends to a fixed limit, and each state probability vector $p_{(n)}$ approaches a constant value, i.e.

$$p_{(n+1)} = p_{(n)} = p, \text{ independent of } n.$$

Thus, in the limiting case,

$$\lim p_{(n+1)} = \lim Ap_{(n)}$$

becomes

$$p = Ap.$$

Therefore, when $n$ approaches infinity, $p_{(n)}$ becomes constant (i.e. independent) of time, and the system is said to be a steady state system.

# Bibliography


Ahuja, R., Magnati, T., & Orlin, J. (1993). *Network flows: Theory, algorithms and applications*. Upper Saddle River: Prentice Hall.

Aokie, M. (1971). *Introduction to optimization techniques: Fundamentals and applications of nonlinear programming*. New York: The Macmillan Company.

Barnett, S. (1975). *Introduction to mathematical control theory*. Oxford: Oxford University Press.

Bazaraa, M., Arvis, J. J., & Sherali, H. (1990). *Linear programming and network flows*. New York: Wiley.

Bazaraa, M. S., Sherali, H. D., & Shetty, C. M. (2004). *Nonlinear programming*. New York: Wiley.

Bazaraa, M. S., Sherali, H. D., & Shetty, C. M. (2006). *Nonlinear programming—Theory and algorithms*. New York: Wiley-Interscience.

Beale, E. M. I. (1959). On quadratic programming. *Naval Research Logistics Quarterly, 5,* 227–243.

Bellman, R. E., & Dreyfus, S. E. (1962). *Applied dynamic programming*. Princeton: Princeton University Press.

Bellmen, R. E., & Zadeh, L. A., (1970). Decision making in a fuzzy environment. *Management Science, 17*(4), 141–164.

Beveridge, G. S. G., & Schechter, R. S. (1970). *Optimization: Theory and practice*. New York: McGraw-Hill.

Bhatti, M. A. (2012). *Practical optimization methods*. Berlin: Springer-Verlag.

Bhunia, A. K. (2008). *Advanced operations research*. DDE study materials, MAS-305, 404. Burdwan: Burdwan University.

Bhunia, A. K. *Advanced optimization and operational research-II*. DDE study materials (Module 111-115). Midnapore: Vidyasagar University.

Bhunia, A. K. (2006–2007). *Operations research*. DDE study materials. Burdwan: Burdwan University.

Bradley, S., Hax, A., & Magnanti, T. (1977). *Applied mathematical programming*. Reading: Addison-Wesley.

Bronson, R., & Naadimuthu, G. (1982). Theory and problems of operations research. In *Schaum's outline series*. New York: McGraw-Hill.

Carlsson, C., & Korhonen, P. (1986). A parametric approach to fuzzy linear programming. *Fuzzy Sets Systems, 29,* 303–313.

Carter, M. W., & Price, C. C. (2001). *Operations research: A practical introduction*. Boca Raton: CRC Press.


© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2

Chakravorty, J. G., & Ghosh, P. R. (2009). *Linear programming and game theory*. Kolkata: Moulik Library.

Chong, E. K. P., & Zak, S. H. (2004). *An introduction to optimization*. New York: Wiley.

Churchman, C. W., Achoff, R. L., & Arnoff, E. L. (1957). *Introduction to operations research*. New York: Wiley.

Cooper, L., & Steinberg, D. (1970). *Introduction to methods of optimization*. Philadelphia: W.B. Saunders Co.

Dano, S. (1975). *Non-linear programming and dynamic programming*. Berlin: Springer-Verlag.

Denardo, E. (1982). *Dynamic programming theory and applications*. Upper Saddle River: Prentice Hall.

Dixan, L. C. W. (1972). *Nonlinear optimization*. London: The English Universities Press Ltd.

Eallenkotter, D. (1989). An early classic misplaced: Ford W. Harris's Economic Order Quantity model of 1915. *Management Science, 35*, 898–900.

Evans, J. R., & Minieka, E. (1992). *Optimization algorithms for networks and graphs*. Boca Raton: CRC Press.

Ferenc, F., Jenö, S., & Ferenc, S. (1999). *Introduction to the theory of games: Concepts, methods, applications*. Berlin: Springer-Verlag.

Fletcher, R. (Ed.). (1969). *Optimization*. London: Academic Press.

Forst, W., & Hoffmann, D. (2010). *Optimization—Theory and practice*. Berlin: Springer.

Gue, R., & Thomas, M. E. (1968). *Mathematical methods in operations research*. New York: Macmillan.

Gupta, R. K. (1992). *Operations research*. Agra: Krishna Prakasha Media (P) Ltd, Meerut, India.

Gupta, P. K., & Hira, D. S. (2004). *Operations research*. New Delhi: S. Chand & Co. Ltd.

Hadley, G. (1964). *Nonlinear and dynamic programming* (Vol. II). Reading: Addition-Wesley.

Hadley, G., & Whitin, T. M. (1963). *Analysis of inventory systems*. Englewood Cliffs: Prentice Hall.

Harris, F. (1915). *Operations and cost (Factory Management Series)*. Chicego, A.W. Shaw co.

Hillier, F. S., & Lieberman, G. J. (2001). *Introduction to operations research*. New York: McGraw-Hill.

Jahn, J. (2007). *Introduction to the theory of nonlinear optimization*. Berlin: Springer.

Jones, A. J. (2001). *Game theory: Mathematical models of conflict*. Chichester: Ellis Horwood Ltd.

Joshi, M. C., & Moudgalya, K. M. (2008). *Optimization theory and practice*. New Delhi: Narosa Publishing House.

Kambo, N. S. (2005). *Mathematical programming techniques*. New Delhi: Allied East-West Press Pvt. Ltd.

Levenson, M. E. (1967). *Maxima and minima*. New York: Macmillan.

Mangasarian, O. L. (1994). *Non-linear programming*. New York: McGraw-Hill.

McCormick, G. P. (1983). *Nonlinear programming: Theory, algorithms, and applications*. New York: Wiley.

McKinsey, J. C. C. (2003). *Introduction to the theory of games*. Mineola: Dover Publications.

Mesterton-Gibbons, M. (2001). *An introduction to game-theoretic modelling*. Providence: American Mathematical Society.

Mital, K. V., & Mohan, C. (2003). *Optimization methods in operations research and system analysis*. New Delhi: New Age International Publishers.

Mohan, C., & Deep, K. (2008). *Optimization techniques*. New Delhi: New Age International Publishers.

Mohan, C., & Deep, K. (2009). *Optimization techniques*. New Delhi: New Age Science.

Mustafi, C. K. (2002). *Operations research—Methods and practice*. New Delhi: New Age International Publishers Ltd.

Naddor, E. (1966). *Inventory systems*. New York: Wiley.

Nicolaj, N., & Ralph, P. B. (1994). *Foundations of game theory: Noncooperative games*. Berlin: Springer.

Osborne, M. J., & Rubinstein, A. (1994). *A course in game theory*. Cambridge: The MIT Press.
Panik, M. J. (1976). *Classical optimization: Foundations and extensions*. Amsterdam: North-Holland.
Peressini, A. L., Sullivan, F. E., & Uhl, J. J., Jr. (1988). *The mathematics of nonlinear programming*. Berlin: Springer-Verlag.
Rao, S. S. (1978). *Optimization—Theory and applications*. New Delhi: Wiley Eastern.
Rao, S. S. (2011). *Engineering optimization theory and practice*. New Delhi: New Age International (P) Ltd. Publishers.
Ravindran, A., Ragsdell, K. M., & Reklaitis, G. V. (1983). *Engineering optimization: Methods and applications*. New York: Wiley.
Reza, F. M. (1961). *An introduction to information theory*. New York: McGraw-Hill.
Roy, T. K. (1998). *Some inventory models in fuzzy environment* (Dissertation). Vidyasagar University.
Sanyal, D. C., & Das, K. (2008). *Linear programming and game theory*. Kolkata: U.N. Dhur & Sons Pvt. Ltd.
Sharma, J. K. (2007). *Operations research—Theory and applications*. New York: Macmillan.
Sharma, S. D. (2009). *Operations research*. Delhi: Kedarnath Ramnath.
Silver, E., & Peterson, R. (1985). *Decision systems for inventory management and production planning*. New York: Wiley.
Simmons, D. M. (1975). *Nonlinear programming for operations research*. Englewood Cliffs: Prentice Hall.
Swarup, K., Gupta, P. K., & Mohan, M. (2003). *Operations research*. New Delhi: Sultan Chand & Sons.
Taha, H. A. (1997). *Operations research—An introduction*. Upper Saddle River: Prentice-Hall.
Trappy, J. F. C., Liu, C. R., & Chang, T. C. (1988). Fuzzy non-linear programming: Theory and application in manufacturing. *International Journal of Production Research, 26,* 957–985.
Zangwill, W. I. (1969). *Nonlinear programming: A unified approach*. Englewood Cliffs: Prentice-Hall.

# Index



© Springer Nature Singapore Pte Ltd. 2019

A. K. Bhunia et al., *Advanced Optimization and Operations Research*,
Springer Optimization and Its Applications 153,
https://doi.org/10.1007/978-981-32-9967-2