Master of Science
PGMM -102N

Classical Optimization Techniques

U. P. Rajarshi Tandon
Open University

Course Design Committee

Course Preparation Committee

Dr. Ashish **Author (Unit 1 - 5)**
Assistant Professor, Govt. College Satnali, Haryana
Dr. Sharanjeet Dhawan **Author (Unit 6 - 7)**
Assistant Professor, CCS HAU COA Bawal, Haryana
Dr. Pushpendra Kumar **Author (Unit 8 - 14)**
Associate Professor, Department of Mathematics
SKD University, Hanumangarh, Rajasthan
Prof. A. K. Malik **Editor/Program Coordinator**
School of Sciences, UPRTOU, Prayagraj
Dr. Raghvendra Singh **Course Coordinator/Invited Member/Secretary**
Assistant Professor, Mathematics
School of Sciences, UPRTOU, Prayagraj

PGMM –102N: CLASSICAL OPTIMIZATION TECHNIQUES
ISBN-

Printed By: K.C. Printing & Allied Works, Panchwati, Mathura - 281003.

Syllabus

PGMM-102N/MAMM-102N: Classical Optimization Techniques

Block-1: Optimization Techniques

Unit-1: Introduction to Optimization Techniques

Introduction, Optimization techniques, applications of optimization techniques, optimization problems, classification of optimization problems.

Unit-2: Unconstrained Optimization Problem

Introduction, unconstrained optimization problem, single and multi-variable optimization problems.

Unit-3: Constrained Optimization Problem

Introduction, constrained optimization problem, constrained multi-variable optimization problem with equality and inequality constraints.

Block-2: Non-Linear Programming Problem

Unit-4: Non-Linear Programming-I

Introduction, unconstrained non-linear optimization problems, direct search method: Fibonacci method of search, Golden section method, Univariate method and Pattern search method, indirect search method: Steepest descent method.

Unit-5: Non-Linear Programming-II

Introduction, constraints non-linear optimization problem, direct methods: complex method and Zoutendijk method, indirect methods: transformation techniques and penalty function methods.

Unit-6: Quadratic Programming

Introduction, Kuhn-Tucker conditions, Quadratic programming, Wolfe's modified simplex method, Beale's method.

Unit-7: Separable Programming Problem

Introduction, Separable programming problem.

Block-3: Dynamic Programming Problem

Unit-8: Introduction to Dynamic Programming

Introduction, multi-decision process, Bellman's principle of optimality, dynamic programming algorithm.

Unit-9: Applications of Dynamic Programming

Introduction, solution of linear programming problem using dynamic programming and applications of dynamic programming problem.

Block-4: Advanced Optimization Techniques

Unit-10: Networking

Introduction, shortest route problem, minimum spanning tree problem and maximum flow problem.

Unit-11: Game Theory

Introduction, Game theory, lower and upper value of game, procedure to find saddle point, games without saddle point.

Unit-12: Goal Programming

Introduction, formulation of Goal programming, single goal models, goal programming algorithm and multi goal models.

Unit-13: Integer Programming Problem-I

Introduction, formulation of Integer programming problem, Gomory's cutting plane method and Branch and Bound Techniques.

Unit-14: Integer Programming Problem-II

Introduction, Branch and Bound Techniques.

Master of Science
PGMM -102N

Classical Optimization Techniques

U. P. Rajarshi Tandon Open University

Block

1 Optimization Techniques

Unit- 1

Introduction to Optimization Techniques

Unit- 2

Unconstrained Optimization Techniques

Unit- 3

Constrained Optimization Techniques

Block-1

Optimization Techniques

In this block, we will discuss into the historical development of Operations Research, explore key definitions, optimization techniques, and investigate their practical applications. The discussion will also encompass optimization problems, offering a thorough examination of their classification, illustrated with examples. Operations Research, as a branch of mathematics, is dedicated to applying scientific methods and techniques to address decision-making challenges, with a focus on achieving the best or optimal solutions.

In the first unit, we discussed the basic overview on operations research and optimization techniques. Applications and classifications of optimization techniques are also be discussed here in details. In the second unit, focus on classical optimization techniques. We will explore the necessary and sufficient conditions for achieving the optimum solution in both single and multivariable unconstrained optimization problems, supported by illustrative examples. These classical techniques prove highly effective in deriving optimal solutions for challenges that entail continuous and differentiable functions. Analytical in nature, these methods facilitate the determination of maximum and minimum points for both unconstrained and constrained continuous objective functions.

In the third unit, we will discuss into classical optimization techniques specifically tailored for constrained multivariable problems. Extensive attention will be given to exploring equality and inequality constraints, accompanied by illustrative examples. These classical optimization techniques play a crucial role in attaining optimal solutions for challenges characterized by continuous and differentiable functions. They involve analytical approaches for determining maximum and minimum points in both unconstrained and constrained continuous objective functions. When dealing with equality-constrained problems, we employ the Direct ubstitution method and Lagrange's Multiplier method.

UNIT-1: Introduction to Optimization Techniques

Structure

1.1 Introduction

In the present unit we shall discuss about the historical development of Operations Research, some important definitions of Operations Research, optimization techniques, application of optimization techniques. Optimization problem and classification of optimization problems are also discussed here in details with examples. Operations Research is a mathematical discipline focused on employing scientific methods and techniques to tackle decision-making problems, aiming to identify the best or optimal solutions. The roots of optimization techniques trace back to ancient times, where they were utilized by notable figures such as Newton, Cauchy, and Lagrange. Newton and Leibnitz made significant contributions to optimization methods within differential calculus. Cauchy, a pioneering mathematician, introduced the steepest descent method as the first application of unconstrained minimization problems. Lagrange developed a method for constrained problems, known as Lagrange's method of undetermined multipliers, involving the addition of unknown multipliers. In the years 1914-1915, Thomas Edison made an attempt to employ a tactical game board for minimizing shipping losses from enemy submarines, avoiding the risk to actual ships in wartime conditions. In this endeavor, he utilized a specific model and techniques of Operations Research. Since then, Operations Research has evolved into a crucial instrument in the organization and management of various institutions, offering valuable insights and methodologies for decision-making processes.

1.2 Objectives

After reading this unit the learner should be able to understand about:

- Optimization Techniques and its historical development
- Applications of Optimization Techniques

- Optimization Problems
- Classification of Optimization Problems

1.3 Optimization Techniques

Optimization techniques play a pivotal role across various disciplines, including electrical engineering, civil engineering, electronics engineering, mechanical engineering, telecommunication engineering, chemical engineering, biochemical engineering, automotive engineering, aerospace engineering, computer engineering, information technology, medical science, education, biotechnology, management, banking, manufacturing industries, and information technology industries. The fundamental objective of optimization is to attain the best possible output, which can be either the maximum or minimum value of a given criterion.

The study of optimization techniques is commonly undertaken within the realm of Operations Research, also known as mathematical programming techniques. These techniques are instrumental in determining the maximum or minimum of a function with several variables while adhering to a defined set of constraints. Overall, optimization techniques continue to be a cornerstone in problem-solving across diverse fields, contributing to advancements and improvements in various processes and systems.

The following are various mathematical programming techniques, along with other well-defined areas of Operations Research:

1. Mathematical Programming Techniques

(i) Linear Programming Problem

(ii) Non-linear Programming Problem

(iii) Dynamic Programming Problem

(iv) Integer Programming Problem

(v) Geometric Programming Problem

(vi) Multi-objective Programming Problem

(vii) Quadratic Programming Problem

(viii) Goal Programming Problem

(ix) Stochastic Programming Problem

(x) Separable Programming Problem

(xi) Information theory

(xii) Sequencing theory

(xiii) Game theory

(xiv) Assignment Methods

(xv) Transportation Methods

(xvi) Inventory Control Methods

(xvii) Network Scheduling Methods

(xviii) Differential Calculus Methods

(xix) Neural Networks

(xx) Fuzzy Logic

(xxi) Genetic Algorithms

(xxii) Simulated Annealing

(xxiii) Calculus of Variations

2. Stochastic Process Techniques

(i) Queuing Theory

(ii) Reliability Theory

(iii) Statistical Decision Theory

(iv) Renewal Theory

(v) Simulation Methods

(vi) Markov Process

3. Statistical Methods

(i) Correlation Analysis

(ii) Regression Analysis

(iii) Cluster Analysis

(iv) Factor Analysis

(v) Design of Experiments

(vi) Machine Learning

1.4 Applications of Optimization Techniques

The importance of application of optimization techniques in solving various problem in engineering field is universally accepted now. They are used in solving the problem in such fields as:

(i) Optimal the total inventory cost.

(ii) Optimal designing of control system.

(iii) Optimal planning, scheduling and controlling.

(iv) Optimal designing of chemical processing equipment and plant.

(v) Optimal selection of a new site for an industry.

(vi) Optimal designing of pipeline networks for process industries.

(vii) Optimal designing of computer structure for minimum cost.

(viii) Optimal designing of plastic structures.

(ix) Optimal planning to get the maximum profit in the presence of one or more competitor.

(x) Optimal designing of electrical networks.

(xi) Optimal designing of aircraft and aerospace structures, achieving a balance between structural efficiency, performance, and minimal weight.

(xii) Optimal designing of earthquake structures for minimum weight.

(xiii) Optimal controlling the waiting and idle times and queuing in production lines to reduce the costs.

(xiv) Optimal designing of civil engineering structures such as bridge, tower, dam, frames, chimney's etc. for minimum cost.

(xv) Optimal designing of material handling equipment such as conveyors, trucks and cranes for minimum cost.

(xvi) Optimal designing of water resources system for maximum benefit.

(xvii) Optimal designing of pumps heat transfer and turbines equipment for getting the maximum efficiency.

(xviii) Optimal designing of electrical machinery such as generators, transformers and motors.

(xix) Optimal designing of mechanical components such as gears, cams and machine tools etc.

1.5 Optimization Problems

An optimization problem is of the form

$$\text{Max or Min } Z = f(X) \qquad \ldots(1.1)$$

Subject to constraints (s.t.)

$g_j(X) \geq = \leq 0,\ j = 1, 2, 3, \ldots\ldots\ldots, m.$ …(1.2)

and $X \geq 0$

where $X = (x_1, x_2, x_3, \ldots\ldots\ldots, x_n)^T$ …(1.3)

In general, such problems are called mathematical programming problems. The function $f(X)$ in (1.1) which is to be maximized or minimized is known as the objective function. Conditions given in (1.2) are called constraints. Variables $x_1, x_2, x_3, \ldots\ldots, x_n$ are called decision (design) variables and conditions in (1.3) are called non-negativity restrictions. Thus any mathematical problem has three main parts:

(i) Objective function

(ii) Constraints

(iii) Non-negative restrictions.

1.6 Classification of Optimization Problems

The optimization problems can be classified in different ways are discussed below:

1.6.1 Classification Based on the Number of Objective Functions

(i) Single Objective Programming Problem.

An optimization problem involving only single objective function is known as single objective programming problem.

(ii) Multi-Objective Programming Problem.

An optimization problem involving two or more objective function is known as multi-objective programming problem.

1.6.2 Classification Based on the Existence of Constraints

(i) Constrained Optimization Problem.

An optimization problem involving constraints is known as constrained optimization problem. For example,

Max $Z = f(X)$

Subject to the constraints $g_j(X) \geq = \leq 0,\ j = 1, 2, 3, \ldots\ldots\ldots, m.$

and $X \geq 0$,

where $X = (x_1, x_2, x_3, \ldots\ldots\ldots, x_n)^T$

(ii) Unconstrained Optimization Problem.

An optimization problem without constraints is known as unconstrained optimization problem.

For example, Max Z = f(X) or Min Z = f(X).

1.6.3 Classification Based on the Nature of the Constraints Involved

(i) Linear Programming Problem.

An optimization problem involving objective function f(X) and all the constraints g_j(X) are in linear form is known as linear programming problem.

(ii) Non-Linear Programming Problem.

An optimization problem involving either objective function f(X) or at least one of the constraints g_j(X) are in non-linear form is known as non-linear programming problem.

(iii) Quadratic Programming Problem.

An optimization problem involving objective function f(X) is quadratic and the constraints g_j(X) are in linear form is known as quadratic programming problem.

(iv) Geometric Programming Problem.

An optimization problem involving objective function f(X) and constraints g_j(X) are expressed as polynomial is known as non-linear programming problem.

1.6.4 Classification Based on the Permissible Values of the Design

(i) Integer Programming Problem.

A linear programming problem involving some or all the design variables x_1, x_2, x_3, ………., x_n are restricted to take on only integer values is known as integer programming problem.

(ii) Real-Valued Programming Problem.

An optimization problem involving all the design variables x_1, x_2, x_3, ………., x_n are permitted to take any real value is known as non-linear programming problem.

1.6.5 Classification Based on the Physical Structure of the Problem

(i) Optimal Control Problem.

An optimization problem involving a number of stages in which each stage evolves from the preceding stage in a specific manner is called optimal control problem.

(ii) Non-Optimal Control Problem.

An optimization problem which is not optimal control problem is called non-optimal control problem.

1.6.6 Classification Based on the Nature of the Design Variables

(i) Static (Parameter) Optimization Problem.

If in the optimization problem we find the value of a set of design parameters, in which some prescribed function of these parameters is made that minimizes subject to certain constraint is known as static optimization problem.

(ii) Dynamic Optimization Problem.

An optimization problem involving each variable is a function of one or more parameters is known as dynamic optimization problem.

1.6.7 Classification Based on the Deterministic Nature of the Variables

(i) Deterministic Programming Problem.

An optimization problem involving deterministic design variables is known as deterministic programming problem.

(ii) Stochastic Programming Problem.

An optimization problem involving some or all the design variables are probabilistic is known as stochastic programming problem.

1.6.8 Classification Based on the Separability of the Functions

(i) Separable Programming Problem

An optimization problem involving objective function $f(X)$ and the constraints $g_j(X)$ are in separable form is known as separable programming problem.

(ii) Non-Separable Programming Problem

An optimization problem involving objective function $f(X)$ and the constraints $g_j(X)$ are in non-separable form is known as non-separable programming problem.

1.7 Summary

Optimization means obtaining the best output which may be maximum or minimum value of the criterion. Any mathematical problem has divided into three main parts:

(i) Objective function

(ii) Constraints and

(iii) Non-negative restrictions.

An optimization problem in which only single objective function is known as single objective programming problem.

An optimization problem in which two or more objective function is known as multi-objective programming problem.

An optimization problem involving constraints is known as constrained optimization problem. An optimization problem without constraints is known as unconstrained optimization problem.

An optimization problem involving objective function $f(X)$ and all the constraints $g_j(X)$ are in linear form is known as linear programming problem.

An optimization problem involving either objective function $f(X)$ or at least one of the constraints $g_j(X)$ are in non-linear form is known as non-linear programming problem.

An optimization problem involving deterministic design variables is known as deterministic programming problem.

An optimization problem involving some or all the design variables are probabilistic is known as stochastic programming problem.

An optimization problem involving objective function $f(X)$ and the constraints $g_j(X)$ are in separable form is known as separable programming problem.

An optimization problem involving objective function $f(X)$ and the constraints $g_j(X)$ are in non-separable form is known as non-separable programming problem.

1.8 Terminal Questions

1. What is optimization? Explain applications of optimization in engineering.
2. Write a short note on optimization techniques.
3. Given ten engineering applications of optimization techniques.
4. Write a short note on classification of optimization problems.

UNIT- 2: UNCONSTRINED OPTIMIZATION TECHNIQUES

Structure

2.1 Introduction

In this unit, we will delve into classical optimization techniques and explore the necessary and sufficient conditions for attaining the optimum solution in both unconstrained single and multivariable optimization problems, supplemented by illustrative examples. Classical optimization techniques prove highly valuable for deriving optimal solutions in scenarios featuring continuous and differentiable functions. These techniques, rooted in analytical methods, are adept at determining maximum and minimum points for unconstrained and constrained continuous objective functions. Throughout our discussion, we will leverage these analytical approaches to provide insights into achieving optimal solutions in various problem-solving contexts.

Classical optimization techniques refer to a set of traditional and well-established methods used to find optimal solutions to mathematical problems. These techniques are particularly useful for addressing optimization problems involving continuous and differentiable functions.

These classical optimization techniques offer a range of methods for solving unconstrained and constrained optimization problems. The choice of method depends on the characteristics of the problem, such as the nature of the objective function and the presence of constraints.

2.2 Objectives

After studying this unit, the learner will be able to understand:

- the unconstrained optimization problem
- the single variable unconstrained optimization problem
- procedure for solving single variable unconstrained optimization problem
- the multi-variable unconstrained optimization problem
- procedure for solving multi-variable unconstrained optimization problem

2.3 Unconstrained Optimization Problems

In classical optimization, sometimes we come across with the problems, to find the maxima or minima of the functions of one or more variables, with no restriction on the variable(s) involved in the problem. Such problems are known as unconstrained optimization problems.

There are two types of unconstrained optimization problems:

(i) Single variable optimization problem

(ii) Multi-variable optimization problem

2.4 Single Variable Optimization Problems

An optimization problem with single variable without any restriction is called a single variable unconstrained optimization problem. Consider $f(x)$ be a continuous function of the single variable x defined in the interval $[a, b]$.

Local Maxima

A function $f(x)$ with single variable is said to have a local (relative) maxima at $x = x_o$ if

$f(x_0) \geq f(x_0+h)$, for all sufficiently small positive and negative value of h.

Local Minima

A function $f(x)$ with single variable is said to have a local (relative) minima at $x = x_o$ if

$f(x_0) \leq f(x_0+h)$, for all sufficiently small positive and negative value of h.

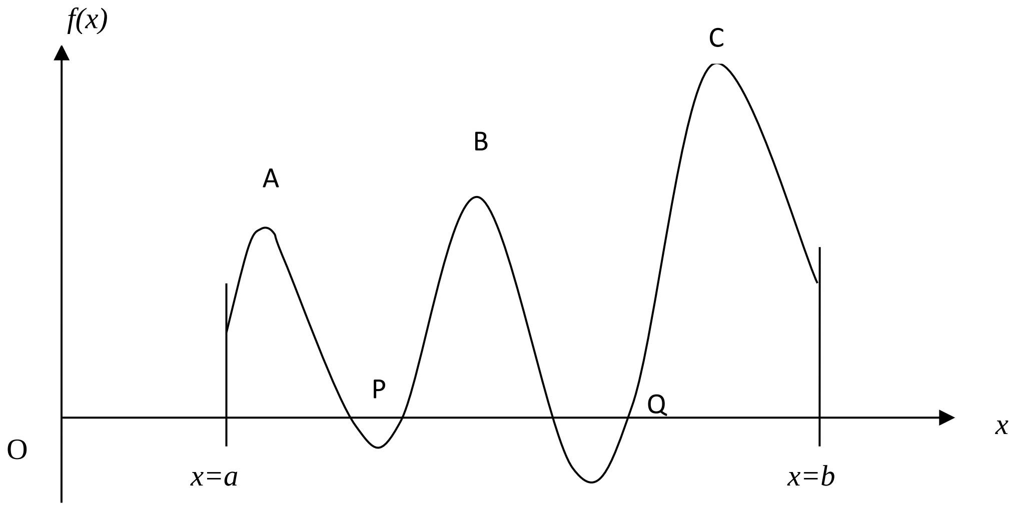
Global Maxima

A function $f(x)$ with single variable is said to have a global (absolute) maxima at $x = x_o$ in $[a, b]$ if $f(x_0) \geq$

$f(x)$, for all x defined in the interval $[a, b]$.

Global Minima

A function $f(x)$ with single variable is said to have a global (absolute) minima at $x = x_o$ in $[a, b]$ if $f(x_0) \leq f(x)$, for all x defined in the interval $[a, b]$.



Here points A, B and C are local maxima; P and Q are local minima; C global maxima and Q global minima.

2.4.1 Condition for Local Maxima or Minima of Single Variable Function

Following are the conditions for the existence of local maxima or minima of single variable function:

Necessary Condition

Let the function $f(x)$ be defined in the interval $[a, b]$ and $f'(x_0)$ exists, $a < x_o < b$. then $f(x)$ has a local maxima or local minima at $x = x_o$, if $f'(x_0) = 0$.

Sufficient Conditions

Let $f'(x_o) = f''(x_0) = f'''(x_0) = \ldots\ldots\ldots\ldots\ldots = f^{(n-1)}(x_0)$, but $f^{n}(x_0) \neq 0$ then $f(x_0)$ has

(i) $f(x)$ have minimum value if $f^{(n)}(x_0) > 0$ and n is even.

(ii) $f(x)$ have maximum value if $f^{(n)}(x_0) < 0$ and n is even.

(iii) Neither maxima and nor minima *i.e.*, point of inflexion if $f^{(n)}(x_0) \neq 0$ and n is odd.

2.4.2 Procedure to Find the Extreme Points of Function of single variable

Let us consider a single variable function

$$y = f(x) \quad \ldots (1)$$

Step I: Differentiate both sides of equation (1) with respect to x, we get

$$\frac{dy}{dx} = f'(x) \quad \ldots (2)$$

Step II: For extreme point, we have

$$\frac{dy}{dx} = 0$$

$$\Rightarrow \quad f'(x) = 0 \quad \ldots (3)$$

We get the values of x which is satisfied to $f'(x) = 0$.

Let us consider after solving equation (3) for x to get $x = x_o$ (say).

Step III: Differentiate again both sides of equation (2) with respect to x both side, we get

$$\frac{d^2y}{dx^2}.$$

Step IV: If $\left(\frac{d^2y}{dx^2}\right)_{x=x_0} > 0$ then $f(x)$ has minima at $x = x_o$.

Step V: If $\left(\frac{d^2y}{dx^2}\right)_{x=x_0} < 0$ then $f(x)$ has maxima at $x = x_o$.

Step VI: If $\left(\frac{d^2y}{dx^2}\right)_{x=x_0} = 0$ then find $\frac{d^3y}{dx^3}$ and we have $\left(\frac{d^3y}{dx^3}\right)_{x=x_0} \neq 0$ then there is neither a maximum nor minima at $x = x_o$ *i.e.*, $x = x_o$ is known as point of inflexion.

Step VII: If $\left(\frac{d^3y}{dx^3}\right)_{x=x_0} = 0$ then find $\frac{d^4y}{dx^4}$ and repeating steps IV, V and VI whenever you get the result.

Examples

Example.1. Determine the maximum and minimum value of the function

$$\boldsymbol{y = 3x^5 - 5x^3 + 1.}$$

Solution: It is given that $y = 3x^5 - 5x^3 + 1$ .... (4)

Differentiate both sides of equation (4) with respect to x, we get

$$\frac{dy}{dx} = 15x^4 - 15x^2 \quad \text{.... (5)}$$

For maxima and minima of y, we have

$$\frac{dy}{dx} = 0$$

$\Rightarrow$ $15x^4 - 15x^2 = 0$

$\Rightarrow$ $15x^2(x-1)(x+1) = 0$

$\Rightarrow$ $x = 0, 1, -1.$

Differentiate again both sides of equation (5) with respect to x, we get

$$\frac{d^2y}{dx^2} = 60x^3 - 30x$$

At $x = 0$, $\frac{d^2y}{dx^2} = 0$ *i.e.*, $x = 0$ is a point of inflexion.

At $x = 1$, $\frac{d^2y}{dx^2} = 30 > 0$ *i.e.*, y is minimum at $x = 1$.

At $x = -1$, $\frac{d^2y}{dx^2} = -30 < 0$ *i.e.*, y is maximum at $x = -1$.

Hence the given function y is minimum at $x = 1$ and maximum at $x = -1$.

Example.2. For a business organization, the following relationship for revenue function and cost functions. To find the level of output x at which profit is maximized, where *x* is measured in tons per week.

$$\mathbf{R}(x) = \mathbf{1000x - 2}\,x^2$$

and C(x) =x^3 – 59 x^2 + 1315 x + 1500.

Solution. The profit function is

$$P(x) = R(x) - C(x)$$

$$= 1000\,x - 2\,x^2 - x^3 + 59\,x^2 - 1315\,x - 1500$$

$$= -\,x^3 + 57\,x^2 - 315\,x - 1500 \qquad \text{…. (6)}$$

Differentiating both sides of equation (6) with respect to x, we get

$$\frac{dP}{dx} = -\,3x^2 + 114x - 315 \qquad \text{…. (7)}$$

For maxima and minima, we have

$$\frac{dP}{dx} = 0$$

$$\Rightarrow \qquad -3\,x^2 + 114\,x - 315 = 0$$

$$\Rightarrow \qquad x = 3,\ 35.$$

Differentiating again both sides of equation (7) with respect to x, we get

$$\frac{d^2P}{dx^2} = -\,6x + 114$$

At $x = 3$, $\frac{d^2P}{dx^2} = 96 > 0$, *i.e.*, the function P is minimum at $x = 3$.

At $x = 35$, $\frac{d^2P}{dx^2} = 96 < 0$, *i.e.* the function P is maximum at $x = 35$.

Hence, the profit is maximum at $x = 35$ tons per week.

Example.3. The speed of signaling (for submarine telegraph cable) varies as $x^2 \log\left(\frac{1}{x}\right)$ where x is denoted as the ratio of the radius of the cube for covering that. Prove that the greatest speed is reached when this ratio is 1:$\sqrt{e}$.

***Solution*:** Consider u is the speed of signaling, then

$$u \propto x^2 \log\left(\frac{1}{x}\right)$$

$$u = \lambda\, x^2 \log\left(\frac{1}{x}\right),\ \lambda \text{ is the constant of proportionally, } \lambda > 0,\ x \neq 0$$

$$u = -\lambda x^2 \log x \qquad \ldots (8)$$

Differentiate both sides of equation (8) with respect to *x*, we get

$$\frac{du}{dx} = -2\lambda x \log x - \lambda x^2(1/x)$$

$$= -\lambda[2x\log x + x] \qquad \ldots. (9)$$

For maxima and minima, we have

$$\frac{du}{dx} = 0$$

$\Rightarrow$ $\quad -\lambda[2x \log x + x] = 0$

$\Rightarrow$ $\quad \log x = -\frac{1}{2}$

$\Rightarrow$ $\quad x = e^{-1/2} = 1/\sqrt{e}$.

Differentiate again both sides of equation (9) with respect to *x*, we get

$$\frac{d^2u}{dx^2} = -\lambda\,[2x\,(1/x) + 2 \log x]$$

$$= -\lambda\,[2 \log x + 3]$$

At $x = 1/\sqrt{e}$, $\frac{d^2u}{dx^2} = -2\lambda$ (negative) *i.e.*, *u* is maximum.

Hence, *u* is maximum, ratio for $x = 1 : \sqrt{e}$.

Example.4. To show that the right circular cylinder of given surface area (including its ends) and maximum volume has a height equal to twice its radius.

Solution: We know that $\quad V = \pi r^2 h \qquad \ldots. (10)$

and $\quad S = 2\pi r h + 2\pi r^2 \quad$ (according to given)

$\Rightarrow$ $\quad 2\pi r h = 2k^2\pi - 2\pi r^2$

$$h = \frac{k^2 - r^2}{r} \qquad \text{.... (11)}$$

From equations (10) and (11), we have

$$V = \pi\, r\, (k^2 - r^2) \qquad \text{.... (12)}$$

Differentiate both side of equation (12) with respect to *r*, we get

$$\frac{dV}{dr} = \pi\left(k^2 - 3r^2\right) \qquad \text{... (13)}$$

For maxima and minima of V, we have

$$\frac{dV}{dr} = 0$$

$$\Rightarrow \qquad \pi\left(k^2 - 3r^2\right) = 0$$

$$\Rightarrow \qquad r = \frac{k}{\sqrt{3}} \qquad \text{.... (14)}$$

Differentiate both sides of equation (13) with respect to *r*, we get

$$\frac{d^2V}{dx^2} = -6\pi r \text{ (Negative)}$$

i.e., V is maximum.

Using equations (11) and (14), we have

$$h\, r = k^2 - r^2$$

$$\Rightarrow \qquad h = 2r.$$

Hence the right circular cylinder of given surface area (including its ends) and maximum volume has a height equal to twice its radius.

Example.5. A given quantity of metal is to be casted into a half cylinder (rectangular base and semi-circular ends). Prove that in order to have minimum surface area, the ratio of the height of the cylinder and the diameter of semi-circular ends is π:π+2.

Solution: Suppose the volume of the half cylinder is

$$V = \frac{1}{2}\pi\, r^2 h \qquad \text{... (15)}$$

Where r and h be the radius and height of the half cylinder respectively.

The surface area of rectangular base = $2\, r\, h$

Curved surface = $\pi\, r\, h$

Two semi-circular ends = $\pi\, r^2$

The total surface area $S = 2\, r\, h + \pi\, r\, h + \pi\, r^2$

$$= r\, h\,(2+\pi) + \pi\, r^2 \quad \dots (16)$$

From equation (15), we have

$$h = \frac{2V}{\pi r^2}$$

Then we have

$$S = \frac{r.2V}{\pi r^2}(2+\pi) + \pi\, r^2$$

$$= \frac{2V}{\pi r}(2+\pi) + \pi\, r^2 \quad \dots. (17)$$

Differentiate both side of (17) with respect to r, we get

$$\frac{dS}{dr} = 2\pi r - \frac{2V}{\pi r^2}(\pi + 2) \quad \dots. (18)$$

For maxima and minima of S, we have

$$\frac{dS}{dr} = 0$$

$$\Rightarrow \quad 2\pi r - \frac{2V}{\pi r^2}(\pi + 2) = 0$$

Using equation (15), we have

$$2\pi\, r - h\,(\pi+2) = 0$$

$$\Rightarrow \quad \frac{h}{2r} = \frac{\pi}{\pi + 2} \quad \dots (19)$$

Differentiate again both side of (18) with respect to r, we get

$$\frac{d^2S}{dr^2} = 2\pi - \frac{4V}{\pi r^3}(\pi + 2)$$

$= 6\,\pi$ (positive) i.e., S is maximum.

Hence the surface area S is minimum at $\frac{h}{2r} = \frac{\pi}{\pi + 2}$.

Check your Progress

1. Write a short note on unconstrained optimization problem.

2. Explain the procedure to find the extreme point for a single variable function.

3. A DC generator has integral resistance R ohms with an open circuit voltage of V volts. To determine the lead resistance *r* at which the power delivered by a DC generator is maximized.

4. The efficiency of a screw jack is defined as $\eta = \tan\alpha \cot(\alpha + \phi)$ where ϕ is a constant. Show that the efficiency is maximum at $\alpha = \frac{\pi}{4} - \frac{\phi}{2}$ and $\eta = \frac{1 - \sin\phi}{1 + \sin\phi}$.

2.5 Multivariable Optimization Problems

An optimization problem with two or more variable without any restriction is called a multi-variable unconstrained optimization problem.

2.5.1 Procedure to Find the Extreme Points of Functions of Two Variables

Let us consider $u = f(x_1, x_2)$ be a function of x_1 and x_2.

Step 1: Differentiate $f(x_1, x_2)$ partially with respect to x_1 and x_2, we get

$$\frac{\partial f}{\partial x_1} \text{ and } \frac{\partial f}{\partial x_2}$$

Step 2: For extreme points of *f*, we have

$$\frac{\partial f}{\partial x_1} = 0$$

and $$\frac{\partial f}{\partial x_2} = 0$$

Solving above these equations, we get some points as (a_1, b_1), (a_2, b_2), ……….etc.

Step 3: Differentiate again partially, we get

$$r = \frac{\partial^2 f}{\partial x_1^2},\ s = \frac{\partial^2 f}{\partial x_1 \partial x_2} \quad \text{and} \quad t = \frac{\partial^2 f}{\partial x_2^2}.$$

Step 4: At (a_1, b_1) determine the values of calculate r and $rt\text{-}s^2$

Case – I: If $rt\text{-}s^2 > 0$ and $r < 0$ then the given function $f(x_1, x_2)$ is maximum at the point (a_1, b_1).

Case – II: If $rt\text{-}s^2 > 0$ and $r > 0$ then the given function $f(x_1, x_2)$ is minimum at the point (a_1, b_1).

Case – III: If $rt\text{-}s^2 < 0$ then the given function $f(x_1, x_2)$ has saddle point at the point (a_1, b_1).

2.5.2 Procedure to Find Extreme Points of Function of *n*- Variables

Suppose $u = f(x_1, x_2, x_3, \ldots\ldots\ldots, x_n)$ is a function of $x_1, x_2, x_3, \ldots\ldots, x_n$.

Necessary Conditions:

$$\frac{\partial f}{\partial x_1} = 0,\ \frac{\partial f}{\partial x_2} = 0,\ \frac{\partial f}{\partial x_3} = 0, \ldots\ldots\ldots\ldots, \frac{\partial f}{\partial x_n} = 0$$

Sufficient Conditions:

Hessian Matrix at P for n variables will be

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Its leading minors are defined as

$$H_1 = \left| \frac{\partial^2 f}{\partial x_1^2} \right|,$$

$$H_2 = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{vmatrix},$$

$$H_3 = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{vmatrix},$$

Hence, the following cases will arise:

Case – I: If H_1, H_2, H_3, ………. are positive (*i.e.*, H is positive definite) then the function $f(x_1, x_2, x_3, \ldots\ldots, x_n)$ has minimum at P.

Case – II: If H_1, H_2, H_3, ………. are alternately negative, positive, negative (*i.e.*, H is negative definite) then the function $f(x_1, x_2, x_3, \ldots\ldots, x_n)$ has maximum at P.

Case – III: If H_1 and H_3, ………. are not of same sign and $H_2 = 0$ (*i.e.*, semi definite or indefinite) then the function $f(x_1, x_2, x_3, \ldots\ldots, x_n)$ has a saddle point at P.

Examples

Example.6 Show that the minimum value of $u = xy\left(\frac{a^3}{x}\right) + \left(\frac{a^3}{y}\right)$ ***is*** $3a^2$**.**

Solution: Given that

$$u = xy\left(\frac{a^3}{x}\right) + \left(\frac{a^3}{y}\right) \qquad \ldots (20)$$

Differentiate both side partially (20) with respect to x & y respectively, we have

$$\frac{\partial u}{\partial x} = y - \frac{a^3}{x^2} \qquad \text{… (21)}$$

and $$\frac{\partial u}{\partial y} = x - \frac{a^3}{y^2} \qquad \text{…. (22)}$$

For maxima minima of u, we have

$$\frac{\partial u}{\partial x} = 0 \quad \text{and} \quad \frac{\partial u}{\partial y} = 0$$

$\Rightarrow$ $$y - \frac{a^3}{x^2} = 0$$

and $$x - \frac{a^3}{y^2} = 0$$

Solving above, we get

$$x = a, y = a.$$

Differentiate again partially equation (21), we get

$$r = \frac{\partial^2 f}{\partial x_1^2} = \frac{2a^3}{x^3}$$

$$s = \frac{\partial^2 f}{\partial x_1 \partial x_2} = 1$$

and $$t = \frac{\partial^2 f}{\partial x_2^2} = \frac{2a^3}{y^3}$$

At (a, a), we get

$$r = 2, s = 1, t = 2.$$

Then we have $rt - s^2 = 3 > 0$

Since at (a, a), we have

$$rt - s^2 > 0 \text{ and } r > 0,$$

Therefore u is minimum at (a, a).

Hence the minimum value of the given function $u = a.\, a + \dfrac{a^3}{a} + \dfrac{a^3}{a} = 3\, a^2$.

Example.7. Determine the extreme points of the function

$$\boldsymbol{u = x^2 + 4y^2 + 4z^2 + 4xy + 4xz + 16yz.}$$

Solution: Given that

$$u = x^2 + 4y^2 + 4z^2 + 4xy + 4xz + 16yz \qquad \dots (23)$$

Differentiate both sides partially of (23) with respect to x, y and z respectively, we get

$$\frac{\partial u}{\partial x} = 2x+2y+2z$$

$$\frac{\partial u}{\partial y} = 8y+4x+16z$$

and
$$\frac{\partial u}{\partial z} = 8z+4x+16y$$

For extreme points of u, we have

$$\frac{\partial u}{\partial x} = 0, \quad \frac{\partial u}{\partial y} = 0 \text{ and } \frac{\partial u}{\partial z} = 0$$

$\Rightarrow$ $2x+2y+2z = 0$ or $2(x+y+z) = 0$

$\Rightarrow$ $8y+4x+16z = 0$ or $4(x+2y+4z) = 0$

and $8z+4x+16y = 0$ or $4(x+4y+2z) = 0$

Solving above three expression, we get

$$x = 0, y = 0, z = 0.$$

Now let the point P is (0, 0, 0).

Again differentiate partially, we get

$$\frac{\partial^2 u}{\partial x^2} = 2, \ \frac{\partial^2 u}{\partial y^2} = 8, \ \frac{\partial^2 u}{\partial z^2} = 8,$$

$$\frac{\partial^2 u}{\partial x \partial y} = 4, \ \frac{\partial^2 u}{\partial y \partial x} = 4, \ \frac{\partial^2 u}{\partial y \partial z} = 16,$$

$$\frac{\partial^2 u}{\partial z \partial y} = 16, \frac{\partial^2 u}{\partial x \partial z} = 4, \frac{\partial^2 u}{\partial z \partial x} = 4.$$

The Hessian matrix of the function *u* (*x*, *y*, *z*) is

$$H = \begin{bmatrix} 2 & 4 & 4 \\ 4 & 8 & 16 \\ 4 & 16 & 8 \end{bmatrix}$$

The leading minors of H are

$$H_1 = |2| = 2,$$

$$H_2 = \begin{vmatrix} 2 & 4 \\ 4 & 8 \end{vmatrix} = 0$$

and $$H_3 = \begin{vmatrix} 2 & 4 & 4 \\ 4 & 8 & 16 \\ 4 & 16 & 8 \end{vmatrix} < 0.$$

Here H_1 and H_3 are not same sign, and $H_2 = 0$ (*i.e.*, semi definite). Therefore the given function *u* has a saddle point at (0, 0, 0).

Example.8. Determine the extreme points $f(x_1, x_2) = 20x_1 + 26x_2 + 4x_1x_2 - 4x_1^2 - 3x_2^2$.

Solution: Given that

$$f(x_1, x_2) = 20x_1 + 26x_2 + 4x_1x_2 - 4x_1^2 - 3x_2^2 \qquad \text{…. (24)}$$

Differentiate both side partially (24) with respect to x_1 and x_2 respectively, we get

$$\frac{\partial f}{\partial x_1} = 20 + 4x_2 - 8x_1 \qquad \text{…. (25)}$$

and $$\frac{\partial f}{\partial x_2} = 26 + 4x_1 - 6x_2 \qquad \text{… (26)}$$

For extreme points of *f*, we have

$$\frac{\partial f}{\partial x_1} = 0$$

and $$\frac{\partial f}{\partial x_2} = 0$$

$\Rightarrow$ $$20 + 4x_2 - 8x_1 = 0$$

and $$26 + 4x_1 - 6x_2 = 0$$

Solving these, we get

$$x_1 = 7,\ x_2 = 9.$$

Differentiating again partially (25), we get

$$r = \frac{\partial^2 f}{\partial x_1^2} = -8,$$

$$s = \frac{\partial^2 f}{\partial x_1 \partial x_2} = 4$$

and $$t = \frac{\partial^2 f}{\partial x_2^2} = -6$$

Thus we have

$$rt\text{-}s^2 = (-8)(-6)-(4)^2 = 32.$$

At (7, 9), $rt-s^2 > 0$ and $r < 0$ *i.e.*, the given function f is maximum at point (7, 9).

Hence the given function f is maximum at (7, 9).

Example.9. Determine the extreme points of the function

$$f(x_1, x_2) = x_1^3 + 2x_2^3 + 3x_1^2 + 12x_2^2 + 24.$$ **And determine their nature also.**

Solution: Given that

$$f(x_1, x_2) = x_1^3 + 2x_2^3 + 3x_1^2 + 12x_2^2 + 24 \qquad \ldots (27)$$

Differentiate both sides partially (27) with respect to x_1 and x_2 respectively, we get

$$\frac{\partial f}{\partial x_1} = 3x_1^2 + 6x_1 \qquad \ldots. (28)$$

and $$\frac{\partial f}{\partial x_2} = 6x_2^2 + 24x_2 \qquad \ldots (29)$$

For extreme points, we have

$$\frac{\partial f}{\partial x_1} = 0$$

and $$\frac{\partial f}{\partial x_2} = 0$$

$\Rightarrow$ $$3x_1^2 + 6x_1 = 0$$

and $$6x_2^2 + 24x_2 = 0$$

Solving these, we get

$x_1 = 0, -2, x_2 = 0, -4.$

Differentiating again partially equations (28) and (29), we get

$$r = \frac{\partial^2 f}{\partial x_1^2} = 6(x_1 + 1)$$

$$s = \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0$$

and $$t = \frac{\partial^2 f}{\partial x_2^2} = 12(x_2 + 2)$$

At (0, 0), we have

$rt\text{-}s^2 = 72(x_1+1)(x_2+2) = 72 > 0$ and $r > 0$ *i.e.*, the given function f is minimum at point (0, 0).

At (0, -4), we have

$rt\text{-}s^2 = 72(x_1+1)(x_2+2) = -144 < 0 \Rightarrow$ no extreme point, *i.e.*, the given function f has a saddle point at (0, -4).

At (-2, 0), we have

$rt\text{-}s^2 = 72(x_1+1)(x_2+2) = -144 < 0 \Rightarrow$ no extreme point, *i.e.*, the given function f has a saddle point at (-2, 0).

At (-2, -4), we have

$rt-s^2 = 72(x_1+1)(x_2+2) = 144 > 0$ and $r < 0$ i.e., f is maximum at (-2, -4).

Hence the given function f has two saddle point at (0, -4) and (-2, 0) and f is minimum at (0, 0) maximum at (-2, -4).

Example.10. Determine the maximum and minimum value of *u*,

$$u = \sin x \sin y \sin (x+y).$$

Solution: Given that

$$u = \sin x \sin y \sin (x+y) \qquad \ldots (30)$$

Differentiate both side partially equation (30) with respect to x and y respectively, we get,

$$\frac{\partial u}{\partial x} = \sin y \left[\sin x \cos(x+y) + \cos x \sin(x+y)\right] \qquad \ldots (31)$$

and $$\frac{\partial u}{\partial y} = \sin x \left[\sin y \cos(x+y) + \cos y \sin(x+y)\right] \qquad \ldots. (32)$$

For maxima minima of u, we have

$$\frac{\partial u}{\partial x} = 0 \quad \text{and} \quad \frac{\partial u}{\partial y} = 0$$

$$\Rightarrow \qquad \sin y \left[\sin x \cos(x+y) + \cos x \sin(x+y)\right] = 0$$

and $$\sin x \left[\sin y \cos(x+y) + \cos y \sin(x+y)\right] = 0$$

Solving above two expression, we get

$$\tan (x+y) = -\tan x \qquad \ldots (33)$$

and $$\tan (x+y) = -\tan y \qquad \ldots (34)$$

Now, we have

$$\tan 2x = -\tan x = \tan (\pi - x)$$

$$2x = \pi - x$$

$$\Rightarrow \qquad x = \pi/3 = y$$

Also

$\sin y = 0 \quad \Rightarrow \quad y = 0$

and $\sin x = 0 \quad \Rightarrow \quad x = 0$

Thus the stationary points are (0, 0), (π/3, π /3).

Differentiate again partially equations (31) and (32), we get

$$r = \frac{\partial^2 u}{\partial x^2} = 2\sin y \cos(2x + y)$$

$$s = \frac{\partial^2 u}{\partial x \partial y} = \sin 2(x + y)$$

and $$t = \frac{\partial^2 u}{\partial y^2} = 2\sin x \cos(2y + x)$$

At point (0, 0), we get $r = 0$, $s = 0$, $t = 0$.

$r\,t - s^2 = 0$ *i.e.*, u has a saddle point at (0, 0).

Now at point (π/3. π/3), we get

$r = 2 \sin(\pi/3) \cos \pi = -\sqrt{3}$.

$s = \sin(4\pi/3) = -\sin(\pi/3) = -\frac{\sqrt{3}}{2}$

and $t = 2 \sin(\pi/3) \sin \pi = -\sqrt{3}$

$\Rightarrow \quad r\,t - s^2 = \frac{9}{4} > 0$ and $r < 0$.

Hence the given function u is maximum at (π/3, π/3) and has a saddle point at (0, 0).

2.6 Summary

An optimization problem with single variable without any restriction is called a single variable unconstrained optimization problem.

Let $f'(x_o) = f''(x_0) = f'''(x_0) = \ldots\ldots\ldots\ldots\ldots = f^{(n-1)}(x_0)$, but $f^{\,n}(x_0) \neq 0$ then $f(x_0)$ has:

(i) Minimum value of $f(x)$ if $f^{(n)}(x_0) > 0$ and n is even.

(ii) Maximum value of $f(x)$ if $f^{(n)}(x_0) < 0$ and n is even.

(iii) Neither maxima nor minima of f *i.e.* point of inflexion if $f^{(n)}(x_0) \neq 0$ and n is odd.

An optimization problem with two or more variable without any restriction is called a multi-variable unconstrained optimization problem.

Let $u = f(x_1, x_2)$ be a function of x_1 and x_2. Then at any point (a_1, b_1):

(i) If $rt-s^2 > 0$ and $r < 0$ then $f(x_1, x_2)$ is maximum at (a_1, b_1).

(ii) If $rt-s^2 > 0$ and $r > 0$ then $f(x_1, x_2)$ is minimum at (a_1, b_1).

(iii) If $rt-s^2 < 0$ then $f(x_1, x_2)$ has saddle point at (a_1, b_1).

Suppose $u = f(x_1, x_2, x_3, \ldots, x_n)$ is a function of $x_1, x_2, x_3, \ldots, x_n$. Then:

(i) If $H_1, H_2, H_3, \ldots$ are positive (*i.e.*, H is positive definite) then the function $f(x_1, x_2, x_3, \ldots, x_n)$ has minimum at P.

(ii) If $H_1, H_2, H_3, \ldots$ are alternately negative, positive, negative (*i.e.*, H is negative definite) then the function $f(x_1, x_2, x_3, \ldots, x_n)$ has maximum at P.

(iii) H_1 and $H_3, \ldots$ are not of same sign and $H_2 = 0$ (*i.e.*, semi definite or indefinite) then the function $f(x_1, x_2, x_3, \ldots, x_n)$ has a saddle point at P.

2.7 Terminal Questions

Q.1. Explain the multivariable optimization problems.

Q.2. Write the procedure to determine the extreme points of function of n-variable.

Q.3. Determine the output x, which maximizes profit P given by the relationship

$$P = 5000 + 1200x - x^2.$$

Q.4. The price of a commodity is the function of its quantity q to be purchased which is given as $10.59\, e^{-0.01q}$. Find the quantity for which total revenue is maximum.

Q.5. The total revenue R of a firm is given by $R = 20x - 2x^2$, Where x represents the quantity sold and $C = x^2 - 4x + 20$. Determine the value of x for which the revenue will be maximum and also find the profit price and total revenue.

Q.6. Determine the local maxima or minima of the function

$$f(x) = \frac{x^5}{5} - 7\frac{x^4}{4} + 17\frac{x^3}{3} - 17\frac{x^2}{2} + 6x + 10 \ \text{for } x \in R.$$

Q.7. Determine the maxima and minima of the function $xy(a-x-y)$.

Q.8. Find the extreme points $f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6.$

Q.9. Determine the maxima and minima of the function $u = x^2+y^2+z^2+x-2z-xy$.

Q.10. Find the stationary point of u, where $u = x^2+y^2$.

Answers

3. 3,65,000.

4. 100.

5. At $x = 5$ total revenue = 50, profit = 25, price = 5 per unit.

6. At point $x = 1$ give point of inflexion, at point $x = 2$ give local maxima and $x = 3$ give local minima.

7. At points (0, 0), (0, a) and (a, 0) u is neither maxima nor minima. At point $\left(\frac{a}{3}, \frac{a}{3}\right)$ u is minimum if $a < 0$ and u is maximum if $a > 0$.

8. At (0, 0) f is minimum and at $\left(-\frac{4}{3}, -\frac{8}{3}\right)$ f is maximum.

9. At $\left(-\frac{2}{3}, -\frac{1}{3}, 1\right)$ u is minimum.

10. At (0, 0) u give saddle point.

UNIT- 3: CONSTRINED OPTIMIZATION TECHNIQUES

Structure

3.1 Introduction

In this unit, we delve into classical optimization techniques, particularly focusing on constrained multivariable problems. The discussion extends to equality and inequality constraints, thoroughly exploring these concepts through illustrative examples. Classical optimization techniques play a crucial role in obtaining optimal solutions for problems that involve continuous and differentiable functions. These techniques are analytical in nature, enabling the identification of maximum and minimum points for both unconstrained and constrained continuous objective functions. For equality constrained problems, we use Direct substitution method and Lagrange's multiplier method.

3.2 Objectives

After studying this unit the learner will be able to understand:

- the constrained optimization problems
- the Direct substitution method
- the Lagrange Multipliers method

- the constrained multi-variables optimization problem with inequality constraints

3.3 Constrained Optimization Problems

The optimization problem of a continuous and differentiable function subject to equality constraints:

Optimize (Max or Min) Z= f(X)

Subject to constraints (s.t.)

$$g_j(X) = 0 \; ; \; j = 1, 2, 3, \ldots\ldots\ldots, m$$

Where
$$X=\begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ \ldots \\ x_n \end{pmatrix}$$

Here m is less than or equal to n. For solving this type of problem (constrained optimization problem), there are several method.

Here we discuss two important methods, which are as follow:

1. Direct Substitution Method

2. Lagrange Multipliers Method

3.4 Direct Substitution Method

In Direct substitution method, putting the value of any variable from the constraint set is put in the objective function. The given constrained optimization problem reduces to unconstrained optimization problem which can be solved by using unconstrained optimization method.

3.5 Lagrange Multipliers Method

Let us consider a general problem with n variables and m equality constraints:

Optimize Z=f(X)

s.t. $$g_j(X) = 0 \; ; \; j = 1, 2, 3, \ldots\ldots\ldots, m \; (m<n)$$

$$X \geq 0$$

Where $$X = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ \cdots \\ x_n \end{pmatrix}$$

Now we defined a function

$$L(x_1, x_2, \ldots\ldots, x_n, \lambda_1, \lambda_2, \ldots\ldots, \lambda_m) = f(X) + \sum_{j=1}^{m} \lambda_j \, g_j(X) \qquad \ldots.. (3)$$

Here, $\lambda_1, \lambda_2, \ldots\ldots, \lambda_m$ are known as Lagrange's undetermined multipliers.

The necessary conditions for extreme of L are

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0 \qquad \ldots. (4)$$

and $$\frac{\partial L}{\partial \lambda_j} = 0; \begin{pmatrix} i = 1, 2, \ldots..n \\ j = 1, 2, \ldots..m \end{pmatrix} \qquad \ldots.. (5)$$

Solving equations (4) and (5), we get

$$X^* = \begin{pmatrix} x_1^* \\ x_2^* \\ \cdots \\ \cdots \\ x_n^* \end{pmatrix} \text{ and } \lambda^* = \begin{pmatrix} \lambda_1^* \\ \lambda_2^* \\ \cdots \\ \cdots \\ \lambda_m^* \end{pmatrix}$$

The sufficient condition for the given function to have extreme point at X^* and the values of k obtained from the equation

$$\begin{bmatrix} L_{11}-k & L_{12} & \dots & L_{1n} & g_{11} & g_{21} & \dots & g_{m1} \\ L_{21} & L_{22}-k & \dots & L_{2n} & g_{12} & g_{22} & \dots & g_{m2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ L_{n1} & L_{n2} & \dots & L_{nn}-k & g_{1n} & g_{2n} & \dots & g_{mn} \\ g_{11} & g_{12} & \dots & g_{1n} & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ g_{m1} & g_{m2} & \dots & g_{mn} & 0 & 0 & \dots & 0 \end{bmatrix} = 0$$

must be the same sign. If all the eigen values of k are negative then the given function gives a maxima and if all the eigen values k are positive then the given function gives a minima. In case if the some eigen values are zero or some of different sign then gives a saddle point. In above L_{ij} and g_{ij} denoted by $\frac{\partial^2 L}{\partial x_i \partial x_j}$ and $\frac{\partial g_j}{\partial x_i}$ respectively.

3.6 Constrained Multivariable Optimization Problems with Inequality Constrains

Let us consider a problem

Optimization (Max or Min) Z=f(X) …. (1)

s.t. $g_j(X) \le 0$; $j = 1, 2, 3, \dots\dots\dots, m$ ….. (2)

Where $X=\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ x_n \end{pmatrix}$

The inequality constrains in equation (2) can be convert into equality constraints by adding slack variables as

$$g_j(X) + y_j^2 = 0; \quad j = 1, 2, 3, \dots\dots\dots, m$$

Now the given problem becomes

Optimize (Max or Min) Z=f(X)

s.t. $G_j(X, Y) = g_j(X) + y_j^2 = 0; \quad j = 1, 2, 3, \ldots, m$

Where $X=\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ \cdots \\ x_n \end{pmatrix}$ and $Y=\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ \cdots \\ y_m \end{pmatrix}$

After converting the inequality constraints in equality constraint then the above problem can be solved by Lagrange's multiplier method.

Examples

Example.1. Find the optimum solution of the constrained multivariable problem:

Minimize $Z = x_1^2+(x_2+1)^2+(x_3-1)^2$

s.t. $x_1+5x_2-3x_3 = 6.$

Solution. Given that

$$Z = x_1^2+(x_2+1)^2+(x_3-1)^2 \quad \ldots. (1)$$

and $x_1+5x_2-3x_3 = 6$

$$\Rightarrow \quad x_3 = \frac{x_1+5x_2-6}{3} \quad \ldots. (2)$$

Using equations (1) and (2), we get

$$Z = x_1^2+(x_2+1)^2+\frac{1}{9}(x_1+5x_2-9)^2 \quad \ldots. (3)$$

Differentiate both side partially of (3) with respect to x_1 and x_2 respectively, we get

$$\frac{\partial Z}{\partial x_1} = 2x_1 + \frac{2}{9}(x_1+5x_2-9)$$

And $$\frac{\partial Z}{\partial x_2} = 2(x_2+1) + \frac{10}{9}(x_1+5x_2-9)$$

For maxima and minima of Z, we have

$$\frac{\partial Z}{\partial x_1} = 0$$

and $$\frac{\partial Z}{\partial x_2} = 0$$

$\Rightarrow$ $$2x_1 + \frac{2}{9}(x_1 + 5x_2 - 9) = 0$$

and $$2(x_2 + 1) + \frac{10}{9}(x_1 + 5x_2 - 9) = 0$$

Solving above these, we get

$$x_1 = \frac{2}{5} \text{ and } x_2 = 1$$

Differentiate again partially equation (4), we get

$$r = \frac{\partial^2 Z}{\partial x_1^2} = 2 + \frac{2}{9} = \frac{20}{9}$$

$$t = \frac{\partial^2 Z}{\partial x_2^2} = 2 + \frac{50}{9} = \frac{68}{9}$$

and $$s = \frac{\partial^2 Z}{\partial x_1 x_2} = \frac{10}{9}$$

At $\left(\frac{2}{5}, 1\right)$, we get

rt-s^2 = $\left(\frac{20}{9}\right)\left(\frac{68}{9}\right)$-$\left(\frac{10}{9}\right)^2$ = 1260 > 0

and $r > 0$

Hence the function Z is minimum at point $\left(\frac{2}{5}, 1\right)$ and the minimum value of the function Z is $\frac{28}{5}$.

Example.2. Determine the dimensions of a large volume box that can be inscribed in a sphere of radius *a*.

Solution: Let x, y and z be the dimensions of the box with respect to origin O and OX, OY, OZ are the reference axes. The volume of given box is

$$V = 8xyz \qquad \text{..... (1)}$$

It is given that the box is to be inscribed in a sphere of radius 'a' *i.e.*,

$$x^2+y^2+z^2 = a^2 \qquad \text{..... (2)}$$

Eliminating z from equations (1) and (2), we get

$$V = 8xy\,(a^2 - x^2 - y^2)^{1/2} \qquad \text{..... (3)}$$

Differentiate both side partially (3) with respect to x and y respectively, we get

$$\frac{\partial V}{\partial x_1} = 8y\left[x.\frac{1}{2}\left(a^2 - x^2 - y^2\right)^{-1/2}(-2x) + \left(a^2 - x^2 - y^2\right)^{1/2}\right]$$

$$= 8y\left[-\frac{x^2}{\left(a^2 - x^2 - y^2\right)^{1/2}} + \left(a^2 - x^2 - y^2\right)^{1/2}\right]$$

$$= 8y\left[\frac{\left(a^2 - 2x^2 - y^2\right)}{\left(a^2 - x^2 - y^2\right)^{1/2}}\right] \qquad \text{..... (4)}$$

and $$\frac{\partial V}{\partial y} = 8x\left[\frac{\left(a^2 - x^2 - 2y^2\right)}{\left(a^2 - x^2 - y^2\right)^{1/2}}\right] \qquad \text{..... (5)}$$

For maxima and minima of V, we have

$$\frac{\partial V}{\partial x_1} = 0$$

and $$\frac{\partial V}{\partial y} = 0$$

$\Rightarrow$ $$8y\left[\frac{\left(a^2 - 2x^2 - y^2\right)}{\left(a^2 - x^2 - y^2\right)^{1/2}}\right] = 0$$

and $$8x\left[\frac{\left(a^2 - x^2 - 2y^2\right)}{\left(a^2 - x^2 - y^2\right)^{1/2}}\right] = 0$$

$\Rightarrow \qquad a^2 - 2x^2 - y^2 = 0$ ….. (6)

and $\qquad a^2 - x^2 - 2y^2 = 0$ ….. (7)

Solving these, we get

$$x = \frac{a}{\sqrt{3}} \quad \text{and} \quad y = \frac{a}{\sqrt{3}}$$

Differentiate again partially equations (4) and (5), we get

$$r = \frac{\partial^2 V}{\partial x^2} = 8y\left[\frac{\left(a^2 - x^2 - y^2\right)^{1/2}(-4x) - \left(a^2 - 2x^2 - y^2\right).\frac{1}{2}\left(a^2 - x^2 - y^2\right)^{-1/2}(-2x)}{\left(a^2 - x^2 - y^2\right)}\right]$$

$$= 8y\left[\frac{\left(a^2 - x^2 - y^2\right)(-4x) + x\left(a^2 - 2x^2 - y^2\right)}{\left(a^2 - x^2 - y^2\right)^{3/2}}\right]$$

$$= \frac{-32xy}{\left(a^2 - x^2 - y^2\right)^{1/2}}$$ {Using equation (6)}

Similarly we get

$$t = \frac{\partial^2 V}{\partial y^2} = \frac{-32xy}{\left(a^2 - x^2 - y^2\right)^{1/2}}$$

and $$s = \frac{\partial^2 V}{\partial x \partial y} = \frac{-16x^2}{\left(a^2 - x^2 - y^2\right)^{1/2}}$$

Now we have

$$rt - s^2 = \left(\frac{-32xy}{\left(a^2 - x^2 - y^2\right)^{1/2}}\right)\left(\frac{-32xy}{\left(a^2 - x^2 - y^2\right)^{1/2}}\right) + \left(\frac{-16x^2}{\left(a^2 - x^2 - y^2\right)^{1/2}}\right)^2$$

$$= \frac{256x^2(x^2 + 4y^2)}{\left(a^2 - x^2 - y^2\right)}$$

At $\left(\frac{a}{\sqrt{3}}, \frac{a}{\sqrt{3}}\right)$, rt-s^2 is positive and $r < 0$ i.e., V is maximum.

Hence the maximum value of the box of large volume

$$V = 8xyz$$

$$= 8\left(\frac{a}{\sqrt{3}}\right)\left(\frac{a}{\sqrt{3}}\right)\left(\frac{a}{\sqrt{3}}\right) = \frac{8a^3}{3\sqrt{3}}.$$

Example.3. Determine the optimal values of Z = $x^2+y^2+z^2$ s.t. $4x+y^2+2z=14$.

Solution: It is given that

$$Z = x^2+y^2+z^2 \qquad \text{.... (1)}$$

and $\quad 4x+y^2+2z=14 \quad \Rightarrow$

$$g(x, y, z) = 4x+y^2+2z-14 = 0 \qquad \text{.... (2)}$$

Now construct the Lagrangian function L is

$$L(x, y, z; \lambda) = x^2+y^2+z^2 + \lambda(4x+y^2+2z-14) \qquad \text{.... (3)}$$

The necessary conditions for extreme L are

$$\left.\begin{aligned} \frac{\partial L}{\partial x} &= 2x + 4\lambda = 0 \\ \frac{\partial L}{\partial y} &= 2y + 2y\lambda = 0 \\ \frac{\partial L}{\partial z} &= 2z + 2\lambda = 0 \\ \frac{\partial L}{\partial \lambda} &= 4x + y^2 + 2z - 14 = 0 \end{aligned}\right\} \qquad \text{.... (4)}$$

Solving above these, we get

$$x = -2\lambda, \quad z = -\lambda \quad \text{and} \quad \lambda = -1.$$

$\Rightarrow \qquad x = 2,\ z = 1$ and $y = \pm 2$,

Putting $x = -2\lambda$, $z = -\lambda$, $y = 0$ in the equation (2), we get

$$\lambda = -1.4.$$

Here (2, 2, 1, -1), (2, -2, 1, -1) and (-2λ, 0, $-\lambda$, λ) or (2.8, 0, 1.4, -1.4) are the extreme points.

Differentiate again partially equation (4), we get

$$\frac{\partial^2 L}{\partial x^2} = 2, \; \frac{\partial^2 L}{\partial x \partial y} = 0, \; \frac{\partial^2 L}{\partial x \partial z} = 0,$$

$$\frac{\partial^2 L}{\partial y \partial x} = 0, \frac{\partial^2 L}{\partial y^2} = 2 + 2\lambda, \; \frac{\partial^2 L}{\partial y \partial z} = 2,$$

$$\frac{\partial^2 L}{\partial z \partial x} = 0, \; \frac{\partial^2 L}{\partial z \partial y} = 0, \; \frac{\partial^2 L}{\partial z^2} = 2,$$

$$\frac{\partial g}{\partial x} = 4, \; \frac{\partial g}{\partial y} = 2y, \quad \frac{\partial g}{\partial z} = 2.$$

The sufficient condition for extreme point is

$$H = \begin{vmatrix} \frac{\partial^2 L}{\partial x^2} - k & \frac{\partial^2 L}{\partial x \partial y} & \frac{\partial^2 L}{\partial x \partial z} & \frac{\partial g}{\partial x} \\ \frac{\partial^2 L}{\partial yx \partial x} & \frac{\partial^2 L}{\partial y^2} - k & \frac{\partial^2 L}{\partial y \partial z} & \frac{\partial g}{\partial y} \\ \frac{\partial^2 L}{\partial z \partial x} & \frac{\partial^2 L}{\partial z \partial y} & \frac{\partial^2 L}{\partial z^2} - k & \frac{\partial g}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} & 0 \end{vmatrix} = 0$$

i.e.,
$$H = \begin{vmatrix} 2-k & 0 & 0 & 4 \\ 0 & 2+2\lambda-k & 0 & 2y \\ 0 & 0 & 2-k & 2 \\ 4 & 2y & 2 & 0 \end{vmatrix} = 0$$

or $\quad 4(2-k)[-10+5k-10\lambda-2y^2+y^2k] = 0$

At point (2, 2, 1, -1) from equation (2), we have

$$(2-k)(-10+5k+10-8+4k) = 0$$

$\Rightarrow \quad k = 2, 8/9.$

i.e., the values of k are positive then there is a minima.

At point (2, -2, 1, -1) from equation (2), we have

$$(2-k)(-10+5k+10-8+4k) = 0$$

$\Rightarrow \quad k = 2, 8/9$

Also values of *k* are positive then there is a minima.

At point (2.8, 0, 1.4, -1.4) from equation (2), we have

$$(2-k)(-10+5k+14-0+0) = 0$$

$$\Rightarrow \quad k = 2, -4/5.$$

i.e. the values of *k* are positive and negative (neither maxima nor minima) *i.e.*, saddle point.

Example.4. Solve

$$\textbf{Min } f(x) = \frac{1}{2}(x_1^2+x_2^2+x_3^2)$$

s.t. $g_1(x) = x_1 - x_2 = 0$

and $g_2(x) = x_1+x_2+x_3-1 = 0$ **using Lagrange's Multiplier Method.**

Solution: Given that

$$\text{Min } f(x) = \frac{1}{2}(x_1^2+x_2^2+x_3^2) \quad \ldots (1)$$

$$\text{s.t.} \quad g_1(x) = x_1 - x_2 = 0 \quad \ldots (2)$$

$$\text{and} \quad g_2(x) = x_1+x_2+x_3-1 = 0 \quad \ldots (3)$$

Construct the Lagrangian function L is

$$L(x_1, x_2, x_3; \lambda_1, \lambda_2) = \frac{1}{2}(x_1^2+x_2^2+x_3^2) + \lambda_1(x_1-x_2) + \lambda_2(x_1+x_2+x_3-1) \quad \ldots .(4)$$

Necessary conditions for extreme of L are

$$\left.\begin{aligned} \frac{\partial L}{\partial x_1} &= x_1 + \lambda_1 + \lambda_2 = 0 \\ \frac{\partial L}{\partial x_2} &= x_2 - \lambda_1 + \lambda_2 = 0 \\ \frac{\partial L}{\partial x_3} &= x_3 + \lambda_2 = 0 \\ \frac{\partial L}{\partial \lambda_1} &= x_1 - x_2 = 0 \\ \frac{\partial L}{\partial \lambda} &= x_1 + x_2 + x_3 - 1 = 0 \end{aligned}\right\} \quad \ldots ..(5)$$

Solving above these, we get

$$x_1 = x_2 = x_3 = \frac{1}{3}; \quad \lambda_1 = 0, \lambda_2 = -\frac{1}{3}$$

Differentiate again partially (6), we get

$$\frac{\partial^2 L}{\partial x_1^2} = 1, \ \frac{\partial^2 L}{\partial x_1 \partial x_2} = 0, \ \frac{\partial^2 L}{\partial x_1 \partial x_3} = 0,$$

$$\frac{\partial^2 L}{\partial x_2 \partial x_1} = 0, \ \frac{\partial^2 L}{\partial x_2^2} = 1, \ \frac{\partial^2 L}{\partial x_2 \partial x_3} = 0,$$

$$\frac{\partial^2 L}{\partial x_3 \partial x_1} = 0, \ \frac{\partial^2 L}{\partial x_3 \partial x_2} = 0, \ \frac{\partial^2 L}{\partial x_3^2} = 1,$$

$$\frac{\partial g_1}{\partial x_1} = 1, \ \frac{\partial g_1}{\partial x_2} = -1, \quad \frac{\partial g_1}{\partial x_3} = 0,$$

$$\frac{\partial g_2}{\partial x_1} = 1, \ \frac{\partial g_2}{\partial x_2} = 1, \quad \frac{\partial g_2}{\partial x_3} = 1.$$

The sufficient condition for extreme point is

$$H = \begin{vmatrix} 1-k & 0 & 0 & 1 & 1 \\ 0 & 1-k & 0 & -1 & 1 \\ 0 & 0 & 1-k & 0 & 1 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{vmatrix} = 0$$

This implies we get $k = 1, 1, 1$.

Here all the values if k are same sign and positive. *i.e.*, f is minimum at $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, -\frac{1}{3}\right)$.

Hence the minimum value of the function $f = \frac{1}{2}\left(\frac{1}{9} + \frac{1}{9} + \frac{1}{9}\right) = \frac{1}{6}$.

3.7 Summary

An optimization problem with no restriction on the decision variables is called an unconstrained optimization problem. An optimization problem with single variable without any restriction is called a single variable unconstrained optimization problem. An optimization problem with two or more than two variables with no restriction on the decision variables is called a multi-variable unconstrained optimization problem.

The sufficient conditions for the function $f(x_1, x_2)$ at the point (a, b):

(a) if $rt-s^2 > 0$ and $r < 0$ then $f(x_1, x_2)$ is maximum.

(b) if $rt-s^2 > 0$ and $r > 0$ then $f(x_1, x_2)$ is minimum.

(c) if $rt-s^2 < 0$ then $f(x_1, x_2)$ has saddle point.

An optimization problem with restriction on the decision variables is called a constrained optimization problem. Lagrange's multiplier method is used to solve the equality constrained problems.

3.8 Terminal Questions

1. Write a short note on Constrained Optimization Problem.

2. Explain the Lagrange Multiplier method.

3. Find the extreme values of $x^2+y^2+z^2$ when $ax+by+cz= p$.

4. Find the maximum and minimum values of $u = x+y+z$ s.t. $\frac{a}{x}+\frac{b}{y}+\frac{c}{z}=1$.

5. Find the extreme value of $x^p y^q z^r$ s.t. $\frac{a}{x}+\frac{b}{y}+\frac{c}{z}=1$.

6. Find the maxima and minima of $x^2+y^2+z^2$ s.t. $ax^2+by^2+cz^2=1$ and $lx+my+nz=0$.

7. Using (i) Direct method and

   (ii) Lagrange multiplier method to solve the following problem:

   Minimize $f = 2x_1^2+2x_2^2+x_3^2+2x_1x_2-8x_1-6x_2-4x_3+2x_1x_3+9$

   s.t. $x_1+x_2+2x_3 = 3$

8. Using Lagrange multiplier method to solve the following problems:

   (i) Minimize $f(x_1, x_2) = 3x_1^2+x_2^2+2x_1x_2+6x_1+2x_2$

s.t. $2x_1 - x_2 = 4.$

(ii) Optimize $f(x_1, x_2) = 6x_1x_2$

s.t. $2x_1 + x_2 = 10$

Also state whether the stationary point is a maxima or minima.

Answer

---

3. Minimum value of u is $\frac{p^2}{a^2+b^2+c^2}$.

4. $u = \left(\sqrt{a} + \sqrt{b} + \sqrt{c}\right)^2.$

5. $\frac{a^p b^q c^r}{p^p q^q r^r}(p+q+r)^{p+q+r}.$

6. $\frac{l^2}{au-1} + \frac{m^2}{bu-1} + \frac{n^2}{cu-1} = 0.$

7. $\left(\frac{4}{3}, \frac{7}{9}, \frac{4}{9}\right).$

8. (i) (1, -2), minimize $f = 5$.

(ii) $\left(\frac{5}{2}, \frac{5}{2}\right)$, maximize f =75.

Master of Science
PGMM -102N
Classical Optimization Techniques

U. P. Rajarshi Tandon
Open University

Block

2 Non-Linear Programming Problems

Unit- 4

Non-Linear Programming-I

Unit- 5

Non-Linear Programming-II

Unit- 6

Quadratic Programming

Unit- 7

Separable Programming Problem

Block-2

Non-Linear Programming Problems

There are a number of non-linear optimization problems which have analytical solution. However a lot many non-linear programming programs are also there which do not fall in analytically solvable category and thus defy any general analytical approach. The only way to tackle such problems is to choose a point to start with, find the value of the function to be optimized at that point, take a tentative short step from the point in some direction and find the function value at the new point. If the new value is better (less in minimum problem and large in maximum problem) than the previous value, then discard the first point in favour of the second point, otherwise stay at the first point and move in some other direction. Go on doing this till a point is found from where no further change in its value is possible in any direction and the function value is the optimum within some prescribed tolerable limit. It is not intended here to go into the theoretical discussion about the derivations and rates of convergence of various methods.

Analytical methods have been devised to solve some of these non-linear programming problems like quadratic programming problem, separable programming problem and geometric programming problem etc. However there are many other non-linear programming problems which defy to yield solution by analytical methods. Our objective in this block is to discuss the methods to solve such problems by using various techniques called search techniques. Such problems can be divided further into two categories; constrained non-linear programming problems and unconstrained non-linear programming problems.

In the fourth unit, we shall discussed about the unconstrained non-linear programming problems and in the fifth unit we deal with constrained non-linear programming problems. Quadratic programming problems and separable programming problems are discussed in unit sixth and seventh respectively.

UNIT-4: Non-Linear Programming-I

Structure

4.1 Introduction

Mathematical programming problem consisting in getting the maximum or minimum of $f(X)$ restricted with constraints $g_j(X) \leq, =$ or ≥ 0, where $f(X)$ and $g_i(X)$ are real valued functions of $X = (x_1, x_2, \ldots.., x_n)$ in n dimensional space E_n. In case some or all of the functions $f(X)$, $g_j(X)$; $j = 1, 2, \ldots., m$ are non-linear, then the mathematical programming problem is called a non-linear programming problem. In this unit we shall discuss the unconstrained non-linear optimization problems. For solving the unconstrained non-linear optimization problems, the methods divided into two categories: (i) Direct search method and (ii) Indirect search method.

Using direct search methods we apply the Fibonacci-Search Plan Method; Golden Section Method; Univariate search technique; Pattern Search Methods: (i) Hooke and Jeeves Method (ii) Powell's Method and Indirect search methods we apply the Steepest decent method.

4.2 Objectives

After reading this unit the learner should be able to understand about the:

- Search techniques (one dimension)
- Fibonacci Method
- Golden Section Method
- Univariate search technique
- Pattern Search Methods
- Hooke and Jeeves Method
- Powell's Method
- Steepest Decent Method (Cauchy Method)

4.3 Search Techniques (One dimension)

We have already discussed classical techniques which were easily applicable if the expressions for objective function and the constraints were simple. But when objective function and constraints are complicated, then classical analytic methods fail and we have to use numerical methods. In most of these numerical techniques, the only requirement is that the function involved should be computable. After introducing a few elementary definitions, we shall discuss these methods.

Unimodal function: A function $f(x)$ with a unique optimal value (either a unique maxima or minima) in $[a,b]$ is known as unimodal function. The following Figures 4.1 and 4.2 showing the unimodal function in the interval $[a,b]$.

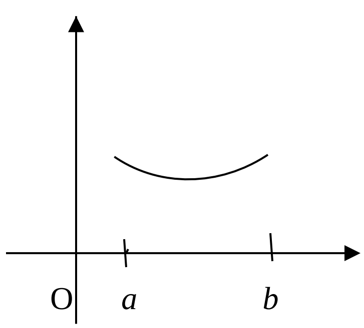


Fig-4.1

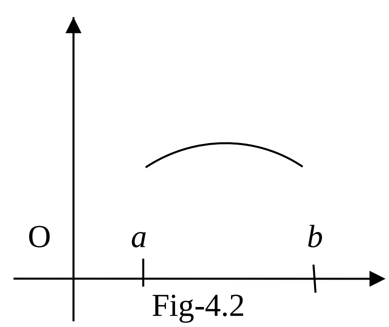


Figure 4.1 and 4.2 unimodal functions in [*a*, *b*]

Let $y = f(x)$ be a unimodal function is $[a,b]$ and $x_1, x_2\ (x_1 < x_2)$ be two points in $[a,b]$. Calculate $f(x_1)$

and $f(x_2)$. Then only following three and no other than these possibilities as shown in figures 4.3, 4.4 and 4.5 are possible.

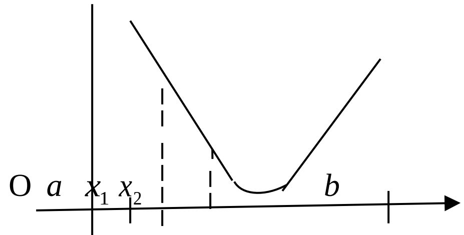


Fig-4.3 (a)

$$f(x_1) > f(x_2)$$
$$\Rightarrow \min \in [x_2, b]$$

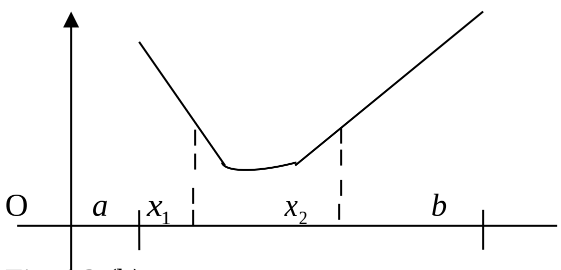


Fig-4.3 (b)

$$f(x_1) = f(x_2)$$
$$\Rightarrow \min \in [x_1, x_2]$$

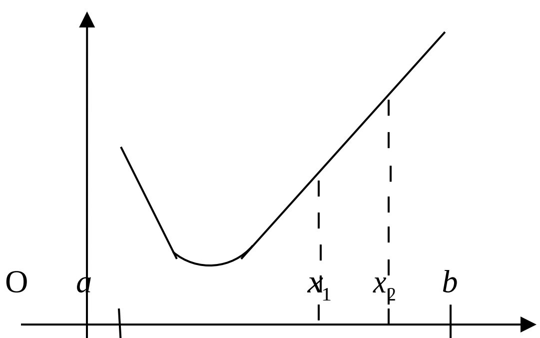


Fig-4.3 (c)

$$f(x_1) < f(x_2)$$
$$\Rightarrow \min \in [a, x_1]$$

Figure 4.3 (a), 4.3 (b) and 4.3 (c) shown the interval to which minima belongs

Interval of uncertainty

Initially the interval $[a,b]$ in which the optimum of the objective (response) function is needed is called uncertainty interval and after two experiments (funding $f(x_1)$ and $f(x_2)$) the uncertainty interval reduces to $[x_2, b]$ or $[x_1, x_2]$ or $[a_1, x_1]$ (see Figures 4.3 (a), 4.3 (b) and 4.3 (c)).

Experiment

Measure or calculation of the response (objective) function $f(x)$ for any set of values x_i 's in uncertainty interval $[a,b]$ is called an experiment in context of search techniques.

Measure of effectiveness

Let L_0 be the initial uncertainty interval and L_n be the uncertainty interval after n experiments, then the ratio $\frac{L_0}{L_n} = \alpha \leq 1$ is called the measure of effectiveness.

4.4 Fibonacci Method

In Fibonacci method the initial interval of uncertainty say $[a,b]$ is given and the function to the optimized to a given degree of accuracy must be unimodal in the interval of uncertainty. Also in this we use Fibonacci numbers. So first we define Fibonacci sequence. The sequence $\{F_n\}$ where

$$F_n = F_{n-1} + F_{n-2} \qquad \text{..........(4.1)}$$

n is an integer greater than $1\,(n>1)$ and

$$F_0 = F_1 = 1 \qquad \text{.....(4.2)}$$

is called Fibonacci sequence.

Thus

F_0	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	....
1	1	2	3	5	8	13	21	34	55	89	144	233	....

Procedure

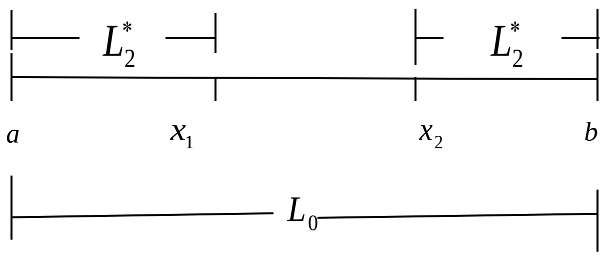
Let $[a,b]$ be the initial interval of uncertainty with length

$$\left.\begin{aligned} L_0 &= b - a \\ L_2^* &= \frac{F_{n-2}}{F_n} L_0 \end{aligned}\right\} \qquad \text{....(4.3)}$$

and 'n' is the number of experiments to be performed. Choose x_1 and x_2 in L_0 such that

$$x_1 = a + \frac{F_{n-2}}{F_n} L_0 \qquad ....(4.4)$$

$$x_2 = b - \frac{F_{n-2}}{F_n} L_0 \qquad ....(4.5)$$



Thus $b - x_2 = b - b + \frac{F_{n-2}}{F_n} L_0$ (using Equation 4.5)

$= x_1 - a$ (using Equation 4.4)

Here x_1-a =b-x_2 showing that x_1 and x_2 are symmetrically placed in respect to end points a and b of interval of length L_0. Also

$$x_2 = b - \frac{F_{n-2}}{F_n} L_0 = a + L_0 - \frac{F_{n-2}}{F_n} L_0 = a + \left(\frac{F_n - F_{n-2}}{F_n} \right) L_0 = a + \frac{F_{n-1}}{F_n} L_0 \qquad ....(4.6)$$

Let the problem be of minimization and

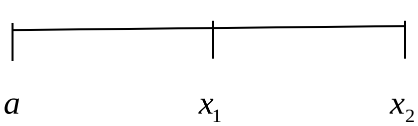
(i) If $x_1 < x_2$ and $f(x_1) < f(x_2)$, then by the assumption of unimodal it can be concluded that minimum does not lie on the right of x_2 and hence reject interval $(x_2, b]$ and the next uncertainty interval is $[a, x_2]$.

(ii) If $x_1 < x_2$ and $f(x_1) > f(x_2)$, discard $[a, x_1]$, then the next uncertainty interval will be $[x_1, b]$. .

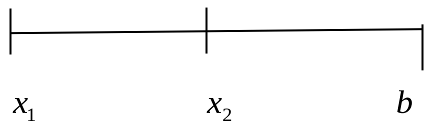
The length L_2 of this new uncertainty interval (for both case (i) (ii)) will be

$$L_2 = L_0 - L_2{*} = L_0 - \frac{F_{n-2}}{F_n} L_0 = L_0 \left(\frac{F_n - F_{n-2}}{F_n} \right) = \frac{F_{n-1}}{F_n} L_0 \quad \text{[using 4.1]} \qquad ...(4.7)$$

For case (i), the interval of further search will be $[a, x_2]$ in which one observation is at x_1



For case (ii), the interval of further search will be $[x_1, b]$ in which one observation is at x_2

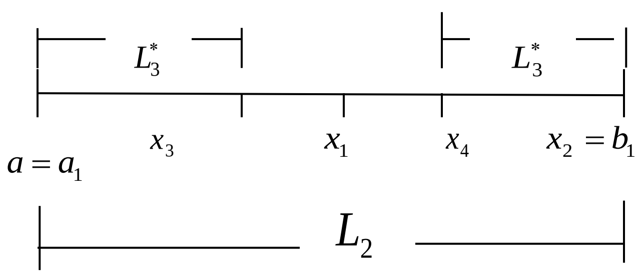


Next iteration is on the interval of uncertainty L_2. Let $L_2 = [a, x_2]$. Let the interval $[a, x_2]$ be predesignated as $[a_1, b_1]$.

As before choose x_3 and x_4 in interval $[a, x_2]$ of length L_2 such that

$$x_3 = a + \frac{F_{n-3}}{F_n} L_2 \qquad \text{......(4.8)}$$

And $$x_4 = x_2 - \frac{F_{n-3}}{F_{n-1}} L_2$$



Then we have

$$x_3 - a = \frac{F_{n-3}}{F_{n-1}} L_2 = x_2 - x_4 \qquad \text{...(4.9)}$$

Take $$\frac{F_{n-3}}{F_{n-1}} L_2 = L_3^* \qquad \text{...(4.10)}$$

Equality in equation (4.9) again implies that x_3 and x_4 selected according to (4.8) are symmetric with respect to end points of L_2.

Next consider $x_2 - x_4 = \frac{F_{n-3}}{F_{n-1}} L_2 = \frac{F_{n-3}}{F_n} L_0$ (using Equations 4.9 & 4.7) …(4.11)

Also $x_2 - x_1 = a + \frac{F_{n-1}}{F_n} L_0 - a - \frac{F_{n-2}}{F_n} L_0$ (using Equations 4.4 and 4.6)

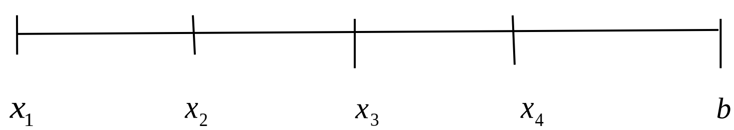
$$= \frac{F_{n-1} - F_{n-2}}{F_n} L_0 = \frac{F_{n-3}}{F_n} L_0 \qquad \text{…(4.12)}$$

$x_2 - x_1 = x_2 - x_4$ (Using Equations 4.11 & 4.12)

This shows that x_4 coincides with x_1. This implies that in second iteration and also in subsequent iterations, we need to find the function value at only one new point (x_3 in this case).

Similarly if for case (ii) the new search interval $[x_1, b]$ is retained, then we take

$$x_3 = x_1 + \frac{F_{n-3}}{F_n} L_2 \quad \text{and} \quad x_4 = b - \frac{F_{n-3}}{F_{n-1}} L_2$$



Here $x_2 = x_3$ i.e., x_2 and new point x_3 will coincide.

Length of the uncertainty interval after third experiment is given as

$$L_3 = L_2 - L_3{*} = \frac{F_{n-1}}{F_n} L_0 - \frac{F_{n-3}}{F_{n-1}} L_2 = \frac{F_{n-1}}{F_n} L_0 - \frac{F_{n-3}}{F_n} L_0 = \frac{F_{n-1} - F_{n-3}}{F_n} L_0 = \frac{F_{n-2}}{F_n} L_0$$

Thus $L_3 = \frac{F_{n-2}}{F_n} L_0$ …(4.13)

Repeating the whole process till the two experiments in the last are equidistant from both the end points of uncertainty interval $[a_n, b_n]$, then optimum point $x*$ can be approximately taken as

$$x* = \frac{a_n + b_n}{2}$$

Also the distance of the 8th experiment from one end the interval of uncertainty will be

$$L_j^* = \frac{F_{n-j}}{F_n} L_0, \qquad \text{….(4.14)}$$

And the length of uncertainty interval at this stage will be

$$L_j = \frac{F_{n-(j-1)}}{F_n} L_0, \; j = 2,3,.... \qquad \ldots(4.15)$$

Thus $\frac{L_n}{L_0} = \frac{F_1}{F_n} = \frac{1}{F_n} = \alpha \quad say$

The ratio $\frac{L_n}{L_0} = \alpha = \frac{1}{F_n}$ permits us the required number of experiment '*n*' to obtain the desired accuracy in searching the optimum of the given function.

Let us find '*n*' when as an example the minimum of $x^2 - 3x, \; 0 \le x \le 1.6$ is required within an interval of uncertainty equal to $0.25\, L_0$, where L_0 is the original interval of the uncertainty, then we have

$$\alpha = \frac{1}{F_n} \le 0.25$$

$$\Rightarrow \qquad F_n = \frac{1}{0.25} = 4$$

and hence $n = 4$.

Examples

Example.1. Determine the minimum of $x^2 - 2x, \; 0 \le x \le 1.5,$ taking *n*=4 using Fibonacci method.

Solution: Here $f(x) = x^2 - 2x, \; L_0 = 1.5 - 0 = 1.5, \; n = 4.$

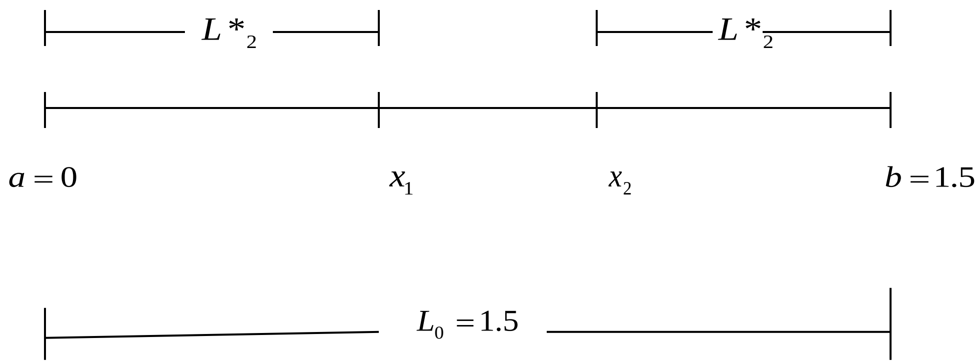
Clearly $f(x) = x^2 - 2x$ is unimodal in $[0, 1.5]$

First two experiments:

Let us take two points x_1 and x_2 as a distance $L*_2 = \frac{F_{n-2}}{F_n} L_0$ from the two points $a = 0$ and $b = 1.5$.

$$x_1 = a + L*_2 = 0 + \frac{F_{n-2}}{F_n} L_0$$

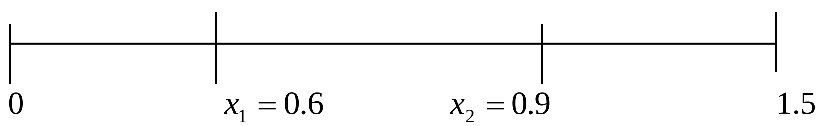
$$= \frac{2}{5} \times 1.5 = 0.6$$



Since x_1 and x_2 are symmetrically placed, x_2 will be at a distance 0.6 from $b = 1.5$ *i.e.*, $x_2 = 0.9$

$f(x_1) = f(0.6) = -0.84$ and $f(x_2) = f(0.9) = -.99$

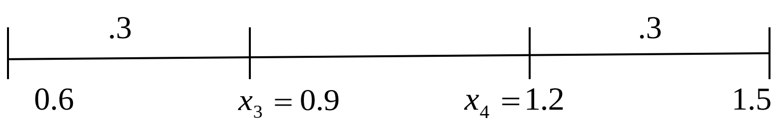
$\Rightarrow \quad f(0.6) > f(0.9)$



Thus reject [0, 0.6] and the new uncertainty interval is [0.6, 1.5].

Third experiments:

Take two points $x_3(0.9)$ and x_4 in the new uncertainty interval [0.6, 1.5], x_3 is the same as x_2 which is at a distance $L*_3 = 0.3$ from 0.6 and hence $x_4 = 1.2$ being symmetrically at a distance 0.3 from 1.5.



Thus we have

$$x_3 = 0.9 \text{ and } x_4 = 1.2.$$

$$f(x_3) = f(0.9) = -.99 \text{ and } f(x_4) = f(1.2) = -.96$$

$\Rightarrow \quad f(x_4) > f(x_3).$

Thus reject [1.2, 1.5] and hence the next uncertainty interval is [0.6, 1.2].

Here 0.9 is the middle point of uncertainty interval [0.6, 1.2]. Thus we do not get new point x_4 as $x_4 = x_3 = .9 = x*$

Thus Min $f(x) = -.99$ at $x = 0.9$. The correct minimum is $f(x) = -1$ at $x = 1$.

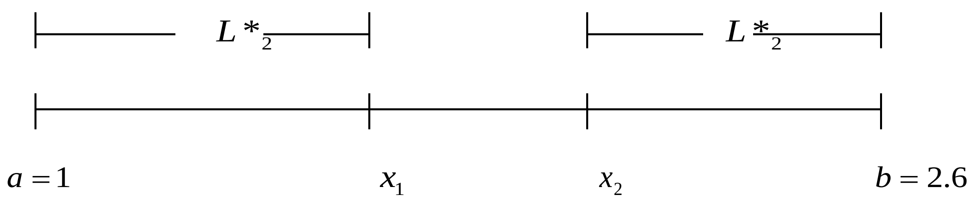
Example.2 Using the method of Fibonacci, find the minimum of $x^2 - 3x + 5$ is the interval [1, 2.6], taking *n*=6.

Solution: Here $f(x) = x^2 - 3x + 5,\; L_0 = 2.6 - 1 = 1.6,$ initial interval [1, 2.6], $n = 6$.

First two experiments:

Let us take two points x_1 and x_2 in [1, 2.6] at a distance.

$$L*_2 = \frac{F_{n-2}}{Fn} L_0 = \frac{F_4}{F_6} L_0 = \frac{5}{13} \times 1.6 = 0.6153846$$ from end points 1 and 2.6 respect.



$$x_1 = a + L*_2 = 1.6153846$$

$$x_2 = b - L*_2 = 2.6 - 6153846$$

$$= 1.9846154$$

$f(x_1) = 2.763314$ and $f(x_2) = 2.984852$

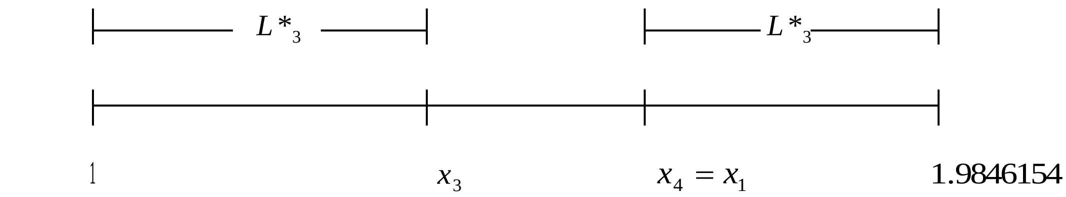
$$\Rightarrow f(x_1) < f(x_2),\; x_1 < x_2.$$

Reject $[1.9846154,\; 2.6]$ and new uncertainty interval is $[1,\; 1.9846154]$.

Third experiment:

Let x_3 and x_4 be two points in $[1,\; 1.9846154]$ at a distance

$$L*_3 = \frac{F_{n-3}}{F_n} L_0 = \frac{F_3}{F_6} L_0 = \frac{3}{13} \times 1.6 = 0.3692308 \;.$$



$x_3 = 1 + L*_3 = 1.3692308$

$x_4 = 1.9846154 - L*_3 = 1.6153846$

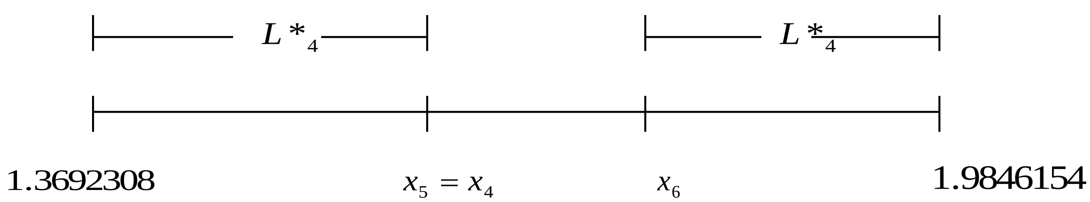
$f(x_3) = 2.7671006$, $f(x_4) = 2.763314$, $x_3 < x_4$ and $f(x_3) > f(x_4)$.

Reject [1, 1.3692308] and new uncertainty interval is [1.3692308, 1.9846154] using unimodality of the function.

Fourth experiment:

Let x_5 and x_6 be the two points in [1.3692308, 1.9846154] at a distance

$$L*_4 = \frac{F_{n-4}}{F_n} L_0 = \frac{F_2}{F_6} L_0 = \frac{2}{13} \times 1.6 = 0.2461538$$



$x_5 = 1.3692308 + L*_4 = 1.6153846$

$x_6 = 1.9846154 - L*_4 = 1.7384616$

$f(x_5) = 2.763314$ and $f(x_6) = 2.8068639$.

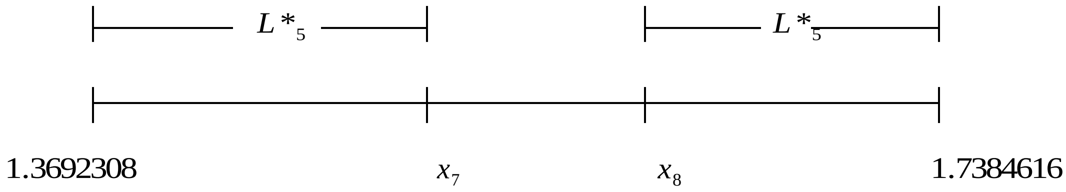
As $x_5 < x_6$ and $f(x_5) < f(x_6)$ by unimodality reject interval [1.7384616, 1.9846154].

Next interval of uncertainty is [1.3692308, 1.7384616].

Fifth experiment:

Let x_7 and x_8 be the two points in [1.3692308, 1.7384616] as a distance

$$L^*_5 = \frac{F_{n-5}}{Fn} L_0 = \frac{F_1}{F_6} L_0 = \frac{1}{13} \times 1.6 = 0.1230769$$



$x_7 = 1.3692308 + L^*_5 = 1.4923077$ by symmetry

$x_8 = 1.7384616 - L^*_5 = 1.6153846$

$\because f(x_7) = 2.750059$ and $f(x_8) = 2.7633136$

$\because x_7 < x_8$ and $f(x_7) < f(x_8)$, using unimodality reject in interval $[1.6153846, 1.7384616]$ and the next new uncertainty interval is [1.3692308, 1.6153846].

Sixth experiment:

$$L^*_6 = \frac{F_{n-6}}{F_n} L_0 = \frac{F_0}{F_6} L_0 = \frac{1}{13} \times 1.6 = 0.1230769 = L^*_5$$

Thus there is no fresh point and hence the final uncertainty interval is [1.3692308, 1.6153846] whose middle point

$$x^* = \frac{1.3692308 + 1.6153846}{2} = \frac{2.9846154}{2} = 1.4923077$$

Thus Min $f(x) = 2.750059$ for $x = 1.4923077$.

4.5 Golden Section Method

The next search method that we discuss is golden section method which differs from Fibonacci method in that the total number of experiments here to be conducted are unlimited, let us assume that '*n*' the total member of experiments to be conducted is very large satisfying the limit.

$$\lim_{n\to\infty} \frac{F_{n-1}}{F_n} = \lim_{n\to\infty} \frac{F_{n-2}}{F_{n-1}} = \lim_{n\to\infty} \frac{F_{n-3}}{F_{n-2}} = \frac{1}{\lambda} \ \ say$$

Then $L_2 = \lim_{n\to\infty} \frac{F_{n-1}}{F_n} L_0 = \frac{L_0}{\lambda}$, $L_3 = \lim_{n\to\infty} \frac{F_{n-2}}{F_n} L_0 = \lim_{n\to\infty} \frac{F_{n-2}}{F_{n-1}} \frac{F_{n-1}}{F_n} L_0 = \frac{L_0}{\lambda^2}$

Proceeding this way k times, we get $L_K = \frac{L_0}{\lambda^{K-1}} = \left(\frac{1}{\lambda}\right)^{K-1} L_0$

From Fibonacci sequence { F_n } we have

$F_n = F_{n-1} + F_{n-2}$, n is an integer greater than $1(n > 1)$.

$$\therefore \frac{F_n}{F_{n-1}} = 1 + \frac{F_{n-2}}{F_{n-1}}$$

Taking the limit as $n \to \infty$, we get

$$\lim_{n\to\infty} \frac{F_n}{F_{n-1}} = \lim_{n\to\infty}\left(1 + \frac{F_{n-2}}{F_{n-1}}\right)$$

or $\quad \lambda = 1 + \frac{1}{\lambda} \quad$ or $\quad \lambda^2 - \lambda + 1 = 0 \quad$ or $\quad \lambda = \frac{1+\sqrt{5}}{2}$.

Working procedure: Take two points in the initial uncertainty interval [a, b] which are symmetrically placed from both the end points a and b at a distance

$$L^*_2 = \lim_{n\to\infty} \frac{F_{n-2}}{F_n} L_0 = \lim_{n\to\infty} \frac{F_{n-2}}{F_{n-1}} \frac{F_{n-1}}{F_n} L_0 = \frac{L_0}{\lambda^2} = (0.618) L_0$$

$L_0 = b - a =$ length of initial uncertainty interval.

In K^{th} iteration take $L^*_K = (0.618)^{K-1} L_0$.

Repeat iterations till a stage that the interval of uncertainty is as small as desired. The optimum points can be taken as the middle points of this final uncertainty interval.

Examples

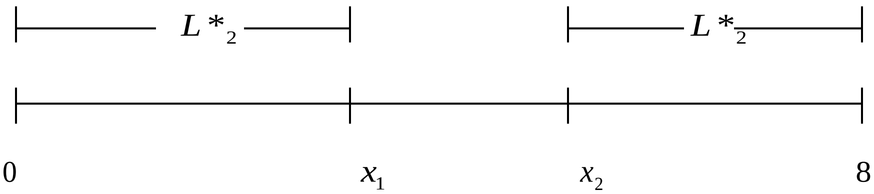
Example.3. Use golden section method to find maximum of $f(x) = x(5-x)$ given that $f(x)$ is an unimodal function is interval $[0, 8]$ in which the maximum lies.

Solution: Here $[a, b] = [0, 8]$ initial uncertainty interval and its length $L_0 = 8 - 0 = 8$.

First two experiments:

Let us take two points x_1 *and* x_2 between $[0,\ 8]$ at a distance

$L*_2 = (0.618)^2 L_0 = (0.618)^2 \times 8 = 3.055392$ from 0 and 8.



$$x_1 = 0 + L*_2 = 3.055392,\ f(x_1) = 5.941975$$

$$x_2 = 8 - L*_2 = 8 - 3.055392 = 4.944,\ f(x_2) = 0.276864$$

Thus we have $x_1 < x_2$ *and* $f(x_1) > f(x_2)$.

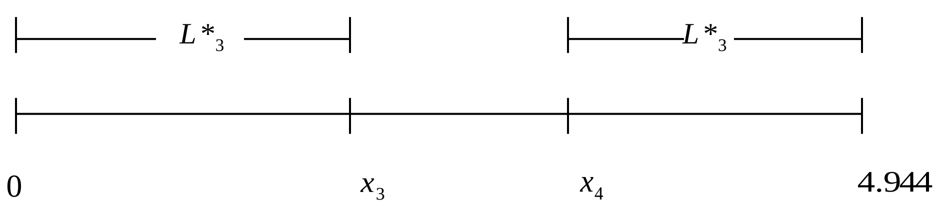
So by unimodality assumption, maximum points does not be on the right of x_2 and reject interval [4.944, 8].

The next uncertainty interval is [0, 4.944] with length $L_2 = (0.618)^{2-1} L_0 = 4.944$ (Approximate).

Third experiment:

Take two points x_3 and x_4 in $[0,\ 4.944]$ as a distance

$L*_3 = (0.618)^3 L_0 = (0.618)^3 \times 8 = 1.888$



$$x_3 = 0 + L*_3 = 1.888,\ f(x_3) = 5.875456$$

$$x_4 = 4.944 - L*_3 = 3.056,\ f(x_4) = 5.940864$$

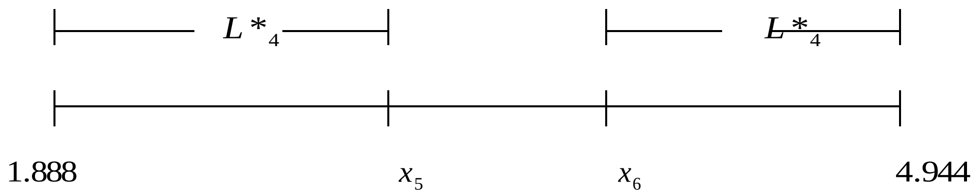
Thus $x_3 < x_4$ *and* $f(x_3) < f(x_4)$.

Hence reject interval [0, 1.888] and next uncertainty interval is [1.888, 4.944].

Fourth experiment:

Take two points x_5 *and* x_6 in [1.888, 4.944] at a distance

$L*_4 = (0.618)^4 L_0 = 1.1669$



$x_5 = 1.888 + L*_4 = 3.055, \; f(x_5) = 5.941975$

$x_6 = 4.944 - L*_4 = 3.777, \; f(x_6) = 4.6190$

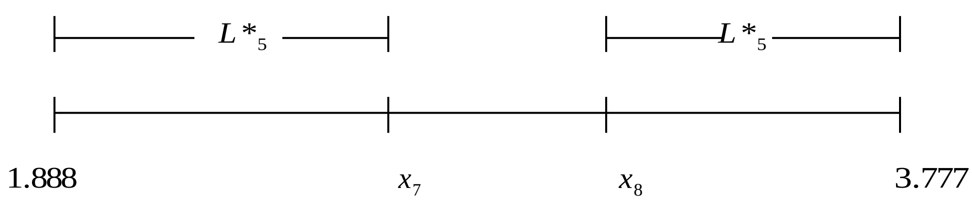
Thus $x_5 < x_6$ and $f(x_5) > f(x_6)$.

New uncertainty interval is [1.888, 3.777] rejecting interval on right of x_4

Fifth experiment:

Take two points x_7 and x_8 in [1.888, 3.777] as a distance

$L*_5 = (0.618)^5 L_0 = 0.721$ from the two end points



$x_7 = 1.888 + L*_5 = 2.609, \; f(x_7) = 6.238597$

$x_8 = 3.777 - L*_5 = 3.056, \; f(x_8) = 5.940864$

Thus we get $x_7 < x_8$ *and* $f(x_7) > f(x_8)$

By unimodality assumption, the next uncertainty interval is [1.888, 3.056].

After 5th experiment the uncertainty interval is sufficiently small (≈ 1.16) and the maximum points

$$x^* = \frac{1.888 + 3.056}{2} = 2.472 \quad \text{and} \quad f(x^*) = 6.249216.$$

4.6 Univariate Method

Univariate method as the name suggests is the method in which we move in axial directions covering all the directions one by one. Movement is made covering all directions, taking one direction at a time. This will complete one cycle of iterations. The process will be over when no further improvement, in the value of the given function to be minimized is possible in any direction of a cycle.

Procedure: Step-I: First we choose a fixed point X_1 (x_1, y_1) and a search direction $u_1 = (1, 0)^T$. Then take a step of size λ_1 in this direction and obtain optimum $\lambda_1 = \lambda_1^*$ for which $f(X_1+\lambda_1^* u_1)$ is minimum. This gives a new point X_2 where $X_2 = X_1+\lambda_1^* u_1$.

Step-II: Repeat step-I by taking optimum step length λ_2^* in the direction $u_2 = (0, 1)^T$ from the point X_2 to arrive at the next point $X_3 = X_2 + \lambda_2^* u_2$.

Step-III: Completion of step-I and II will form one cycle of iterations. Repeat the process for the completion of next cycle of iterations until the variation in the value of function f is negligible.

Here we shall consider functions in two variables x_1 and x_2 only, although the procedural method described below can be extended for functions of more than two variables also. We illustrate the method through examples:

Examples

Example.4. Minimize $f(X) = f(x_1, x_2) = 2x_1^2+3x_2^2-x_1 x_2$ using Univariate method taking

$X_1=\begin{pmatrix}1\\1\end{pmatrix}$ as the starting point.

Solution: Step-I: Consider unit vectors $u_1 = (1, 0)^T$ and $u_2 = (0, 1)^T$ in x_1 and x_2 (axial) directions respectively. Given $X_1 = \begin{pmatrix}1\\1\end{pmatrix}$, where $f(X_1) = 4$.

We take a step size λ_1 in x_1 direction to arrive

$$X_2 = X_1+\lambda_1\begin{pmatrix}1\\0\end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} \lambda_1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1+\lambda_1 \\ 1 \end{pmatrix}$$

To find optimum step size λ_1, we have

$$\text{Min } f\begin{pmatrix} 1+\lambda_1 \\ 1 \end{pmatrix} = 2(1+\lambda_1)^2 + 3 - (1+\lambda_1)$$

For minimum of *f*, we have

$$\frac{df}{d\lambda_1} = 4(1+\lambda_1) - 1 = 0$$

$$\Rightarrow \quad \lambda_1 = -\frac{3}{4}$$

and $$\frac{d^2 f}{d\lambda_1^2} = 4 > 0$$

$\Rightarrow$ f is minimum for $\lambda_1 = -\dfrac{3}{4}$.

Therefore new point

$$X_2 = \begin{pmatrix} 1+\lambda_1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/4 \\ 1 \end{pmatrix},$$

$$f(X_2) = \frac{1}{8} + 3 - \frac{1}{4} = \frac{23}{8} = 2.875.$$

Step-II: Now to find optimum step size λ_2 in x_2 direction.

Minimize $f\,[X_2+\lambda_2\,(0, 1)^T] = f\left(\begin{pmatrix} 1/4 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_2 \end{pmatrix}\right)$

$$= f\begin{pmatrix} 1/4 \\ 1+\lambda_2 \end{pmatrix}$$

$$= \frac{1}{8} + 3(1+\lambda_2)^2 - \frac{1}{4}(1+\lambda_2)$$

For minimum of *f*, we have

$$\frac{df}{d\lambda_2} = 6(1+\lambda_2) - \frac{1}{4} = 0$$

$$\Rightarrow \quad \lambda_2 = -\frac{23}{24}$$

and $$\frac{d^2 f}{d\lambda_2^2} = 6 > 0$$

$\Rightarrow$ f is minimum for $\lambda_2 = -\frac{23}{24}$.

New point $X_3 = X_2 + \lambda_2 (0, 1)^T$

$$= \begin{pmatrix} 1/4 \\ 1+\lambda_2 \end{pmatrix} = \begin{pmatrix} 1/4 \\ 1/24 \end{pmatrix},$$

$$f(X_3) = \frac{1}{8} + \frac{3}{(24)^2} - \frac{1}{96}$$

$$= \frac{69}{576}$$

$$= 0.11979.$$

Thus one cycle of iterations is completed.

Cycle-II:

Step-III: Now considering X_3 as the base point, to find optimum step size λ_3 in x_1-direction.

Minimize $f\,[X_3 + \lambda_3\,(1, 0)^T]$

$$= f\left(\begin{pmatrix} 1/4 \\ 1/24 \end{pmatrix} + \begin{pmatrix} \lambda_3 \\ 0 \end{pmatrix}\right)$$

$$= f\begin{pmatrix} 1/4 + \lambda_3 \\ 1/24 \end{pmatrix}$$

$$= 2\left(\frac{1}{4} + \lambda_3\right)^2 + 3\left(\frac{1}{24}\right)^2 - \frac{1}{24}\left(\frac{1}{4} + \lambda_3\right)$$

For minimum of f, we have

$$\frac{df}{d\lambda_3} = 4\left(\frac{1}{4} + \lambda_3\right) - \frac{1}{24} = 0$$

$$\Rightarrow \quad \lambda_3 = -\frac{23}{96}$$

$$\frac{d^2 f}{d\lambda_3^2} = 4 > 0$$

$\Rightarrow$ f is minimum for $\lambda_3 = -\frac{23}{96}$.

$\therefore$ New point X_4 = $\begin{pmatrix} 1/4 + (-23/96) \\ 1/24 \end{pmatrix}$

$$= \begin{pmatrix} 1/96 \\ 1/24 \end{pmatrix},$$

$$f(X_4) = 2\left(\frac{1}{96}\right)^2 + 3\left(\frac{1}{24}\right)^2 - \left(\frac{1}{96}\right)\left(\frac{1}{24}\right)$$

$$= \frac{23}{4608}$$

$$= 0.00499 .$$

Step-IV: Now considering X_4 as the base point, to find optimum step size λ_4 in x_2-direction.

Minimize $f\,[X_4 + \lambda_4\,(0, 1)^T]$

$$= f\left(\begin{pmatrix} 1/96 \\ 1/24 \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_4 \end{pmatrix}\right)$$

$$= f\begin{pmatrix} 1/96 \\ 1/24 + \lambda_4 \end{pmatrix}$$

$$= 2\left(\frac{1}{96}\right)^2 + 3\left(\frac{1}{24} + \lambda_4\right)^2 - \frac{1}{96}\left(\frac{1}{24} + \lambda_4\right)$$

For minimum of f, we have

$$\frac{df}{d\lambda_4} = 6\left(\frac{1}{24} + \lambda_4\right) - \frac{1}{96} = 0$$

$$\Rightarrow \quad \lambda_4 = -\frac{23}{576}$$

$$\frac{d^2 f}{d\lambda_4^2} = 6 > 0$$

$\Rightarrow$ f is minimum for $\lambda_4 = -\frac{23}{576}$.

$\therefore$ New point $X_5 = \begin{pmatrix} 1/96 \\ (1/24) - (23/576) \end{pmatrix}$

$$= \begin{pmatrix} 1/96 \\ 1/576 \end{pmatrix},$$

$$f(X_5) = 2\left(\frac{1}{96}\right)^2 + 3\left(\frac{1}{576}\right)^2 - \left(\frac{1}{96}\right)\left(\frac{1}{576}\right)$$

$$= \frac{33}{12 \times (96)^2}$$

$$= 0.00020797 \; .$$

Thus minimum f (X) = 0.00020797, when x_1=1/96 and x_2 = 1/576.

Note: From the above steps we see that as the base point moves from (1, 1) towards $\left(\frac{1}{96}, \frac{1}{576}\right) \approx$ (0, 0), f (X) is becoming smaller and smaller and is tending to zero.

Thus Min $f(X)$ = 0 at (0, 0).

4.7. Pattern Search Methods:

In Univariate method starting from a fixed point (the base point) search is made for the minimum of the given function by moving in the axial directions (the directions of axes) only. It can be observed that the rate of convergence in univariate method to arrive at the optimal point is very slow.

To accelerate the rate of convergence, a number of algorithms have been devised. Here we describe two of them:

(i) Hooke and Jeeve's method

(ii) Powell's method

4.8. Hooke and Jeeve's Method:

In this method we start from a fixed point say X_1. Step size in all the directions is taken to be constant (i.e. $\Delta x_1 = \Delta x_2 = \ldots\ldots\ldots = \Delta x_n$ = constant) and search is made in each direction u_i, u_i is the unit vector in the direction of x_i – axis whose i^{th} component is 1 and all other components are 0. Search is made first in positive and then if necessary in the negative direction of each axis to arrive at a temporary base point.

A temporary base point Y_{kj} obtained from X_k by perturbing the j^{th} component of X_k ($Y_{k0} = X_k$), defined as

$$Y_{k\,i} = \begin{cases} Y_{k,i-1} + \Delta x_i u_i & if\ f^{+} = f\left(Y_{k,i-1} + \Delta x_i u_i\right) < f = f\left(Y_{k,i-1}\right) \\ Y_{k,i-1} - \Delta x_i u_i & if\ f^{-} = f\left(Y_{k,i-1} - \Delta x_i u_i\right) < f = f\left(Y_{k,i-1}\right) \\ Y_{k,i-1} & if\ f\left(Y_{k,i-1}\right) = f \le min\left(f^{+}, f^{-}\right) \end{cases}$$

This process of determining new base point is continued for i=1, 2, 3, …, n until all directions are exhausted (covered).

After one cycle of covering all the axial directions reaching at the final temporary base point of the cycle say Y_{1n}, n being the number of design variables in the given problem, if Y_{1n} is different from $Y_{10}=X_1$, then a new base point is taken as $X_2 = Y_{1n}$ and a single step of optimum size λ^* is taken in the direction $S = X_2 - X_1$. This way the pattern of the preceding set of axial steps is repeated for further distance.

After this, search along the axes is again resumed and a second new pattern is generated to go further ahead. This is why the method is called pattern search method. These steps are repeated till a desired accuracy is achieved or the change in the value of the function satisfies the given condition in the problem.

We shall confine our discussion by considering functions in two variables (n = 2), which can be extended for more than two variables easily.

Examples

Example.5. Minimize $f(X) = 2x_1^2+x_2^2+2x_1x_2+x_1-x_2$ by Hooke's and Jeeve's method, taking starting base point $X_1 = (0, 0)$ and $\Delta x_1 = \Delta x_2 = 0.8$.

Solution: Step-I: Set $Y_{10}=X_1= (0, 0)$, $f(X_1) = 0$.

Moving in direction u_1, from base point X_1,

$$f^{+}(X_1+\Delta x_1 u_1) = f^{+}[(0, 0)+0.8(1, 0)]$$

$$= f^{+}(0.8, 0)$$

$$= 2.08 \nless f(X_1)$$

$$f^{-}(X_1-\Delta x_1 u_1) = f^{-}(-0.8, 0)$$

$= 0.48 \nless f(X_1)$

As f^+, f^- are not less than $f(X_1)$, so movement in u_1 direction is not beneficial and hence is discarded and $Y_{11} = X_1$

and $f(Y_{11}) = f(X_1) = 0$.

Step-II: Resuming movement in u_2-direction,

$$f^+(X_1+\Delta x_2 u_2) = f^+[(0, 0)+0.8(0, 1)]$$

$$= f^+(0, 0.8)$$

$$= -0.16 < f(X_1)$$

As $Y_{12} = X_1+\Delta x_2 u_2 = (0, 0.8)$ is different from X_1, and the new base point is taken as

$$Y_{12} = X_2$$

$$= (0, 0.8)$$

and $f(X_2) = -0.16$.

Step-III: As movement has been made in both the axial directions u_1 and u_2, the third movement is to be made in first pattern direction

$S_{p1}=X_2-X_1$

$= (0, 0.8) - (0, 0)$

$= (0, 0.8)$ through step length λ from X_2 so that $f(X_2+\lambda S_{p1})$ is minimum.

For minimum of $f[X_2+\lambda(X_2-X_1)] = f\{0, 0.8(1+\lambda)\}$

$$= [0.8(1+\lambda)]^2-0.8(1+\lambda),$$

$$\frac{df}{d\lambda} = 2\times 0.8(1+\lambda) - 0.8 = 0$$

$$\Rightarrow \quad \lambda = -\frac{3}{8}$$

and $$\frac{d^2 f}{d\lambda^2} = 1.6 > 0$$

$\Rightarrow$ f is minimum for $\lambda^* = -\dfrac{3}{8}$.

Thus we get the new base point

$$Y_{20} = X_3 = X_2+\lambda^*(X_2-X_1)$$

$$= (0, 0.8)-(3/8)(0, 0.8)$$

$$= (0, 0.5)$$

and $f(X_3) = f(0, 0.5) = -0.25.$

This completes the first cycle of iterations along both the axial directions u_1, u_2 and first pattern direction S_{p1}.

Second cycle of iterations:

Step-IV: First we move in u_1 direction starting with base point $X_3 = (0, 0.5)$.

$$f^+(X_3+\Delta x_1 u_1) = f^+[(0, 0.5)+(0.8, 0)]$$

$$= f^+(0.8, 0.5)$$

$$= 2.63 \nless f(X_3)$$

$$f^-(X_3-\Delta x_1 u_1) = f^-(-0.8, 0.5)$$

$$= -0.57 < f(X_3) =-0.25$$

Thus new temporary base point is

$$Y_{21} = X_3-\Delta x_1 u_1$$

$$= (-0.8, 0.5),$$

and $f(Y_{21}) = -0.57$

Step-V: Moving in u_2- direction from base point $Y_{21} = (-0.8, 0.5)$ to Y_{22}, we have

$$f^+(Y_{21}+\Delta x_2 u_2) = f^+(-0.8, 1.3)$$

$$=-1.21 < f(Y_{21})$$

Thus new base point

$$Y_{22} = X_4$$

$$= Y_{21} +\Delta x_2 u_2$$

$$= (-0.8, 1.3)$$

and $f(X_4) = -1.21.$

Step-VI: Now after exhausting both the axial directions in second cycle of iterations we move along the second pattern search direction $S_{p2} =X_4-X_3$ starting from X_4 through optimal step length λ so that

$f(X_4+\lambda S_{p2}) = f[(-0.8, 1.3)+\lambda(-0.8, 0.8)]$

$= f[-0.8(1+\lambda), 1.3+\lambda(0.8)]$ is minimum.

To find minimum of $f(\lambda)$, we have

$$\frac{df}{d\lambda} = 1.28\lambda - 0.32 = 0$$

$$\Rightarrow \quad \lambda = 0.25$$

and $$\frac{d^2 f}{d\lambda^2} = 1.28 > 0$$

$\Rightarrow$ f is minimum for $\lambda = 0.25$

Thus we get the point

$$X_5 = Y_{30}$$

$$= X_4+\lambda S_{p2}$$

$$= (-1, 1.5)$$

and $f(X_5) = -1.25.$

Third cycle of iterations:

Step-VII: Moving in u_1- direction from $X_5 = (-1, 1.5)$.

$f^+(X_5+\Delta x_1 u_1) = f^+[(-1, 1.5)+0.8(1, 0)]$

$= f^+(-0.2, 1.5)$

$= 0.03 \nless f(X_5)$

$f^-(X_5-\Delta x_1 u_1) = f^-[(-1, 1.5)-0.8(1, 0)]$

$= f^-(-1.8, 1.5)$

$= 0.03 \nless f(X_5)$,

So $Y_{31} = X_5$.

Step-VIII: Making movement in u_2- direction from $Y_{31} = X_5$,

$f^+(X_5+\Delta x_2 u_2) = f^+[(-1, 1.5)+0.8(0, 1)]$

$= f^+(-1, 2.3)$

$= -0.61 \nless f(X_5)$

$f^-(X_5-\Delta x_2 u_2) = f^-[(-1, 1.5)-0.8(0, 1)]$

$= f^-(-1, 0.7)$

$= -0.61 \nless f(X_5)$

This shows that movement from X_5 in any of the axial directions produces no change in the value of the function f at X_5. Thus X_5 is the optimal solution.

Therefore we get Min $f = -1.25$, when $x_1 = -1$ and $x_2 = 1.5$.

4.9 Powell's Method:

Let f be the function of n variables which is to be minimized and $u_1, u_2, \ldots, u_n$ be the axial directions. In Powell's method starting from a fixed point search is made sequentially in the directions $u_n; u_1, u_2, \ldots, u_n$ in the first cycle and along $S_{p1}, u_2, u_3, \ldots u_{n-1}, u_n, S_{p1}$ in the second cycle and $S_{p2}, u_3, u_4, \ldots u_{n-1}, u_n, S_{p1}, S_{p2}$ in the third cycle and so on until the minimum point is reached. Where S_{pi}'s are the pattern search directions defined as $S_{pi} = X_i - X_{i-n}$, n is the number of design (decision) variables.

In particular taking $n = 2$, Search is made in the directions u_2, u_1, u_2 and S_{p1}, u_2, S_{p1} and so on. Thus in the second cycle one of the axial directions is replaced by the pattern search direction. We illustrate the method through examples.

Note: In Powell's method any of u_1 or u_2 can be taken as first search direction and hence the cycle of iterations can also be in the order $u_1; u_2, u_1$ and $S_{p1}; u_1, S_{p1}$. Here u_2 is discarded in second cycle of iteration. Powell's method being the method of conjugate directions, it converges in at most two cycles of iterations.

Examples

Example.6. Minimize $f(x_1, x_2) = 2x_1^2 + x_2^2 + 2x_1x_2 + x_1 - x_2$ using Powell's method taking $X_1 = (0, 0)$ as the starting point.

Solution: Given that $f = 2x_1^2 + x_2^2 + 2x_1x_2 + x_1 - x_2$ and base point $X_1 = (0, 0)$.

Cycle I: In first cycle of iteration, search will be made in the directions u_2, u_1, u_2.

Taking step of size λ_1 in the direction S_2 from X_1 to reach at X_2 such that

$f(X_2) = f(X_1+\lambda_1 u_2) = \lambda_1^2 - \lambda_1$ is minimum.

For f is minimum, we have

$$\frac{df}{d\lambda_1} = 2\lambda_1 - 1 = 0$$

$$\Rightarrow \quad \lambda_1 = \frac{1}{2}$$

and $$\frac{d^2 f}{d\lambda_1^2} = 2 > 0$$

$\Rightarrow$ f is minimum.

So we have

$$X_2 = X_1 + \lambda_1 u_2$$

$$= (0, 0) + (1/2)(0, 1)$$

$$= (0, 1/2),$$

and $f(X_2) = f(0, 1/2)$

$= -0.25 < f(X_1)$.

Next moving from X_2 in the direction of $u_1 = (1, 0)$ through step length λ_2 arriving at point

$$X_3 = X_2 + \lambda_2 u_1.$$

We find λ_2 so that

$$f(X_3) = f(\lambda_2, 1/2)$$

$$= 2\lambda_2^2 + 2\lambda_2 - 0.25 \text{ is minimum.}$$

For f minimum, we have

$$\frac{df}{d\lambda_2} = 4\lambda_2 + 2 = 0$$

$$\Rightarrow \quad \lambda_2 = -\frac{1}{2}$$

and $$\frac{d^2 f}{d\lambda_2^2} = 4 > 0$$

$\Rightarrow$ f is minimum.

So we have

$$X_3 = X_2 + \lambda_2 u_1$$

$= (0, 1/2) – (1/2).(1, 0)$

$= (-1/2, 1/2)$

and $f(X_3) = -0.75 < f(X_2) = -0.25.$

Now from X_3, move to X_4 taking step length of size λ_3 in the direction of $u_2 = (0, 1)$.

$X_4 = X_3 + \lambda_3 (0, 1)$

$= (-1/2, 1/2) + (0, \lambda_3)$

$= (-1/2, 1/2+\lambda_3).$

Now we determine λ_3 such that

$f(X_4) = f(-1/2, 1/2+\lambda_3) = \lambda_3^2 - \lambda_3 - 0.75$ is minimum.

For minimum f, we have

$$\frac{df}{d\lambda_3} = 2\lambda_3 - 1 = 0$$

$$\Rightarrow \quad \lambda_3 = \frac{1}{2}$$

and $$\frac{d^2 f}{d\lambda_3^2} = 2 > 0$$

$\Rightarrow$ f is minimum.

Thus we have $X_4 = (-1/2, 1)$

and $f(X_4) = -1 < f(X_3) = -0.75.$

This completes one cycle of iterations.

Second cycle: We generate the pattern search direction S_{p1} as

$S_{p1} = X_4 – X_2$

$= (-1/2, 1) – (0, 1/2)$

$= (-1/2, 1/2)$

Now from X_4, we move to X_5 by taking step length λ_4 in the direction of $S_{p1} = (-1/2, 1/2)$.

$X_5 = X_4 + \lambda_4 S_{p1}$

$= (-1/2, 1) + \lambda_4 (-1/2, 1.2)$

$= [-(1/2).(1+\lambda_4), 1+(1/2)\lambda_4]$

We now determine λ_4 such that

$f(X_5) = 0.25\lambda_4^2 - 0.5\lambda_4 - 1$ is minimum.

For minimum of *f*, we have

$$\frac{df}{d\lambda_4} = 0.5\lambda_4 - 0.5 = 0$$

$$\Rightarrow \qquad \lambda_4 = 1$$

and $$\frac{d^2 f}{d\lambda_4^2} = 0.5 > 0$$

$\Rightarrow$ *f is* minimum, where $\lambda_4 = 1$.

$\therefore \qquad X_5 = X_4 + \lambda_4 \, S_{p1}$

$= (-1/2, 1) + (-1/2)(1/2, 1/2)$

$= (-1, 3/2)$

and $f(X_5) = -1.25 < f(X_4)$.

Now from X_5 we move to X_6 by taking step length λ_5 in the direction of u_2 (0, 1).

$X_6 = X_5 + \lambda_5 u_2$

$= (-1, 3/2) + \lambda_5 (0, 1)$

$= (-1, 3/2 + \lambda_5)$.

We determine λ_5 such that

$f(X_6) = f(-1, 3/2 + \lambda_5) = \lambda_5^2 - (5/4)$ is minimum.

For minimum of *f*, we have

$$\frac{df}{d\lambda_5} = 2\lambda_5 = 0$$

$$\Rightarrow \qquad \lambda_5 = 0.$$

$\therefore X_6 = (-1, 3/2) = X_5$.

Which ($\lambda_5 = 0$) shows that *f* cannot be minimized in the direction of u_2 and there is no other direction to move.

Thus we get min $f = -1.25$, when $x_1 = -1$, $x_2 = 3/2$.

4.10. Steepest Descent (Cauchy's) Method:

This method is also called "Cauchy's method". In this method like other methods we start from an initial point called the base point and movement is made in the direction of steepest direction. Indirect search methods use the derivatives along with finding the value of function at the search point. These methods, therefore, are also called "Gradient Methods".

Procedure Method:

Step-1: Let X_1 be the starting point. Let us consider $f(x_1, x_2)$ be the function to be minimized. A function decreases most rapidly in the negative direction of gradient. Calculate $f(X_1)$ and $S_1 = -\nabla f(X_1)$ (*i.e.*, negative of gradient f at X_1).

Step-2: Find the optimum step length λ_1 in this direction to arrive at the point $X_2 = X_1 + S_1 \lambda_1$ so that $f(X_2) < f(X_1)$. Proceed this way till one of the following conditions is satisfied and terminate the process

(i) $\left|\dfrac{\partial f}{\partial x_i}\right| < \in$, $\in$ is a small positive number.

(ii) $\left|X_{i+1} - X_i\right| < \in$, i.e., change in design vectors in the consecutive iterations is small.

(iii) Relative change in the value of f at two consecutive steps is small *i.e.,*

$$\left|\frac{f(X_{i+1}) - f(X_i)}{f(X_i)}\right| < \in.$$

Examples

Example.7: **Minimize $f(x_1, x_2) = 2x_1^2 + x_2^2 + 2x_1x_2 + x_1 - x_2$, starting from the point $X_1 = (0, 0)$ and using Steepest descent method, taking** $\in = 0.01$.

Solution: Given that $f(x_1, x_2) = 2x_1^2 + x_2^2 + 2x_1x_2 + x_1 - x_2$

$$\therefore \quad \nabla f(X) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right)$$

$$= (4x_1 + 2x_2 + 1, 2x_2 + 2x_1 - 1)$$

Initial step: Given $X_1 = (0, 0)$ and hence $-\nabla f(X_1) = (-1, 1)$.

Iteration–1: We find λ_1 which minimizes

$$f[X_1-\lambda_1\nabla f(X_1)] = f[(0, 0)+\lambda_1(-1, 1)]$$

$$= f(-\lambda_1, \lambda_1)$$

$$= 2\lambda_1^2+\lambda_1^2-2\lambda_1^2-\lambda_1-\lambda_1$$

$$= \lambda_1^2-2\lambda_1$$

For minimum of f, we have

$$\frac{\partial f}{\partial \lambda_1} = 2\lambda_1-2 = 0$$

$\Rightarrow$ $\lambda_1=1$

and $$\frac{\partial^2 f}{\partial \lambda_1^2} = 2 > 0 \text{ (minimum)}$$

Thus new point $X_2 = X_1-\lambda_1 \nabla f(X_1) = (-1, 1)$.

$\nabla f(X_2) = (-1, -1)$.

$|\nabla f(X_2)| \nless \in$

$\Rightarrow$ X_2 is not optimal point.

Iteration-2: Now we find λ_2 which minimizes

$$f[X_3] = f[X_2-\lambda_2\nabla f(X_2)]$$

$$= f[(-1, 1)-\lambda_2(-1, -1)]$$

$$= f(-1+\lambda_2, 1+\lambda_2)$$

$$\text{or } f(X_3) = 5\lambda_2^2-2\lambda_2-1$$

Now we find λ_2 which minimizes $f(X_3)$ as a function of λ_2.

For minimum of f, we have

$$\frac{\partial f}{\partial \lambda_2} =10\lambda_2-2 = 0$$

$\Rightarrow \qquad \lambda_2 = 1/5$

and $\qquad \dfrac{\partial^2 f}{\partial \lambda_2^2} = 10 > 0$ (minimum)

$X_3 = X_2 - \lambda_2 \nabla f(X_2)$

$= (-1+(1/5), 1+(1/5))$

$= (-0.8, 1.2)$.

$\nabla f(X_3) = (0.2, -0.2)$

and $\quad |\nabla f(X_3)| \nless 0.01$

$\Rightarrow \quad X_3$ is not optimal point.

Iteration 3: Next we find λ_3 which minimizes

$f[X_4] = f[X_3 - \lambda_3 \nabla f(X_3)]$

$= f[(-0.8, 1.2) - \lambda_3(0.2, -0.2)]$

$= f(-0.8 - 0.2\lambda_3, 1.2 + 0.2\lambda_3)$

or $\quad f(X_4) = 0.04\lambda_3^2 - 0.08\lambda_3 - 1.2$

Now we find λ_3 which minimizes $f(X_3)$ as a function of λ_3.

For minimum of f, we have

$$\frac{\partial f}{\partial \lambda_3} = 0.08\lambda_3 - 0.08 = 0$$

$\Rightarrow \qquad \lambda_3 = 1$

and $\qquad \dfrac{\partial^2 f}{\partial \lambda_3^2} = 0.08 > 0$ (minimum)

$X_4 = X_3 - \lambda_3 \nabla f(X_3)$

$= (-1, 1.4)$

and $\quad f(X_4) = -1.24$

$\nabla f(X_4) = (-0.2, -0.2)$

And $|\nabla f(X_4)| \nless 0.01$

$\Rightarrow$ X_4 is not an optimal point.

Iteration 4: Further we find λ_4 such that $f(X_5) = f[(X_4 - \lambda_4 \nabla f(X_4)]$ is minimum.

For $f(X_5) = f[X_4 - \lambda_4 \nabla f(X_4)]$

$= 0.2\lambda_4^2 - .08\lambda_4 - 1.24$ minimum,

For minimum of f, we have

$$\frac{\partial f}{\partial \lambda_4} = 0.4\lambda_4 - 0.08 = 0$$

$\Rightarrow$ $\lambda_4 = 0.2$

and $\frac{\partial^2 f}{\partial \lambda_4^2} = 0.4 > 0$ (minimum)

$\therefore$ $X_5 = X_4 - \lambda_4 \nabla f(X_4)$

$= (-1, 1.4) - 0.2(-0.2, -0.2)$

$= (-0.96, 1.44)$

and $\nabla f(X_4) = (0.04, -0.04)$

and $|\nabla f(X_5)| \nless 0.01$

$\Rightarrow$ X_5 is not an optimal point.

Iteration 5: Next to minimize

$f(X_6) = f(X_5 - \lambda_5 \nabla f(X_5))$

$= .0016\lambda_5^2 - 0.0032\lambda_5 - 1.248$

For minimum of f, we have

$$\frac{\partial f}{\partial \lambda_5} = 0.0032\lambda_5 - 0.0032 = 0$$

$\Rightarrow$ $\lambda_5 = 1$

and $\frac{\partial^2 f}{\partial \lambda_5^2} = 0.0032 > 0$ and hence minima.

$X_6 = X_5 - \lambda_5 \nabla f(X_5)$

$= (-1, 1.48)$

and $\nabla f(X_6) = (-0.04, 0.04)$

$|\nabla f(X_6)| \nless 0.01$

$\Rightarrow$ X_6 is not an optimal point.

Iteration 6: Next to minimize

$$f(X_7) = f(X_6) - \lambda_6 \nabla f(X_6)$$

$$= .0016\,\lambda_6^2 - 1.2496$$

For minimum of f, we have

$$\frac{\partial f}{\partial \lambda_6} = 0.0032\lambda_6 = 0$$

$\Rightarrow$ $\lambda_6 = 0$

and $\frac{\partial^2 f}{\partial \lambda_6^2} = 0.0032 > 0$

$\Rightarrow$ $f(X_7)$ is minimum.

With $X_7 = f(X_6) - \lambda_6 \nabla f(X_6)$

$= (-1, 1.48)$

and $\lambda_6 = 0$

$\Rightarrow$ further improvement in f is not possible alternatively $X_6 = X_7 = (-1, 1.48)$

$\Rightarrow$ $|X_7 - X_6| = (0, 0) < \in = (.01, .01)$.

Thus we get

Min $f = f(X_6) = f(X_7) = -1.2496 \approx -1.25$,

when $x_1 = -1$ and $x_2 = 1.48$.

4.11 Summary

In direct search methods, derivatives of the function to be minimized are not needed, these methods are suitable for functions which are not differentiable.

A function $f(x)$ with a unique optimal value (either a unique maxima or minima) in $[a,b]$ is called unimodal function.

In Fibonacci method the initial interval of uncertainty say $[a,b]$ is given. Also the function to the optimized to a given degree of accuracy must be unimodal in the interval of uncertainty.

Golden section method differs from Fibonacci method in that the total number of experiments to be conducted in golden section method are unlimited.

In Univariate method, we move in axial directions covering all directions, taking one direction at a time. This will complete one cycle of iterations. The process will be over when no further improvement in the value of the given function to be minimized is possible in any direction of a cycle. In this method the rate of convergence is slow.

Pattern search method is used to accelerate the rate of convergence. Powell's method being the method of conjugate directions, it converges in at most two cycles of iterations.

Indirect search methods are used for the functions having derivatives along with finding the value of function at the search point. These methods are also called Gradient Methods.

In Steepest decent method, we start from an initial point called the base point and movement is made in the direction of steepest direction. This method is also called Cauchy's method.

4.12 Terminal Questions

Q.1. Use Fibonacci method to find maximum of $f(x) = x(5-x)$ given that $f(x)$ is an unimodal function is interval $[0, 8]$ in which the maximum lies.

Q.2. Minimize the function $f(x) = 0.65 - \frac{0.75}{1+x^2} - 0.65x\,tan^{-1}(1/x)$ by golden section method in interval $[0,8]$ with $n = 6$.

Q.3. Minimize $f(X) = f(x_1, x_2) = 2x_1^2+x_2^2+x_1-x_2+2x_1x_2$ using Univariate method taking $X_1=\begin{pmatrix}0\\0\end{pmatrix}$ as the starting point (base point).

Q.4. Minimize $f(X) = 3x_1^2+x_2^2-2x_1x_2-4x_1-3x_2$ by Hooke's and Jeeve's method, taking starting base point as $X_1 = (0, 0)$ and $\Delta x_1 = \Delta x_2 = 1$.

Q.5. Minimize $f(x_1, x_2) = 4x_1^2 + 3x_2^2 - 5x_1x_2 - 8x_1$ using Powell's method taking $X_1 = (0, 0)$ as the starting point.

Q.6. Minimize $f(x_1, x_2) = 2x_1^2 + x_2^2$ starting from the point $X_1 = (1, 2)$ and using Steepest descent method solve up to two iterations.

Answer

1. $x = 2.66667$, $f(x) = 6.22222$.

2. $x = 0.4055$, $f(x) = -0.30658$.

3. $x_1 = -1$, $x_2 = 1.5$, $f(X) = -1.25$.

4. $x_1 = 1.75$, $x_2 = 3.2.5$, $f(X) = -8.375$.

5. $x_1 = 48/23$, $x_2 = 40/23$.

6. $f(X) = 1.0336$.

UNIT-5: Non-Linear Programming-II

Structure

5.1 Introduction

In unit five, we will explore constrained non-linear optimization problems. A diverse range of non-linear programming problems exists, categorized based on the characteristics of the objective function $f(X)$ and the constraints $g_j(X)$. This unit will delve into search methods designed to address constrained non-linear programming problems. Initially, we will examine search techniques in one dimension, followed by an exploration in two dimensions. The methods employed to tackle both unconstrained and constrained problems are further classified into the following two categories.

1. Direct Search Methods

2. Decent Methods

5.2 Objectives

After reading this unit the learner should be able to understand about the constrained problems and their solution method:

- Complex Method
- Zoutendijk Method (Method of Feasible Directions)
- Indirect method
- Transform Method
- Penalty function and Penalty Methods:

  (i) Interior Penalty Function Method

  (ii) Exterior Penalty Function Method

5.3 Complex Method

Complex method deals with the constrained optimization problems of the type

Min Z = f(X) …… (5.1)

Subject to $g_i(X) \le 0$, $j = 1, 2, 3, \ldots, m$ …… (5.2)

$X = (x_1, x_2, \ldots, x_n)^T$ and $x_i^{(l)} \le x_i \le x_i^{(u)}$, $i = 1, 2, \ldots, n$ …… (5.3)

$x_i^{(l)}$ = lower bound on x_i and $x_i^{(u)}$ = upper bound on x_i.

Conditions in (5.3) are called side constraints. For a minimizations problem in n variables if we consider k points where $k \ge n+1$ then the figure formed on joining them is known a complex. We shall consider the minimization problem in two variables and take $k = 2n = 4$. These four points will form the vertices of the complex.

Working procedure for the case of two variables x_1 and x_2:

Step-I: In complex method one point X_1 is given and remaining $(2n-1) = 3$ ($n = 2$ here) points, X_2, X_3 and X_4 are obtained one at a time by using random members $r_{i,j}$, $0 < r_{i,j} < 1$. Calculate

$x_{i,j} = x_i^l + r_{i,j}\,[x_i^{(u)} - x_i^{(l)}]$, $i = 1, 2$ and $j = 2, 3, 4$

$x_{i,j}$ = i^{th} component of the point X_j. It is worth noticing that point X_2, X_3 and X_4 thus generated satisfy side constraint (5.3) but may not satisfy all constraint in (5.2). In case a point say X_4 is not satisfying all constraints in (5.2), then a new point denoted as $X_4^{(1)}$ is obtained by moving X_4 half way towards (in the direction of) the centroid $X_0 = (X_1+X_2+X_3)/3$ of the remaining points X_1, X_2, X_3 i.e., $X_4^{(1)} = (X_0+X_j)/2$. We check if new point $X_4^{(1)}$ satisfies all constraints in (5.2) or not. If not, we further get a new point $X_4^{(2)}$ by moving $X_4^{(1)}$ half way towards the centroid X_0 given above. We proceed this way until a feasible point

X_4 satisfying (5.2) is obtained. Thus we get four feasible points X_1, X_2, X_3 and X_4 all satisfying conditions (5.2) and (5.3) and they form the vertices of starting complex.

Step-II: Calculate $f(X_1)$, $f(X_2)$, $f(X_3)$ and $f(X_4)$ and mark the largest and the least value obtained. Let $f(X)$ be largest at $X_n = X_4$ say and smallest at X_l then the process of reflection is used to determine a new point X_r as $X_r = (1+\alpha)X_0 - \alpha X_h$ where $\alpha \geq R$ and X_0 is the centroid of all vertices except X_h that is $X_0 = (X_1+X_2+X_3)/3$. Now check the feasibility of X_r

(i) If X_r is feasible and $f(X_r) < f(X_h)$, then X_h is replaced by X_r and move on step-II.

(ii) If $f(X_r) \geq f(X_h)$, a new trial point X_r is found by taking new α = old $\alpha/2$ and is tested for further satisfaction of the condition $f(X_r) < f(X_h)$. We continue the process until the condition $f(X_r) < f(X_h)$ is satisfied and this way value of α will become smaller and smaller.

(iii) If X_r is not obtained in any manner such that $f(X_r) < f(X_h)$ then we neglect the whole reflection process and new reflection process is started by taking X_h which gives the second largest value of the function.

Now for the convergence, the procedure ends when distance between any two vertices among X_1, X_2, X_3, X_4 becomes smaller than the prescribed value of $\in$.

Examples

Example.1. Minimize $f(X) = f(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 1.5)^2 - 0.25$,

Subject to $x_1 + x_2 \leq 4$, $0 \leq x_1 \leq 2$ and $1 \leq x_2 \leq 3$ by complex method with $X_1 = \begin{pmatrix} 0.7 \\ 1.1 \end{pmatrix}$.

Solution: Here constraint is $g(X) = x_1 + x_2 - 4 \leq 0$ and side constraints are $0 \leq x_1 \leq 2$ & $1 \leq x_2 \leq 3$.

Consider 4 vertices X_1, X_2, X_3 and X_4 of the complex with $X_1 = \begin{pmatrix} 0.7 \\ 1.1 \end{pmatrix}$

Let us choose the random numbers

$$r_{1,2} = 0.4,\ r_{1,3} = 0.6,\ r_{1,4} = 0.8$$

$$r_{2,2} = 0.5,\ r_{2,3} = 0.7,\ r_{2,4} = 0.9$$

to find the remaining three vertices X_2, X_3 and X_4 using the formula

$$x_{i,j} = x_i^{(l)} + r_{i,j}[x_i^{(u)} - x_i^{(l)}],\ i = 1, 2;\ j = 1, 2, 3, 4.$$

Where $x_{i,j}$ is the i^{th} component of vertex X_j.

Here $x_1^{(l)} = 0$, $x_1^{(u)} = 2$ and $x_2^{(l)} = 1$, $x_2^{(u)} = 3$.

Therefore $x_{1,2} = x_1^{(l)} + r_{1,2}\,[x_1^{(u)} - x_1^{(l)}] = 0.8$

$x_{1,3} = x_1^{(l)} + r_{1,3}\,[x_1^{(u)} - x_1^{(l)}] = 1.2$

$x_{1,4} = x_1^{(l)} + r_{1,4}\,[x_1^{(u)} - x_1^{(l)}] = 1.6$

$x_{2,2} = x_2^{(l)} + r_{2,2}\,[x_2^{(u)} - x_2^{(l)}] = 2$

$x_{2,3} = x_2^{(l)} + r_{2,3}\,[x_2^{(u)} - x_2^{(l)}] = 2.4$

$x_{2,4} = x_2^{(l)} + r_{2,4}\,[x_2^{(u)} - x_2^{(l)}] = 2.8$

Thus the first simplex consist of the vertices

$$X_1 = \begin{pmatrix} 0.7 \\ 1.1 \end{pmatrix}, X_2 = \begin{pmatrix} 0.8 \\ 2 \end{pmatrix}, X_3 = \begin{pmatrix} 1.2 \\ 2.4 \end{pmatrix} \quad \text{and} \quad X_4 = \begin{pmatrix} 1.6 \\ 2.8 \end{pmatrix},$$

$\therefore\ g(X_1) = x_1 + x_2 = 0.7 + 1.1 = 1.8 \le 4$

Similarly $g(X_2) = 2.8 \le 4$, $g(X_3) = 3.6 \le 4$ and $g(X_4) = 4.8 > 4$.

$\therefore g(x)$ is not satisfied at vertex X_4 and hence it is replaced by some point in the feasible region.

X_0, the centroid of satisfying vertices is

$$X_0 = \frac{X_1 + X_2 + X_3}{3} = \begin{pmatrix} 0.9 \\ 1.83 \end{pmatrix}$$

Thus new X_4 vertex is

$$X_4 = X_4^{(1)} = \frac{X_0 + X_4}{2} = \begin{pmatrix} 1.25 \\ 2.315 \end{pmatrix}$$

and $g(X_4^{(1)}) = 3.565 \le 4.$

Thus $X_4^{(1)}$ lies in the feasible region and the initial complex has the vertices X_1, X_2, X_3 and $X_4 = X_4^{(1)}$ with

$$X_1 = \begin{pmatrix} 0.7 \\ 1.1 \end{pmatrix}, X_2 = \begin{pmatrix} 0.8 \\ 2 \end{pmatrix}, X_3 = \begin{pmatrix} 1.2 \\ 2.4 \end{pmatrix} \text{ and } X_4 = X_4^{(1)} = \begin{pmatrix} 1.25 \\ 2.315 \end{pmatrix},$$

and $f(X_1) = 0$, $f(X_2) = 0.04$, $f(X_3) = 0.60$ and $f(X_4^{(1)}) = 0.4\,767$..

$f(X_3) = 0.60$ gives the maximum value and $f(X_1) = 0$ is the minimum value

So we take $X_3 = X_h$ with $f(X_h) = 0\,.6$ and $X_1 = X_l$, $f(X_l) = 0$

The new centroid X_0 is obtained as

$$X_0 = \frac{X_1 + X_2 + X_4^{(1)}}{3} = \begin{pmatrix} 0.917 \\ 1.805 \end{pmatrix} \quad \text{and} \quad f(X_0) = -0.15.$$

$f(X_0) < f(X_h)$

$\Rightarrow$ $f(X)$ is decreasing from $X_h(=X_3)$ toward X_0. So let us find X_r using reflection process as $X_r = (1+\alpha)X_0 - \alpha X_h$

Taking $\alpha=1$, $X_r = 2X_0 - X_h \begin{pmatrix} 0.634 \\ 1.21 \end{pmatrix}$

and $f(X_r) = -0.034944$.

As X_r feasible and $f(X_r) < f(X_h)$. So we proceed further by considering the different values of α as

(i) Taking $\alpha=0.1$ then we have

$X_r^{(1)} = 1.1\, X_0 - 0.1\, X_h$

$$= (1.1) \begin{pmatrix} 0.917 \\ 1.805 \end{pmatrix} - 0.1 \begin{pmatrix} 1.2 \\ 2.4 \end{pmatrix}$$

$$= \begin{pmatrix} 0.8887 \\ 1.7455 \end{pmatrix}$$

and $f(X_r^{(1)}) = -0.1773$.

(ii) Taking $\alpha=0.2$ then we have

$X_r^{(2)} = 1.2\, X_0 - 0.2\, X_h$

$$= (1.2) \begin{pmatrix} 0.917 \\ 1.805 \end{pmatrix} - 0.2 \begin{pmatrix} 1.2 \\ 2.4 \end{pmatrix}$$

$$= \begin{pmatrix} 0.8604 \\ 1.686 \end{pmatrix}$$

and $f(X_r^{(2)}) = -0.1959$.

(iii) Taking $\alpha=0.3$ then we have

$X_r^{(3)} = 1.3\, X_0 - 0.3\, X_h$

$$= (1.3) \begin{pmatrix} 0.917 \\ 1.805 \end{pmatrix} - 0.3 \begin{pmatrix} 1.2 \\ 2.4 \end{pmatrix} = \begin{pmatrix} 0.8321 \\ 1.6265 \end{pmatrix}$$

and $f(X_r^{(3)}) = -0.2058$.

(iv) Taking $\alpha = 0.4$ then we have

$X_r^{(4)} = 1.4\, X_0 - 0.4\, X_h$

$$= (1.4)\begin{pmatrix} 0.917 \\ 1.805 \end{pmatrix} - 0.4\begin{pmatrix} 1.2 \\ 2.4 \end{pmatrix} = \begin{pmatrix} 0.8038 \\ 1.567 \end{pmatrix}$$

and $f(X_r^{(4)}) = -0.2070$.

(v) Taking $\alpha=0.5$ then we have

$X_r^{(5)} = 1.5\, X_0 - 0.5\, X_h$

$$= (1.5)\begin{pmatrix} 0.917 \\ 1.805 \end{pmatrix} - 0.5\begin{pmatrix} 1.2 \\ 2.4 \end{pmatrix} = \begin{pmatrix} 0.7755 \\ 1.5075 \end{pmatrix}$$

and $f(X_r^{(5)}) = -0.1995$.

Since the decreasing of the values of f continues up to $X_r^{(4)}$, so let us replace the vertex $X_h=X_3$ with highest values by $X_r^{(4)}$ to get the new complex with vertices

$$X_1 = \begin{pmatrix} 0.7 \\ 1.1 \end{pmatrix} \quad \text{and} \quad f(X_1) = 0$$

$$X_2 = \begin{pmatrix} 0.8 \\ 1.1 \end{pmatrix} \quad \text{and} \quad f(X_2) = 0.04$$

$$X_3 = X_r^{(4)} = \begin{pmatrix} 0.8038 \\ 1.567 \end{pmatrix} \quad \text{with } f(X_3) = f(X_r^{(4)}) = -0.207016$$

and $$X_4 = \begin{pmatrix} 1.25 \\ 2.315 \end{pmatrix} \quad \text{with } f(X_4) = f(X_4^{(1)}) = -0.476725$$

Thus $f(X_4) = f(X_4^{(1)})$ gives the maximum values and $f(X_3) = f(X_r^{(4)})$ gives the minimum value.

So $X_h = X_4 = \begin{pmatrix} 1.25 \\ 2.315 \end{pmatrix}$ with $f(X_4) = -0.476725$

and $X_l = X_3 = X_r^{(4)} = \begin{pmatrix} 0.8038 \\ 1.567 \end{pmatrix}$ with $f(X_3) = f(X_l) = -0.207016$

The centroid $X_0 = (X_1 + X_2 + X_3)/3 = \begin{pmatrix} 0.7679 \\ 1.5557 \end{pmatrix}$ with $f(X_0) = -0.19303$

$|f(X_l) - f(X_0)| = 0.014$.

Thus if the desired accuracy is $\in = 0.01$, then the above solution is accepted and therefore

Min $f(X) = -0.2070$ at X_3 *i.e.*, when $x_1=0.8038$ and $x_2=1.5557$.

5.4 Zoutendijk Method (Method of Feasible Direction)

Consider the problem

$$\text{Min } f(X) \qquad \text{....(5.4)}$$

Subject to $\quad g_j(X) \le 0$, j = 1, 2, ……….., m $\qquad$ ….(5.5)

$X = (x_1, x_2, \ldots\ldots, x_m)^T$

In this method of feasible direction, we have choose the starting point X_i (i = 1 for starting point) which satisfy all the constraints in (5.5) and move to a better point X_{i+1} (the point where value of the function in (5.4) is lesser than that it has at X_i) as per the iterative formula given as

$$X_{i+1} = X_i + \lambda_i S_i$$

Where X_i is the Starting point for $_i{}^{th}$ iteration.

S_i is the direction in which to move.

λ_i = length of the step to be taken in the direction of S_i.

X_{i+1} is the point determined at the end of $_i{}^{th}$ iteration

Care should be taken to choose λ_i such that new point X_{i+1} is in feasible region. Further the search direction is found such that

(i) Even a very small movement in that direction violets none of the constraints in (5.5) and

(ii) The value of the function $f(X)$ in (5.4) decreases.

Feasible Direction:

A direction S which satisfies condition (i) is called feasible direction. One way to determine the feasible direction is: if at X_i,

$$\left\{ \frac{d}{d\lambda} \left[g_j(X_i + \lambda_i S_i) \right] \right\}_{\lambda_i = 0} = S_i^T \nabla g_j(X_i) \le 0$$

then direction S_i is feasible direction.

Usable Feasible Direction:

A direction S_i which satisfies both the conditions given in (i) and (ii) is known as usable feasible direction. Further a direction S_i will be a usable feasible direction if

(a) $\left\{\dfrac{d}{d\lambda}\left[f\left(X_i+\lambda_i S_i\right)\right]\right\}_{\lambda_i=0} = S_i^T \nabla f(X_i) < 0$ …..(5.6)

(b) $\left\{\dfrac{d}{d\lambda}\left[g_j\left(X_i+\lambda_i S_i\right)\right]\right\}_{\lambda_i=0} = S_i^T \nabla g_j(X_i) \le 0$ …..(5.7)

We can reduce the value of f by taking step lengths in such a direction S_i as described above. The process terminates at X_i if $\nabla f(X_i) = 0$. Zoutendijk method is a method of feasible directions.

Examples

Example.2. Using the method of feasible directions due to Zoutendijk method

Min $f(X) = x_1^2+x_2^2-2x_1-3x_2+3$

Subject to $g_1(X) = x_1+x_2 \le 4$ taking the initial point $X_1 = \begin{pmatrix}0\\0\end{pmatrix}$.

Solution: Given that

$$X_1 = \begin{pmatrix}0\\0\end{pmatrix},$$

$$f(X) = x_1^2+x_2^2-2x_1-3x_2+3,\ x_1+x_2 \le 4$$

Thus we have

$$g_1(X) = x_1+x_2 - 4.$$

Then we have

$$f(X_1) = 3 \text{ and } g(X_1) = -4 < 0.$$

Step-I: Here $g(X_1) < 0$ so search direction S_1 is

$$S_1 = -\nabla f(X_1) = -\begin{Bmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \end{Bmatrix}_{X_1=(0,\,0)} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

After normalization we have

$$S_1 = \begin{pmatrix} 2/3 \\ 1 \end{pmatrix}.$$

Step-II: To obtain a new point X_2 we take a step of length λ_1 in the direction of $-\nabla f(X_1)$ to arrive at $X_2 = X_1+\lambda_1 S_1 = [(2/3)\lambda_1, \lambda_1]$

$$f(X_2) = f[(2/3)\lambda_1, \lambda_1]$$

$$= (13/9)\,\lambda_1^2-(13/3)\lambda_1+3$$

To find λ_1 for minimum f, we have

$$\frac{\partial f}{\partial \lambda_1} = \frac{26}{9}\lambda_1 - \frac{13}{3} = 0 \Rightarrow \lambda_1 = 1.5$$

and $$\frac{\partial^2 f}{\partial \lambda_1^2} = \frac{26}{9} > 0\,.$$

Thus f is minimum when λ_1=1.5 giving

$$X_2 = (0, 0)+1.5(2/3, 1)$$

$$= (1, 1.5)$$

and $g(X_2) = -1.5 < 0.$

So the new search direction S_2 is

$$S_2 = -\nabla f(X_2)$$

$$= \begin{Bmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \end{Bmatrix}_{X_1=(1,\,1.5)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

which implies that there is no search direction available to get minimum f.

Hence the minimum f = -0.25

At X = X_2 = (1, 1.5), *i.e.*, x_1 = 1 and x_2 = 1.5.

5.5 Indirect Method

In using the technique of indirect method for solving a NLPP, the function to be minimized must be differentiable. Here we discussed two important indirect methods:

1. Transform Techniques

2. Penalty Function Methods'

5.6 Transform Techniques

In the constraints $g_j(X)$ in a non-linear optimization problem exhibit simple forms in independent variables, it becomes feasible to apply variable transformations. Through these transformations, the constraints can be automatically satisfied. Consequently, it becomes possible to convert a constrained optimization problem into an unconstrained one by employing these transformation techniques.

Some of typical transformations of independent variables are:

(i) If lower and upper bounds on x_i are given as:

$$l_i \leq x_i \leq u_i, \qquad \text{……(5.8)}$$

then these can be satisfied by transforming the variable x_i as

$$x_i = l_i + (u_i - l_i)\sin^2 y_i \qquad \text{......(5.9)}$$

Where y_i is the new variable which can take any value.

(ii) If $x_i \in (0, 1)$, then we can use one of the following transformations:

(a) $x_i = \sin^2 y_i$

(b) $x_i = \cos^2 y_i$

(c) $x_i = \dfrac{e^{y_i}}{e^{y_i} + e^{-y_i}}$

(d) $x_i = \dfrac{y_i^2}{1 + y_i^2}$ ….(5.10)

(iii) If the design variable is restricted to assume only positive values then one of the following transformations can be used:

(a) $x_i = |y_i|$

(b) $x_i = y_i^2$ or

(c) $x_i = e^{y_i}$ ….(5.11)

(iv) If $x_i \in (-1,1)$, then we can use:

(a) $x_i = \sin y_i$

(b) $x_i = \cos y_i$ or

(c) $x_i = \dfrac{2y_i}{1+y_i^2}$ ….(5.12)

Note that, to use above transformations, the constrained function $g_i(X)$ should be simple. In case it is not possible to eliminate, all the constraints using by change of variables, then it is better not to use the transformation method.

Examples

Example.3. A courier service does not accept rectangular packets of more than 42 cm in length. If {length+2(width+height)} is at most of 72 cms, then compute the maximum volume of the rectangular packet.

Solution: Let us consider x_1, x_2 and x_3 be the length, width and height of the rectangular packet then the formulation of the problem is

$$\text{Max } f(X) = x_1x_2x_3 \qquad \text{….. (5.13)}$$

Subject to $x_1+2x_2+2x_3 \le 72,$ …… (5.14)

$x_1 \le 42,$ ….(5.15)

$x_1, x_2, x_3 \ge 0.$ ……(5.16)

Let us transform x_i's into y_i's by taking

$y_1 = x_1, y_2 = x_2$

and $y_3 = x_1+2x_2+2x_3$

These imply that $x_3 = (1/2)(y_3-y_1-2y_2)$ …..(5.17)

Thus constraints (5.14), (5.15) and (5.16) can be written as:

$0 \le y_1 \le 42,$

$0 \le y_2 \le 36,$

and $0 \le y_3 \le 72$, ...(5.18)

The upper bound for y_i's in (5.18) can easily be obtained (say for y_2, taking $x_1 = x_3 = 0$ in (5.14), we get $2x_2 \le 72$. Thus $x_2 = y_2 \le 36$).

The constrained in (5.18) are automatically satisfied if we define z_1, z_2 and z_3 as:

$$y_i = l_i+(u_i-l_i)\sin^2 z_i \quad .....(5.19)$$

Using (5.18) and (5.19) we have

$$y_1 = l_1+(u_1-l_1)\sin^2 z_1 = 42\sin^2 z_1 \quad .....(5.20)$$

$$y_2 = l_2+(u_2-l_2)\sin^2 z_2 = 36\sin^2 z_2 \quad .....(5.21)$$

$$y_3 = l_3+(u_3-l_3)\sin^2 z_3 = 72\sin^2 z_3 \quad .....(5.22)$$

So the original problem reduces to

Max $f = (1/2)\, y_1 y_2 (y_3-y_1-2y_2)$

$= (1/2)(42\sin^2 z_1)(36\sin^2 z_2)(72\sin^2 z_3-42\sin^2 z_1-72\sin^2 z_2)$

Subject to $0 \le \sin^2 z_i \le 1$, $i = 1, 2, 3$ ...... (5.23)

For minimum f, we have

$$\frac{\partial f}{\partial z_1} = 0$$

$$\Rightarrow \quad 1512\sin z_1 \cos z_1 \sin^2 z_2\left(\sin^2 z_3 - \frac{7}{6}\sin^2 z_1 - 72\sin^2 z_2\right) = 0 \quad ....(5.24)$$

$$\frac{\partial f}{\partial z_2} = 0$$

$$\Rightarrow \quad 108864\sin^2 z_1 \sin z_2 \cos z_2\left(\sin^2 z_3 - \frac{7}{12}\sin^2 z_1 - 2\sin^2 z_2\right) = 0 \quad ....(5.25)$$

$$\frac{\partial f}{\partial z_3} = 0$$

$$\Rightarrow \quad 108864\,\sin^2 z_1 \sin^2 z_2 \sin z_3 \cos z_3 = 0 \quad ....(5.26)$$

From (5.25), we have

$\sin^2 z_1 = 0 \quad \Rightarrow \quad z_1 = 0 \quad \Rightarrow \quad y_1 = 0 \quad \Rightarrow \quad f = 0$ which is not acceptable.

$\sin^2 z_2 = 0 \quad \Rightarrow \quad z_2 = 0 \;\Rightarrow\; y_2 = 0 \quad \Rightarrow \quad f = 0$ which is not acceptable.

If $\quad \sin z_3 = 0 \quad \Rightarrow \quad z_3 = 0 \;\Rightarrow\; y_3 = 0 \;\Rightarrow f$ is negative which is not acceptable.

If $\cos z_3 = 0 \quad \Rightarrow \quad z_3 = \pi/2$ ......(5.27)

From (5.24), we have

$$\sin^2 z_3 - \frac{7}{6}\sin^2 z_1 - 72\sin^2 z_2 = 0$$

$\Rightarrow \qquad 7\sin^2 z_1 + 6\sin^2 z_2 = 6$ ......(5.28)

From (5.25), we have

$$\sin^2 z_3 - \frac{7}{12}\sin^2 z_1 - 2\sin^2 z_2 = 0$$

$\Rightarrow \qquad 7\sin^2 z_1 + 24\sin^2 z_2 = 12$ ......(5.29)

Subtract Equation (5.28) from Equation (5.29), we have

$\Rightarrow \quad 18\sin^2 z_2 = 6$

$\Rightarrow \quad \sin^2 z_2 = 1/3.$

From (5.28), we have

$7\sin^2 z_1 + 6(1/3) = 6$

$\Rightarrow \quad \sin^2 z_1 = 4/7.$

Hence $y_1 = 42\,(4/7) = 24$, $y_2 = 36(1/3) = 12$, $y_3 = 72$.

$\therefore \qquad x_1 = 24,\ x_2 = 12,\ x_3 = 12$

and Max $f = 3456$ cm^3.

5.7 Penalty Function Method

Consider the problem

$\text{Min } f(X)$ ..… (5.30)

Subject to $\quad g_j(X) \le 0,\ i = 1, 2, \ldots\ldots\ldots, m$

$h_i(X) = 0,\ i = 1, 2, \ldots\ldots, p,\ X \in E_n$

The penalty method transforms problem (5.30) into a sequence of problems, each devoid of constraints.

In tackling problem (5.30), the penalty method introduces constraint effects by modifying the objective function, akin to the approach used in the Big-M method of artificial variable technique in the simplex method.

Consequently, to address the constrained optimization problem outlined in (5.30), we introduce an auxiliary unconstrained function.

$$F(X, r_k) = f(X) + P(X, r_k) \qquad \text{.....(5.31)}$$

Where $P(X, r_k)$ is a function of constraints $g_j(X)$ and $h_i(X)$ and r_k is a positive parameter such that

$$\lim_{r_k \to 0} \min F(X, r_k) = \min f(X).$$

5.8 Formulation of Penalty Function

On the basis of the formulation of penalty function, the penalty function method is divided into two categories:

1. The Interior Penalty Method

2. The Exterior Penalty Method

In the interior penalty method, the form of penalty function $P(X, r_k)$, which is a function of constraint functions $g_j(X)$ and $h_i(X)$ and a positive parameter r_k is

$$P(X, r_k) = -r_k \sum_{j=1}^{m} \frac{1}{g_j(X)} + \frac{1}{\sqrt{r_k}} \sum_{i=1}^{p} h_i^2(X) \qquad \text{......(5.32)}$$

In the interior penalty method the minimum of the auxiliary function $F(X, r_k)$ as defined above in (5.31) is approaching to the minimum of the objective function $f(X)$ from points inside the feasible region as r_k tends to zero.

In exterior penalty method we take the form of penalty function $P(X, r_k)$ as

$$P(X, r_k) = \frac{1}{r_k} \sum_{j=1}^{m} \left[\max\left(0, g_j(X)\right)\right]^2 + \frac{1}{r_k} \sum_{i=1}^{p} h_i^2(X) \qquad \text{......(5.33)}$$

and minimize the auxiliary function $F(X, r_k)$ for a sequence of decreasing values of r_k i.e., minimum of $F(X, r_k)$ approaches to minimum of $f(X)$ from the points of infeasible region as r_k tends to zero.

We will illustrate the above interior and exterior penalty based method through the following examples.

Example.4. Minimize f(X) = f (x_1, x_2) = $(x_1+2)^3+3x_2+1$

Subject to $x_1 \geq 2$, $x_2 \geq 0$.

Solution: Given problem can be re-written as

$f(X) = (x_1+2)^3+3x_2+1$

Subject to $2-x_1 \leq 0$, $-x_2 \leq 0$.

Hence the auxiliary unconstrained problem becomes

$$\text{Minimize } F(X, r_k) = (x_1+2)^3+3x_2+1-r_k\left(\frac{1}{2-x_1}-\frac{1}{x_2}\right).$$

For minimization of F, we have

$$\frac{\partial F}{\partial x_1} = 3(x_1+2)^2 - \frac{r_k}{(2-x_1)^2} = 0$$

$$\Rightarrow \quad \left(x_1^2-4\right)^2 = \frac{r_k}{3}$$

$$\text{or} \quad x_1 = \left(4+\sqrt{\frac{r_k}{3}}\right)^{1/2}$$

$$\text{and} \quad \frac{\partial F}{\partial x_2} = 3-\sqrt{\frac{r_k}{x_2^2}}$$

$$\Rightarrow \quad x_2 = \sqrt{\frac{r_k}{3}}$$

Thus we have

$$x_1 = \left(4+\sqrt{\frac{r_k}{3}}\right)^{1/2}$$

and $x_2 = \sqrt{\frac{r_k}{3}}$

are the possible feasible values of x_1 and x_2.

Now as $r_k \to 0$, $x_1 = 2$ and $x_2 = 0$.

$$\lim_{r_k \to 0} \min F(X, r_k) = \min f(X) = 65, \text{ when } x_1 = 2 \text{ and } x_2 = 0.$$

The solution of the given problem tending to $x_1 = 2$ and $x_2 = 0$ can be seen from the following table as $r_k \to 0$ through four values 10, 0.1, 0.001 and 0.0001 given to r_k for $k = 1, 2, 3$ and 4.

k	r_k	x_1	x_2	$F(X, r_k)$	$f(X)$
1	10	2.414	10.826	117.556	92.478
2	0.1	2.045	0.183	70.508	67.739
3	0.001	2.005	0.018	65.529	65.274
4	0.0001	2.000	0.000	65.000	65.000

Above table shows, how the values of auxiliary function $F(X, r_k)$ are approaching to the values of given function $f(X)$ as r_k is tending to zero through positive values.

Example.5. Minimize $f(X) = 2x$

Subject to $x \geq 3$ using interior penalty method.

Solution: The given problem can be written as

Min $f(X) = 2x$

Subject to $g_1(x) = 3 - x \leq 0$.

The auxiliary unconstrained problem can be written as

Minimize $F(X, r_k) = 2x - r_k\left(\frac{1}{3-x}\right)$ such that

$$\lim_{r_k \to 0} \min F(X, r_k) = \min f(X)$$

For minimization of F, we have

$$\frac{\partial F}{\partial x} = 2 - \frac{r_k}{(3-x)^2} = 0$$

$$\Rightarrow \quad x = 3 + \sqrt{\frac{r_k}{2}},$$

and $$\frac{\partial^2 F}{\partial x^2} = \frac{-2r_k}{(3-x)^3} > 0 \text{ at } x = 3 + \sqrt{\frac{r_k}{2}}.$$

Now as $r_k \to 0$ at $x = 3$.

$$\lim_{r_k \to 0} \min \ F(X, r_k) = \min \ f(X) = 6, \text{ when } x = 3.$$

Note that this is the problem in one variable x.

Example.6. Minimize f(X) = $x_1^2+2x_2^2$

Subject to $2x_1+5x_2 \le 10$ using exterior penalty method.

Solution: The given problem can be written as

f (X) = $x_1^2+2x_2^2$

Subject to $g_1(x) = 2x_1+5x_2 - 10 \le 0$

The auxiliary function F(X, r_k) in exterior penalty method is given as

Minimize F(X, r_k) = $x_1^2+2x_2^2+ \dfrac{1}{r_k}$ [Max. (0, $2x_1+5x_2-10$)]^2

For minimization of F, ($r=1/r_k$), we have

$$\frac{\partial F}{\partial x} = 2x_1 + \frac{4}{r_k}(2x_1 + 5x_2 - 10) = 0$$

$$\Rightarrow \quad (2+8r)x_1 + 20rx_2 - 40r = 0 \qquad \text{….(5.34)}$$

$$\frac{\partial F}{\partial x_2} = 4x_2 + 10r(2x_1 + 5x_2 - 10) = 0$$

$$\Rightarrow \quad 20rx_1 + (4+50r)x_2 - 100r = 0 \qquad \text{….(5.35)}$$

From equations (5.34) and (5.35), we have

$$\frac{x_1}{160\,r} = \frac{x_2}{200\,r} = \frac{1}{8 + 132\,r}$$

$$\Rightarrow \qquad x_1 = \frac{40r}{2+33r} = \frac{40}{2r_k + 33}$$

and $$x_2 = \frac{50r}{2+33r} = \frac{50}{2r_k + 33}$$

Now as $r_k \to 0$, we get

$$x_1 = \frac{40}{33} \quad \text{and} \quad x_2 = \frac{50}{33}$$

and minimum f (X) = $(40/33)^2+2(50/33)^2$

= 200/33.

Hence the minimum of the f (X) is 200/33.

5.9 Summary

In solving NLPP, the function to be minimized need not be differentiable and only computational work is needed. In such conditions direct methods (like Complex Method, Zoutendijk Method etc.) are convenient to use. In using the technique of indirect methods for solving a NLPP, the function to be minimized must be differentiable. In such cases indirect methods (like Transform Techniques, Penalty Function Methods etc.) are used.

5.10 Terminal Questions

Q.1. Minimize f (X) = $f(x_1, x_2) = (x_1 -1)^2+(x_2-2)^2$,

Subject to $x_1+x_2 \le 4$, $x_1-x_2 \le 2$, $x_1 \ge 0$ and $x_2 \ge 0$ by complex method with $X_1 = \begin{pmatrix} 0.5 \\ 1.5 \end{pmatrix}$.

Q.2. Using the method of feasible directions due to Zoutendijk method

Minimize f (X) = $x_1^2+x_2^2-4x_1-4x_2+8$

Subject to $g_1(X) = x_1+2x_2 -4 \le 0$ taking the starting point $X_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Q.3. Min $f(X) = x_1^2+x_2^2-6x_1-8x_2+10$

Subject to $4x_1^2+x_2^2 \leq 16$, $3x_1+5x_2 \leq 15$

Using (i) Interior penalty method (ii) Exterior penalty method.

Answer

1. $x_1=1$, $x_2=2$.
2. $x_1=1.6$, $x_2 = 1.2$, $f(X) = 0.8$.
3. $x_1=3$, $x_2 = 4$.

UNIT-6: Quadratic Programming

Structure

6.1 Introduction

In the present unit we shall discuss about the convex and concave functions, Kuhn-Tucker conditions, quadratic programming problems which is solved by two important techniques named as Wolfe's modified simplex method and Beale's method. Quadratic programming (QP) is a type of mathematical optimization problem that deals with quadratic objective functions subject to linear equality and inequality constraints. The goal of quadratic programming is to find the values of the decision variables (x) that minimize (or maximize) the quadratic objective function while satisfying the given linear constraints. Quadratic programming problems arise in various fields, such as finance, engineering, operations research, and machine learning. Common applications include portfolio optimization, structural design, and support vector machines in machine learning.

6.2 Objectives

After reading this unit the learner should be able to understand about:

- the Convex and Concave functions

- the Kuhn-Tucker conditions
- the quadratic programming problem
- Wolfe's Modified Simplex Method
- Beale's Method

6.3 Convex and Concave Functions

A function $f(x)$ is said to be convex over a convex set S, if for any two point x_1 and x_2 in S

$$f[\lambda x_1+(1-\lambda)x_2] \le \lambda f(x_1)+(1-\lambda)f(x_2); \quad 0 \le \lambda \le 1$$

A function $f(x)$ is said to be concave over a convex set S if for any two points x_1 and x_2 in S

$$f[\lambda x_1+(1-\lambda)x_2] \ge \lambda f(x_1)+(1-\lambda)f(x_2); \qquad 0 \le \lambda \le 1$$

Note:

1. A function $f(x)$ is said to be convex iff the Hessian matrix $H(x)$ of second partial derivatives is positive semi-definite.

2. A function $f(x)$ is said to be concave iff the Hessian matrix $H(x)$ of second partial derivatives is negative semi-definite.

6.4 Kuhn-Tucker Conditions

This section we shall discussed the Kuhn-Tucker conditions under certain restrictions for identifying the stationary points of constrained non-linear optimization problems. Consider the optimization problem

$$\text{Optimization (Max or Min) } Z=f(X) \qquad \dots (6.1)$$

$$\text{s.t.} \quad g_j(X) \le 0; \; j = 1, 2, 3, \dots, m \qquad \dots (6.2)$$

Converting the inequality constraints into the equality constraints by adding slack variables s_j^2, we have

$$g_j(X) + s_j^2 = 0 \qquad \dots (6.3)$$

Now we define a Lagrangian function

$$L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m) = f(X)+\sum_{j=1}^{m}\lambda_j \left\{g_j(X)+s_j^2\right\} \qquad \dots (6.4)$$

The Kuhn-Tucker necessary conditions for extreme points are

$$\frac{\partial L}{\partial x_i} = 0; i = 1,2,3,...., n;$$

$\Rightarrow$ $$\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0 \qquad ..... (6.5)$$

$$\frac{\partial L}{\partial s_j} = 0$$

$\Rightarrow$ $$g_j(X) + s_j^2 = 0; \quad j = 1, 2, 3, ..., m \qquad ..... (6.6)$$

and $$\frac{\partial L}{\partial y_j} = 0$$

$\Rightarrow$ $$2\lambda_j s_j = 0; \quad j = 1, 2, 3, ..., m \qquad ..... (6.7)$$

From Equations (6.6) and (6.7), we have

$$\lambda_j g_j(X) = 0$$

$\Rightarrow$ $$\lambda_j = 0$$

or $$g_j(X) = 0.$$

Case-I: If $g_j(X) = 0$ at the optimum point then the constraint is called active constraints and we can determine the optimum solution.

Case-II: If $\lambda_j = 0$ at the optimum point then it is known as an inactive constraints.

Note: If the given optimization problem is a minimization problem with constraints of the form $g_j(X) \geq 0$ then $\lambda_j \leq 0$ but if the given problem is a maximization problem with constraints of the form $g_j(X) \leq 0$ then $\lambda_j \leq 0$.

Consider some of the maximization or minimization problems in the following terms.

(i) Maximize $Z = f(X)$

s.t. $g_j(X) \leq 0$; $j = 1, 2, 3, ..., m$

For the function $f(X)$ maxima, we have

$$\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0 \qquad ; \; i = 1, 2, 3, ..., m$$

$$\lambda_j g_j(X) = 0 \qquad ; \; j = 1, 2, 3, ..., m$$

and $$\lambda_j \leq 0.$$

(ii) Maximize Z= f(X)

s.t. $g_j(X) \geq 0$; $j = 1, 2, 3, \ldots\ldots\ldots, m$

For the function f(X) maxima, we have

$$\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0 \qquad ; \ i = 1, 2, 3, \ldots\ldots\ldots, m$$

$$\lambda_j g_j(X) = 0 \qquad ; \ j = 1, 2, 3, \ldots\ldots\ldots, m$$

and $\lambda_j \geq 0$.

(iii) Minimize Z= f(X)

s.t. $g_j(X) \leq 0$; $j = 1, 2, 3, \ldots\ldots\ldots, m$

For the function f(X) maxima, we have

$$\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0 \qquad ; \ i = 1, 2, 3, \ldots\ldots\ldots, m$$

$$\lambda_j g_j(X) = 0 \qquad ; \ j = 1, 2, 3, \ldots\ldots\ldots, m$$

and $\lambda_j \geq 0$.

(iv) Minimize Z= f(X)

s.t. $g_j(X) \geq 0$; $j = 1, 2, 3, \ldots\ldots\ldots, m$

For the function f(X) maxima, we have

$$\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0 \qquad ; \ i = 1, 2, 3, \ldots\ldots\ldots, m$$

$$\lambda_j g_j(X) = 0 \qquad ; \ j = 1, 2, 3, \ldots\ldots\ldots, m$$

and $\lambda_j \leq 0$.

Examples

Example.1. Solve the following problem

Minimize $f(X) = x_1^2 + x_2^2 + x_3^2$

s.t. $g_1(X) = 2x_1+x_2-5 \le 0$

$g_2(X) = x_1+x_3-2 \le 0$

$g_3(X) = 1-x_1 \le 0$

$g_4(X) = 2-x_2 \le 0$

$g_5(X) = -x_3 \le 0.$

Solution: it is given that

$$\text{Minimize } f(X) = x_1^2+x_2^2+x_3^2 \quad \ldots(6.8)$$

$$\text{s.t.} \quad \left.\begin{aligned} g_1(X) &= 2x_1 + x_2 - 5 \le 0 \\ g_2(X) &= x_1 + x_3 - 2 \le 0 \\ g_3(X) &= 1 - x_1 \le 0 \\ g_4(X) &= 2 - x_2 \le 0 \\ g_5(X) &= -x_3 \le 0 \end{aligned}\right\} \quad \ldots(6.9)$$

Now define a Lagrangian function by introducing slack variable s_j^2, we have

$$L(x_1, x_2, x_3; \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5; s_1, s_2, s_3, s_4, s_5) = f(X) + \sum_{J=1}^{5} \lambda_j \left[g_j(X) + s_j^2 \right] \quad \ldots(6.10)$$

From equation (6.8), (6.9) and (6.10), we have

$$L = x_1^2+x_2^2+x_3^2+\lambda_1(2x_1+x_2-5+s_1^2)+\lambda_2(x_1+x_3-2+s_2^2)+\lambda_3(1-x_1+s_3^2)+\lambda_4(2-x_2+s_4^2)+\lambda_5(-x_3+s_5^2) \quad \ldots(6.11)$$

The Kuhn-Tucker necessary conditions for minimization of L (with $g_j(X) \le 0$) are

$$\frac{\partial L}{\partial x_i} = 0; \quad i = 1,\ 2,\ 3.$$

$$\lambda_j g_j = 0; \quad j = 1, 2, 3, 4, 5.$$

$$\lambda_j \ge 0; \quad j = 1, 2, 3, 4, 5.$$

Differentiate partially Equation (6.11) and we get

$$\left.\begin{aligned}\frac{\partial L}{\partial x_1} &= 2x_1 + 2\lambda_1 + \lambda_2 - \lambda_3 = 0\\ \frac{\partial L}{\partial x_2} &= 2x_2 + \lambda_1 - \lambda_4 = 0\\ \frac{\partial L}{\partial x_3} &= 2x_3 + \lambda_2 - \lambda_5 = 0\end{aligned}\right\} \qquad ...(6.12)$$

$$\left.\begin{aligned}\lambda_1(2x_1 + x_2 - 5) = 0\\ \lambda_2(x_1 + x_3 - 2) = 0\\ \lambda_3(1 - x_1) = 0\\ \lambda_4(2 - x_2) = 0\\ \lambda_5(-x_3) = 0\end{aligned}\right\} \qquad ...(6.13)$$

$$\lambda_j \geq 0;\ j = 1, 2, 3, 4, 5. \qquad .....(6.14)$$

Let $\lambda_3 \neq 0$ and $\lambda_4 \neq 0$, we have

$x_1 = 1$, $x_2 = 2$, $x_3 = 0$ and $\lambda_5 = 0$..

If we take $x_1 = 1$, $x_2 = 2$, $x_3 = 0$ then equation (6.13) is satisfied.

Now from Equation (6.13), we have

$$2x_1+x_2-5 \neq 0$$

and $$x_1+x_2-2 \neq 0$$

so $\lambda_1 = 0$ and $\lambda_2 = 0$.

From Equation (6.12), we have

$2-\lambda_3 = 0 \quad \Rightarrow \quad \lambda_3 = 2$

$4-\lambda_4 = 0 \quad \Rightarrow \quad \lambda_4 = 4$

$\lambda_2-\lambda_5 = 0 \quad \Rightarrow \quad \lambda_2 = \lambda_5$

Hence the optimum solution is $x_1 = 1$, $x_2 = 2$, $x_3 = 0$, $\lambda_1 = \lambda_2 = \lambda_5 = 0$, $\lambda_3 = 2$ and $\lambda_4 = 4$

and minimize $f = (1)^2+(2)^2+(0)^2 = 5$.

6.5 Quadratic Programming

A non-linear programming problem involved objective function is quadratic and constraints are in linear form is known as quadratic programming. Here we shall discuss two methods to solve quadratic programming problems:

1. Wolfe's modified simplex method

2. Beale’s method

6.6 Wolfe’s Modified Simplex Method

To solve a quadratic programming problem using Wolfe’s modified simplex method, first we use Kuhn-Tucker conditions to express the problem in such a form to apply the computational procedure based on the simplex method. The procedure to solve a quadratic programming problem by Wolfe’s modified simplex method as follows:

Consider a quadratic programming problem

$$Max \quad Z = f(X) = \sum_{j=1}^{n} c_j x_j + \frac{1}{2} \sum_{j,k=1}^{n} x_j d_{jk} x_k$$

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i$$

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

Where $d_{jk} = d_{kj}$ for all j and k and where $b_i \geq 0$.

Step 1: If the given quadratic programming problem is in the minimization form, then first convert it into maximization form.

Step 2: Now convert the inequality constraints of the given problem into equations by introducing the slack variables $s_i^{\ 2}$ in the i^{th} constraint $(i = 1,2.....m)$ and the slack variables $s_{m+j}^{\ 2}$ in the i^{th} non-negativity constraints $(j = 1,2.....n)$

Step 3: Now construct the Lagrangian function

$$L(x,s,\lambda) = f(x) - \sum_{i=1}^{m} \lambda_i \left[\sum_{j=1}^{n} a_{ij} x_j - b_i + s_i^{\ 2} \right] - \sum_{j=1}^{n} \lambda_{m+j} \left(-x_j + s_{m+j}^{\ 2} \right)$$

Where $x = (x_1, x_2,.......x_n), s = (s_1, s_2,.....s_{m+n}), \lambda = (\lambda_1, \lambda_2,.....\lambda_{m+n})$

Step 4: Differentiate the $L(x,s,\lambda)$ partially in respect to the components of x,s and λ and equating to zero the first order partial derivatives. Construct the Kuhn-tucker conditions from the resulting equations.

Step 5: Now introduce the non-negative artificial variables $A_j,\ j=1,2,.....n$ in the Kuhn-tucker condition, we have $c_j+\sum_{k=1}^{n}d_{jk}x_k-\sum_{i=1}^{m}\lambda_i a_{ij}+\lambda_{m+j}=0$ for $j=1,2,....n.$

Step 6: Construct the objective function

$$Max\ \ Z=-A_1-A_2-.....-A_n$$

Step 7: Find the initial basic feasible solution to the following LPP:

Max $Z=-A_1-A_2-.....A_n$ subject to the constraints:

$$\sum_{k=1}^{n}d_{jk}x_k-\sum_{i=1}^{m}\lambda_i a_{ij}+\lambda_{m+j}+A_j=-c_j$$

$$\sum_{j=1}^{n}a_{ij}x_j+x_{n+i}=b_i$$

$$A_j,\lambda_i,\lambda_{m+j},x_j\geq 0$$

Where $x_{n+1}=s_i^{\ 2},\ \ i=1,2,....,m$ and satisfying the complementary slackness conditions (restricted basis conditions)

$$\sum_{j=1}^{n}\lambda_{m+j}x_j+\sum_{j=1}^{m}x_{n+i}\lambda_i=0$$

Step 8: Now use phase I of artificial variable techniques (two phase method) to find an optimum solution to the Linear Programming Problem (LPP) of step 7, which is satisfying the complementary slackness conditions.

This optimum solution is an optimum solution for the given quadratic programming problem also.

Examples

Example.2. Use Wolfe's modified simplex method to solve the Quadratic programming problem:

Max $Z=4x_1+6x_2-2x_1^{\ 2}-2x_1x_2-2x_2^{\ 2}$

s.t. $x_1+2x_2\leq 2$,

$x_1, x_2 \geq 0.$

Solution: The given problem is maximization problem.

Consider $x_1 \geq 0$ and $x_2 \geq 0$ also as the inequality constraints, convert the inequality constraints into equations by introducing slack variables s_1^2, s_2^2 and s_3^2 respectively.

The modified problem is

$$\text{Max } Z = 4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2$$

$$x_1 + 2x_2 + s_1^2 = 2$$

$$-x_1 + s_2^2 = 0$$

$$-x_2 + s_3^2 = 0$$

Construct the Lagrangian function

$$L = L(x_1, x_2, s_1, s_2, s_3, \lambda_1, \lambda_2, \lambda_3)$$

$$= \left(4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2\right) - \lambda_1\left(x_1 + 2x_2 + s_1^2 - 2\right) - \lambda_2\left(-x_1 + s_2^2\right) - \lambda_3\left(-x_2 + s_3^2\right) = 0$$

Differentiating partially with respect to $x_1, x_2, s_1, s_2, s_3, \lambda_1, \lambda_2, \lambda_3$ and equating to zero, we get

$$\frac{\partial L}{\partial x_1} = 4 - 4x_1 - 2x_2 - \lambda_1 + \lambda_2 = 0$$

$$\frac{\partial L}{\partial x_2} = 6 - 2x_1 - 4x_2 - 2\lambda_1 + \lambda_3 = 0$$

$$\frac{\partial L}{\partial s_1} = -2\lambda_1 s_1 = 0$$

$$\frac{\partial L}{\partial s_2} = -2\lambda_2 s_2 = 0$$

$$\frac{\partial L}{\partial s_3} = -2\lambda_3 s_3 = 0$$

$$\frac{\partial L}{\partial \lambda_1} = x_1 + 2x_2 + s_1^2 - 2 = 0$$

$$\frac{\partial L}{\partial \lambda_2} = -x_1 + {s_2}^2 = 0$$

$$\frac{\partial L}{\partial \lambda_3} = -x_2 + {s_3}^2 = 0$$

On simplification, we have

$$\left.\begin{aligned} &4x_1 + 2x_2 + \lambda_1 - \lambda_2 = 4 \\ &2x_1 + 4x_2 - 2\lambda_1 - \lambda_3 = 6 \\ &x_1 + 2x_2 + {s_1}^2 = 2 \end{aligned}\right\} \quad \ldots.\ (6.14)$$

$$\left.\begin{aligned} &\lambda_1 {s_1}^2 + \lambda_2 {s_2}^2 + \lambda_3 {s_3}^2 = 0 \quad \Rightarrow \lambda_1 {s_1}^2 + \lambda_2 x_1 + \lambda_3 x_2 = 0 \\ &x_1, x_2, {s_1}^2, \lambda_i \geq 0, \quad i = 1, 2, 3 \end{aligned}\right\} \quad \ldots(6.15)$$

A solution $x_j,\ j = 1, 2$ of equation (6.14) above, satisfying equation (6.15) shall necessarily be an optimal one for maximizing L.

To obtain the solution to the above simultaneous equation (6.14) we introduce the artificial variables A_1 and A_2 (Both non-negative) in the first two constraints of equation (6.14) and construct the new objective function Max $Z = -A_1 - A_2$

Max $Z = -A_1 - A_2$

s.t. $4x_1 + 2x_2 + \lambda_1 - \lambda_2 + A_1 = 4$

$2x_1 + 4x_2 + 2\lambda_1 - \lambda_3 + A_2 = 6$

$x_1 + 2x_2 + x_3 = 2$

$x_1, x_2, x_3 \geq 0$

$A_1, A_2, \lambda_i \geq 0 \qquad i = 1, 2, 3.$

Where we have replaced ${s_1}^2$ by x_3 and satisfying the complementary slackness condition

$$\sum \lambda_i x_i = 0$$

The optimum solution to the above LPP shall now be obtained by the phase-I of artificial variable technique. Thus the starting simplex table is

cost →				0	0	0	0	0	0	-1	-1	Minimum ratio x_B / x_1
Variable →				x_1	x_2	x_3	λ_1	λ_2	λ_3	A_1	A_2	
Table No.	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^7	α^8	
1	-1	A_1	4	[4]	2	0	1	-1	0	1	0	1→
	-1	A_2	6	2	4	0	2	0	-1	0	1	3
	0	x_3	2	1	2	1	0	0	0	0	0	2
c_j-Z_j				6↑	6	0	3	-1	-1	0	0	

In the table 1 we see that x_1 or x_2 are the variables with most positive entry 6 in c_j-Z_j row, let x_1 enter the basis and A_1 leaves. The next simplex table is:

cost →				0	0	0	0	0	0	-1	Minimum ratio x_B / x_2
Variable →				x_1	x_2	x_3	λ_1	λ_2	λ_3	A_2	
Table No.	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^8	
2	0	x_1	1	1	1/2	0	1/4	-1/4	0	0	2
	-1	A_2	4	0	3	0	3/2	1/2	-1	1	4/3
	0	x_3	1	0	[3/2]	1	-1/4	1/4	0	0	2/3→
c_j-Z_j				0	3↑	0	3/2	1/2	-1	0	

In the table 2 we see that x_2 is the variable with most positive entry 3 in c_j-Z_j row, and hence will enter the basis and x_3 leaves. The next simplex table is:

cost →				0	0	0	0	0	0	-1	Minimum ratio x_B / λ_1
Variable →				x_1	x_2	x_3	λ_1	λ_2	λ_3	A_2	
Table No.	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^8	
3	0	x_1	2/3	1	0	-1/3	1/3	-1/3	0	0	2
	-1	A_2	2	0	0	-2	[2]	0	-1	1	1→
	0	x_2	2/3	0	1	2/3	-1/6	1/6	0	0	…..
c_j-Z_j				0	0	-2	2↑	0	-1	0	

In the table 3 we observe that λ_1 is the variable with most positive entry 2 in c_j-Z_j row and hence will enter the basis, and A_2 leaves. The next simplex table is:

cost →				0	0	0	0	0	0	Minimum ratio
Variable →				x_1	x_2	x_3	λ_1	λ_2	λ_3	
Table No.	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	
4	0	x_1	1/3	1	0	0	0	-1/3	1/6	2
	0	λ_1	1	0	0	-1	1	0	-1/6	1→
	0		5/6	0	1	1/2	0	1/6	-1/6	…..

		x_2								
c_j-Z_j				0	0	0	0	0	0	

In the above table 4, c_j-$Z_j \le 0$ for all j, therefore solution is optimal.

Thus the optimal solution is $x_1 = \frac{1}{3}, x_2 = \frac{5}{6}$ and

$$\text{Max } Z = 4x_1 + 6x_2 - 2x_1^{\ 2} - 2x_1x_2 - 2x_2^{\ 2}$$

$$= 4\left(\frac{1}{3}\right) + 6\left(\frac{5}{6}\right) - 2\left(\frac{1}{3}\right)^2 - 2\left(\frac{1}{3}\right)\left(\frac{5}{6}\right) - 2\left(\frac{5}{6}\right)^2$$

$$= \frac{4}{3} + \frac{30}{6} - \frac{2}{9} - \frac{10}{18} - \frac{50}{36} = \frac{25}{6}.$$

Example.3. Use Wolfe's modified simplex method to solve the Quadratic programming problem:

$$\textbf{Max } Z = 2x_1 + 3x_2 - 2x_1^{\ 2}$$

$$\textbf{s.t.} \quad x_1 + 4x_2 \le 4,$$

$$x_1 + x_2 \le 2,$$

$$x_1, x_2 \ge 0$$

Solution: The given problem is a maximization problem. Consider $x_1 \ge 0$ *and* $x_2 \ge 0$ also as the inequality constraints, convert the inequality constraints into equations by introducing slack variables $s_1^{\ 2}$, $s_2^{\ 2}$, $s_3^{\ 2}$ and $s_4^{\ 2}$ respectively. The modified problem is

$$\text{Maximize } Z = 2x_1 + 3x_2 - 2x_1^{\ 2}$$

s.t.

$$x_1 + 4x_2 + s_1^{\ 2} = 4$$

$$x_1 + x_2 + s_2^{\ 2} = 2$$

$$-x_1 + s_3^2 = 0$$

$$-x_2 + s_4^2 = 0$$

Construct the Lagrangian function

$$L = L(x_1, x_2, s_1, s_2, s_3, s_4, \lambda_1, \lambda_2, \lambda_3, \lambda_4)$$

$$= (2x_1 + 3x_2 - 2x_1^2) - \lambda_1(x_1 + 4x_2 + s_1^2 - 4) - \lambda_2(x_1 + x_2 + s_2^2 - 2) - \lambda_3(-x_1 + s_3^2) - \lambda_4(-x_2 + s_4^2)$$

Differentiating partially with respect to $x_1, x_2, s_1, s_2, s_3, s_4, \lambda_1, \lambda_2, \lambda_3, \lambda_4$ and equating to zero, we have

$$\frac{\partial L}{\partial x_1} = 2 - 4x_1 - \lambda_1 - \lambda_2 + \lambda_3 = 0$$

$$\frac{\partial L}{\partial x_2} = 3 - 4\lambda_1 - \lambda_2 + \lambda_4 = 0$$

$$\frac{\partial L}{\partial s_1} = -2\lambda_1 s_1 = 0$$

$$\frac{\partial L}{\partial s_2} = -2\lambda_2 s_2 = 0$$

$$\frac{\partial L}{\partial s_3} = -2\lambda_3 s_3 = 0$$

$$\frac{\partial L}{\partial s_4} = -2\lambda_4 s_4 = 0$$

$$\frac{\partial L}{\partial \lambda_1} = x_1 + 4x_2 + s_1^2 - 4 = 0$$

$$\frac{\partial L}{\partial \lambda_2} = x_1 + x_2 + s_2^2 - 2 = 0$$

$$\frac{\partial L}{\partial \lambda_3} = -x_1 + s_3^2 = 0$$

$$\frac{\partial L}{\partial \lambda_4} = -x_2 + s_4^2 = 0$$

On simplification, we have

$$4x_1 + \lambda_1 + \lambda_2 - \lambda_3 = 2$$
$$4\lambda_1 + \lambda_2 - \lambda_4 = 3$$
$$x_1 + 4x_2 + s_1^{\ 2} = 4$$
$$x_1 + x_2 + s_2^{\ 2} = 2$$
.... (6.16)

$$\lambda_1 s_1^{\ 2} + \lambda_2 s_2^{\ 2} + x_1\lambda_3 + x_2\lambda_4 = 0$$
$$x_1, x_2, s_1^{\ 2}, s_2^{\ 2}, \lambda_i \geq 0,\ i = 1,2,3,4.$$
...... (6.17)

A solution $x_j,\ j = 1, 2$ of equation (6.16) above, satisfying equation (6.17) shall necessarily be an optimal one for maximizing L.

To obtain the solution to the above simultaneous equation (6.16) we introduce the artificial variables A_1 and A_2 (Both non-negative) in the first two constraints of equation (6.16) and construct the new objective function Max $Z = -A_1 - A_2$

Now the problem becomes

Max $Z = -A_1 - A_2$

s.t. $4x_1 + \lambda_1 + \lambda_2 - \lambda_3 + A_1 = 2$

$4\lambda_1 + \lambda_2 - \lambda_4 + A_2 = 3$

$x_1 + 4x_2 + x_3 = 4$

$x_1 + x_2 + x_4 = 2$

$x_1, x_2, x_3, x_4, A_1, A_2, \lambda_i \geq 0, i = 1,2,3,4.$

Where we have replaced $s_1^{\ 2}$ by x_3 *and* $s_2^{\ 2}$ by x_4 and satisfying the complementary slackness condition $\sum \lambda_i x_i = 0.$

The optimum solution to the above LPP shall now be obtained by the phase-I of artificial variable technique. Thus the starting simplex table is

In the table 1 we see that λ_1 is the

cost →				0	0	0	0	0	0	0	0	-1	-1	Minimum ratio x_B / x_1
Variable →				x_1	x_2	x_3	x_4	λ_1	λ_2	λ_3	λ_4	A_1	A_2	
Table No	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^7	α^8	α^9	α^{10}	
1	-1	A_1	2	[4]	0	0	0	1	1	-1	0	1	0	1/2→
	-1	A_2	3	0	0	0	0	4	1	0	-1	0	1	….
	0	x_3	4	1	4	1	0	0	0	0	0	0	0	4
	0	x_4	2	1	1	0	1	0	0	0	0	0	0	2
c_j-Z_j				4↑	0	0	0	5	2	-1	-1	0	0	

variable with most positive entry 5 in c_j-Z_j row and hence will enter the basis. But λ_1 will not enter the basis because x_3 is in the basis (complementary slackness conditions $\lambda_1 x_3 = 0$). The next most positive entry is 4 for x_1 column in c_j-Z_j row, so x_1 enter the basis and A_1 leaves. The next simplex table is 2. In the table 2 we observe that either λ_1 *or* λ_2 can enter the basis but x_3 *and* x_4 are still in the basis so these cannot enter the basis because of the complementary slackness conditions $\lambda_1 x_3 = 0$ and $\lambda_2 x_4 = 0$. Here x_2 can enter the basis because λ_4 is not in the basis (complementary slackness conditions $\lambda_4 x_2 = 0$).

cost →				0	0	0	0	0	0	0	0	-1	Minimum ratio x_B / x_2
Variable →				x_1	x_2	x_3	x_4	λ_1	λ_2	λ_3	λ_4	A_2	
Table No	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^7	α^8	α^{10}	

2	0	x_1	1/2	1	0	0	0	1/4	1/4	-1/4	0	0	…
	-1	A_2	3	0	0	0	0	4	1	0	-1	1	…
	0	x_3	7/2	0	4 (boxed)	1	0	-1/4	-1/4	1/4	0	0	7/8→
	0	x_4	3/2	0	1	0	1	-1/4	-1/4	1/4	0	0	3/2
c_j-Z_j				0	0↑	0	0	4	1	0	-1	0	

Now use entering and leaving variable rules to get the new simplex table.

cost →				0	0	0	0	0	0	0	0	-1	Minimum ratio x_B / λ_1
Variable →				x_1	x_2	x_3	x_4	λ_1	λ_2	λ_3	λ_4	A_2	
Table No	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^7	α^8	α^{10}	
3	0	x_1	1/2	1	0	0	0	1/4	-1/4	-1/4	0	0	2
	-1	A_2	3	0	0	0	0	4 (boxed)	1	0	-1	1	3/8→
	0	x_2	7/8	0	1	1/4	0	-1/16	-1/16	1/16	0	0	….
	0	x_4	5/8	0	0	-1/4	1	-3/16	-3/16	3/16	0	0	….
c_j-Z_j				0	0	0	0	4↑	1	0	-1	0	

In the table 3 we see that either λ_1 *or* λ_2 can enter the basis but x_4 is still in the basis so λ_2 cannot enter the basis because of the complementary slackness condition $\lambda_2 x_4 = 0$. Hence λ_1 enter the basis. Now use entering and leaving variable rules to get the new simplex table:

cost →				0	0	0	0	0	0	0	0	Minimum ratio
Variable →				x_1	x_2	x_3	x_4	λ_1	λ_2	λ_3	λ_4	
Table No	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^7	α^8	
4	0	x_1	5/16	1	0	0	0	0	3/16	-1/4	1/16	2
	0	λ_1	3/4	0	0	0	0	1	1/4	0	-1/4	3/8
	0	x_2	59/64	0	1	1/4	0	0	-3/64	1/16	-1/16	….
	0	x_4	49/64	0	0	-1/4	1	0	-9/16	3/16	-3/64	….
c_j-Z_j				0	0	0	0	0	0	0	0	

In the above table 4, c_j-$Z_j \le 0$ for all j, therefore solution is optimal.

Thus the optimal solution is $x_1 = \frac{5}{16}$, $x_2 = \frac{59}{64}$ and

Max $Z = 2x_1 + 3x_2 - 2x_1^{\,2}$

$$= 2\left(\frac{5}{16}\right) + 3\left(\frac{59}{64}\right) - 2\left(\frac{25}{256}\right)$$

$$= \frac{160 + 708 - 50}{256} = \frac{818}{256} = 3.19$$

Example.4 Use Wolfe's modified simplex method to solve the Quadratic programming problem: Max $Z = 6x_1 + 3x_2 - 4x_1x_2 - 2x_1^{\,2} - 3x_2^{\,2}$

s.t. $x_1 + x_2 \le 1$,

$2x_1 + 3x_2 \le 4$

$x_1, x_2 \geq 0$

Solution: The given problem is a maximization problem. Consider $x_1 \geq 0$ *and* $x_2 \geq 0$ also as the inequality constraints, convert the inequality constraints into equations by introducing slack variables $s_1^{\ 2}$, $s_2^{\ 2}$, $s_3^{\ 2}$ and $s_4^{\ 2}$ respectively. The modified problem is

Max $Z = 6x_1 + 3x_2 - 4x_1x_2 - 2x_1^{\ 2} - 3x_2^{\ 2}$

$$x_1 + x_2 + s_1^{\ 2} = 1$$

$$2x_1 + 3x_2 + s_2^{\ 2} = 4$$

$$-x_1 + s_3^{\ 2} = 0$$

$$-x_2 + s_4^{\ 2} = 0$$

Construct the Lagrangian function

$$L = L(x_1, x_2, s_1, s_2, s_3, s_4, \lambda_1, \lambda_2, \lambda_3, \lambda_4)$$

$$= (6x_1 + 3x_2 - 4x_1x_2 - 2x_1^{\ 2} - 3x_2^{\ 2}) - \lambda_1(x_1 + x_2 + s_1^{\ 2} - 1) - \lambda_2(2x_1 + 3x_2 + s_2^{\ 2} - 4) - \lambda_3(-x_1 + s_3^{\ 2}) - \lambda_4(-x_2 + s_4^{\ 2})$$

Differentiating partially with respect to $x_1, x_2, s_1, s_2, s_3, s_4, \lambda_1, \lambda_2, \lambda_3, \lambda_4$ and equating to zero, we have

$$\frac{\partial L}{\partial x_1} = 6 - 4x_2 - 4x_1 - \lambda_1 - 2\lambda_2 + \lambda_3 = 0$$

$$\frac{\partial L}{\partial x_2} = 3 - 4x_1 - 6x_2 - \lambda_1 - 3\lambda_2 + \lambda_4 = 0$$

$$\frac{\partial L}{\partial s_1} = -2\lambda_1 s_1 = 0$$

$$\frac{\partial L}{\partial s_2} = -2\lambda_2 s_2 = 0$$

$$\frac{\partial L}{\partial s_3} = -2\lambda_3 s_3 = 0$$

$$\frac{\partial L}{\partial s_4} = -2\lambda_4 s_4 = 0$$

$$\frac{\partial L}{\partial \lambda_1} = x_1 + x_2 + s_1^{\ 2} - 1 = 0$$

$$\frac{\partial L}{\partial \lambda_2} = 2x_1 + 3x_2 + s_2{}^2 - 4 = 0$$

$$\frac{\partial L}{\partial \lambda_3} = -x_1 + s_3{}^2 = 0$$

$$\frac{\partial L}{\partial \lambda_4} = -x_2 + s_4{}^2 = 0$$

On simplification, we have

$$\left.\begin{aligned} &4x_1 + 4x_2 + \lambda_1 + 2\lambda_2 - \lambda_3 = 6 \\ &4x_1 + 6x_2 + \lambda_1 + 3\lambda_2 - \lambda_4 = 3 \\ &x_1 + x_2 + s_1{}^2 = 1 \\ &2x_1 + 3x_2 + s_2{}^2 = 4 \end{aligned}\right\} \quad \ldots.(6.18)$$

$$\left.\begin{aligned} &\lambda_1 s_1{}^2 + \lambda_2 s_2{}^2 + x_1\lambda_3 + x_2\lambda_4 = 0 \\ &x_1, x_2, s_1{}^2, s_2{}^2,\ \lambda_i \geq 0 \qquad i = 1, 2, 3, 4 \end{aligned}\right\} \quad \ldots.\ (6.19)$$

A solution $x_j,\ \ j = 1, 2$ to equation (6.18) above and satisfying equation (6.19) shall necessarily be an optimal one for maximizing L.

To determine the solution to the above simultaneous equation (6.18) we introduce the artificial variables A_1 and A_2 (both non-negative) in the first two constraints of equation (6.18) and construct the dummy objective function

$$\text{Max } Z = -A_1 - A_2.$$

Now the problem becomes

$$\text{Max } Z = -A_1 - A_2$$

s.t.
$$4x_1 + 4x_2 + \lambda_1 + 2\lambda_2 - \lambda_3 + A_1 = 6$$

$$4x_1 + 6x_2 + \lambda_1 + 3\lambda_2 - \lambda_4 + A_2 = 3$$

$$x_1 + x_2 + x_3 = 1$$

$$2x_1 + 3x_2 + x_4 = 4$$

$$x_1, x_2, x_3, x_4 \geq 0$$

$A_1, A_2, \lambda_i \geq 0 \qquad\qquad i = 1, 2, 3, 4.$

Where we have replaced $s_1^{\,2}$ by x_3 and $s_2^{\,2}$ by x_4 and satisfying the complementary slackness condition $\sum \lambda_i x_i = 0.$

The optimum solution to the above LPP shall now be obtained by the phase-I of artificial variable technique.

Thus the starting simplex table is

cost →				0	0	0	0	0	0	0	0	-1	-1	Minimum ratio x_B / x_1
Variable →				x_1	x_2	x_3	x_4	λ_1	λ_2	λ_3	λ_4	A_1	A_2	
Table No	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^7	α^8	α^9	α^{10}	
1	-1	A_1	6	4	4	0	0	1	2	-1	0	1	0	3/2
	-1	A_2	3	4	[6]	0	0	1	3	0	-1	0	1	1/2→
	0	x_3	1	1	1	1	0	0	0	0	0	0	0	1
	0	x_4	4	2	3	0	1	0	0	0	0	0	0	4/3
c_j-Z_j				8	10↑	0	0	2	5	-1	-1	0	0	

In the table 1 we see that x_2 is the variable with most positive entry 10 in c_j-Z_j row and hence will enter the basis and A_2 leaves. The next simplex table is:

cost →				0	0	0	0	0	0	0	0	-1	Minimum ratio x_B / x_1
Variable →				x_1	x_2	x_3	x_4	λ_1	λ_2	λ_3	λ_4	A_1	
Table No	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^7	α^8	α^9	
2	-1	A_1	4	4/3	0	0	0	1/3	0	-1	2/3	1	3
	0	x_2	1/2	[2/3]	1	0	0	1/6	½	0	-1/6	0	3/4→
	0	x_3	1/2	1/3	0	1	0	-1/6	-1/2	0	1/6	0	3/2
	0	x_4	5/2	0	0	0	1	-1/2	-3/2	0	1/2	0	…
c_j-Z_j				4/3↑	0	0	0	1/3	0	-1	2/3	0	

In the table 2 we see that x_1 is the variable with most positive entry 4/3 in c_j-Z_j row and hence will enter the basis and x_2 leaves. The next simplex table is:

cost →				0	0	0	0	0	0	0	0	-1	Minimum ratio x_B / x_1
Variable →				x_1	x_2	x_3	x_4	λ_1	λ_2	λ_3	λ_4	A_1	
Table No	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^7	α^8	α^9	

3	-1	A_1	3	0	-2	0	0	0	-1	-1	1	1	3
	0	x_1	3/4	1	3/2	0	0	1/4	3/4	0	-1/4	0	….
	0	x_3	1/4	0	-1/2	1	0	-1/4	-3/4	0	**1/4**	0	4/3→
	0	x_4	5/2	0	0	0	1	-1/2	-3/2	0	1/2	0	5
c_j-Z_j				0	-2	0	0	0	-1	-1	1↑	0	

The next simplex table is:

cost →				0	0	0	0	0	0	0	0	-1	Minimum ratio x_B / x_1
Variable →				x_1	x_2	x_3	x_4	λ_1	λ_2	λ_3	λ_4	A_1	
Table No	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^7	α^8	α^9	
4	-1	A_1	6	4	4	0	0	**1**	2	-1	0	1	6→
	0	x_1	1	1	1	1	0	0	0	0	0	0	….
	0	λ_4	1	0	-2	4	0	-1	-3	0	1	0	…
	0	x_4	2	0	1	-2	1	0	0	0	0	0	….
c_j-Z_j				0	4	0	0	1↑	2	-1	0	0	

In the table 3 we see that λ_4 is the variable with most positive entry 1 in c_j-Z_j row and hence will enter the basis and x_3 leaves. In the table 4 we see that either x_2 or λ_1 can enter the basis but λ_4 is still in the basis so x_2 cannot enter the basis because of the complementary slackness condition $\lambda_2 x_4 = 0$.

Hence λ_1 enter the basis and A_1 leaves.

Now use entering and leaving variable rules to get the new simplex table:

cost →				0	0	0	0	0	0	0	0	Minimum ratio x_B / x_1
Variable →				x_1	x_2	x_3	x_4	λ_1	λ_2	λ_3	λ_4	
Table No	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^7	α^8	
5	0	λ_1	6	4	4	0	0	1	2	-1	0	6→
	0	x_1	1	1	1	1	0	0	0	0	0	….
	0	λ_4	7	4	2	4	0	0	-1	-1	1	…
	0	x_4	2	0	1	-2	1	0	0	0	0	….
c_j-Z_j				0	0	0	0	0	0	0	0	

In the above table 5, c_j-$Z_j \leq 0$ for all j, therefore solution is optimal.

Thus the optimal solution is $x_1 = 1, x_2 = 0$ and

Minimum $Z = 6x_1 + 3x_2 - 4x_1x_2 - 2x_1^2 - 3x_2^2 = 6 + 0 - 0 - 2 - 0 = 4$.

6.7 Beale's Method

Consider a quadratic programming problem

Max $f(X) = c^T X + \frac{1}{2} X^T QX$

s.t. $AX \{\leq, \geq, or =\} b \;\; and \;\; X \geq 0;$

$where X \in R^n$, A is $m \times n$, b is $m \times 1$, c is $n \times 1$, and Q is a $n \times n$ symmetric matrix.

The procedure to solve a quadratic programming problem by Beale's method as follows:

Step 1: If the given quadratic programming problem is in the minimization form, then convert it into maximization form.

Step 2: Put the given QPP in standard form using slack and surplus variables.

Step 3: Choose arbitrarily any m variables as the basic variables, so that the remaining $n-m$ variables become non-basic. The basic and non-basic variables are

$X_B = \left(X_{B_1}, X_{B_2}, ---- X_{B_m}\right)$ and $X_{NB} = \left(X_{NB_1}, X_{NB_2} ---- X_{NB_{n-m}}\right)$ respectively.

Step 4: Write each basic variable x_{B_i} in terms of non-basic variables $X_{NB_i}'s \;\; (and \; u_i's \;\; if \;\; any)$ using the given (as well as additional, if any) constraints.

Step 5: Write the objective function $f(X)$ in terms of non-basic $X_{NB_i}'s \;\; (and \; u_i's \;\; if \;\; any)$.

Step 6: Evaluate the partial derivatives of $f(X)$ formulated above with respect to the non-basic variables at the point $X_{NB} = 0 \; (and \; u = 0)$. Here three cases arise:

Case-1: If $\left(\frac{\partial f(X)}{\partial x_{NB_k}}\right)_{\substack{X_{NB}=0 \\ u=0}} = 0$ for each $k = 1, 2, \cdots\cdots, n-m$

and $\left(\frac{\partial f(X)}{\partial u_i}\right)_{\substack{X_{NB}=0 \\ u=0}} = 0$ for each i

then the current basic solution is optimal. Go to step 9.

Case-2: If $\left(\frac{\partial f(X)}{\partial x_{NB_k}}\right)_{\substack{X_{NB}=0 \\ u=0}} > 0$ for at least one k

Choose the most positive one. The corresponding non-basic variables will enter the basis.

Case-3: If $\left(\frac{\partial f(X)}{\partial x_{NB_k}}\right)_{\substack{X_{NB}=0\\u=0}} = 0$ for each $k = 1, 2, \cdots\cdots, n-m$

and $\left(\frac{\partial f(X)}{\partial u_i}\right)_{\substack{X_{NB_k}=0\\u=0}} \neq 0$ for sum $i = r$

then introduce a new non-basic variables u_j defined by $u = \frac{1}{2}\left(\frac{\partial f}{\partial u_r}\right)$ and treat u_r as a basic variable (to be ignored later) and go to step 4.

Step 7: Consider $X_{NB_i} = x_k$ be the entering variable identified in step 6 (case-2). Find the minimum of the ratios $min\left[\frac{\alpha_{ho}}{|\alpha_{hk}|}, \frac{\gamma_{ko}}{|\gamma_{kk}|}\right]$ for all basic variables x_h where α_{ho} is the constant term and α_{hk} is the coefficient of x_k in the expression of basic variables x_h when expressed in terms of the non-basic ones, and γ_{ko} is the constant term and γ_k is the coefficient of x_k in $\frac{\partial f}{\partial x_k}$.

(i) If the minimum ratio occurs for some $\frac{\alpha_{ho}}{|\alpha_{ho}|}$ the corresponding basic variables, x_h will leave the basis.

(ii) If the minimum ratio occurs for some $\frac{\gamma_{ko}}{|\gamma_{kk}|}$ the exit criterion corresponds to a non-basic variables. In this case introduce an additional non-basic variables called a free variable defined by

$$u_i = \frac{1}{2}\frac{\partial f}{\partial x_k} \qquad (u_i \text{ is unrestricted})$$

Which relation becomes an additional constraint equation.

Step 8: Now go to step 4 and repeat the procedure until an optimal basic solution is reached.

Step 9: Find the optimal values of X_B and $f(X)$ by setting $X_{NB} = 0$ in their expressions obtained in steps 4 and 5.

Examples

Example.5. Use Beale's method to solve the following NLPP

Max $Z = 2x_1 + 3x_2 - x_1^{\ 2}$

s.t. $x_1 + 2x_2 \le 4$

and $x_1, x_2 \ge 0.$

Solution: The given QPP is

Max $Z = 2x_1 + 3x_2 - x_1^{\ 2}$

s.t. $x_1 + 2x_2 \le 4$

and $x_1, x_2 \ge 0.$

Put the given QPP in standard form using slack variable x_3, we get

Max $Z = 2x_1 + 3x_2 - x_1^{\ 2}$

s.t. $x_1 + 2x_2 + x_3 = 4$

and $x_1, x_2, x_3 \ge 0.$

Now we have x_3 as basis variable and x_1, x_2 as non-basis variable i.e.,

$$X_B = (x_3),\ X_{NB} = (x_1, x_2)$$

Then we have

$$x_3 = 4 - x_1 - 2x_2$$

and $f(X) = 2x_1 + 3x_2 - x_1^{\ 2}$

Differentiate partially f with respect to x_1 and x_2 both sides, we get

$$\left(\frac{\partial f}{\partial x_1}\right)_{X_{NB}=0} = (2 - 2x_1)_{x_1, x_2 = 0} = 2$$

and $\left(\frac{\partial f}{\partial x_2}\right)_{X_{NB}=0} = (3)_{x_1, x_2 = 0} = 3$

Here the most positive x_2 so x_2 enter the basis.

$min\left[\frac{\alpha_{3o}}{|\alpha_{32}|}, \frac{\gamma_{20}}{|\gamma_{22}|}\right] = min\left[\frac{4}{|-2|}, \frac{3}{0}\right] = 2$ and x_3 leave the basis.

Now $X_B = (x_2),\ X_{NB} = (x_1, x_3)$ then we have

$$x_2 = \frac{1}{2}(4 - x_1 - x_3) = 2 - \frac{1}{2}x_1 - \frac{1}{2}x_3$$

and $f = 2x_1 + 3\left(2 - \frac{1}{2}x_1 - \frac{1}{2}x_3\right) - x_1^2$

$$= 2x_1 + 6 - \frac{3}{2}x_1 - \frac{3}{2}x_3 - x_1^2$$

$$= 6 + \frac{1}{2}x_1 - x_1^2 - \frac{3}{2}x_3$$

Differentiate partially above f with respect to x_1 and x_3 both sides, we get

$$\left(\frac{\partial f}{\partial x_1}\right)_{X_{NB}=0} = \left(\frac{1}{2} - 2x_1\right)_{x_1, x_3 = 0} = \frac{1}{2}$$

and $\left(\frac{\partial f}{\partial x_3}\right)_{X_{NB}=0} = \left(-\frac{3}{2}\right) = -\frac{3}{2}$

Here x_1 enter the basis.

$$min\left[\frac{\alpha_{20}}{|\alpha_{21}|}, \frac{\gamma_{10}}{|\gamma_{11}|}\right] = min\left[\frac{2}{|-1/2|}, \frac{1/2}{|-2|}\right] = \frac{1}{4} \text{ which corresponds to } \frac{\gamma_{10}}{|\gamma_{11}|}$$

so x_2 not leave the basis.

Now we introduce a new non-basis variable u_1 such that

$$u_1 = \frac{1}{2}\frac{\partial f}{\partial x_1} = \frac{1}{2}\left(\frac{1}{2} - 2x_1\right) = \frac{1}{4} - x_1$$

$$X_B = (x_1, x_2),\ X_{NB} = (u_1, x_3)$$

$$x_1 = \frac{1}{4} - u_1$$

$$x_2 = 2 - \frac{1}{2}x_1 - \frac{1}{2}x_3 = 2 - \frac{1}{2}\left(\frac{1}{4} - u_1\right) - \frac{1}{2}x_3$$

$$= 2 - \frac{1}{8} + \frac{1}{2}u_1 - \frac{1}{2}x_3 = \frac{15}{8} + \frac{1}{2}u_1 - \frac{1}{2}x_3$$

and $f = 6 + x_1\left(\frac{1}{2} - x_1\right) - \frac{3}{2}x_3$

$$= 6 + \left(\frac{1}{4} - u_1\right)\left(\frac{1}{2} - \frac{1}{4} + u_1\right) - \frac{3}{2}x_3$$

$$= 6 + \left(\frac{1}{4} - u_1\right)\left(\frac{1}{4} + u_1\right) - \frac{3}{2}x_3$$

$$= 6 + \frac{1}{16} - u_1^{\ 2} - \frac{3}{2}x_3$$

$$= \frac{97}{16} - u_1^{\ 2} - \frac{3}{2}x_3$$

Differentiate partially above f with respect to x_3 and u_1 both sides, we get

$$\left(\frac{\partial f}{\partial x_3}\right)_{X_{NB}=0} = \left(-\frac{3}{2}\right)_{x_1, x_3=0} = -\frac{3}{2}$$

and $\left(\frac{\partial f}{\partial u_1}\right)_{X_{NB}=0} = \left(-2u_1\right)_{x_3,u_1} = 0$

This gives the optimal solution.

Hence the optimal solution is $x_1 = \frac{1}{4}, x_2 = \frac{15}{8}$ and Max $Z = \frac{97}{16}$.

Example.6 Use Beale's method to solve the following NLPP

Min $Z = 6 - 6x_1 + 2x_1^{\ 2} - 2x_1x_2 + 2x_2^{\ 2}$

s.t. $x_1 + x_2 \le 2$

and $x_1, x_2 \ge 0.$

Solution: The given QPP is

Min $Z = 6 - 6x_1 + 2x_1^{\ 2} - 2x_1x_2 + 2x_2^{\ 2}$

s.t. $x_1 + x_2 \le 2$

and $x_1, x_2 \ge 0.$

Convert the minimization objective function into a maximization to get

$$Max - Z = f(X) = -6 + 6x_1 - 2x_1^2 + 2x_1x_2 - 2x_2^2$$

Put the given QPP in standard form using slack variable x_3, we get

Min $Z = 6 - 6x_1 + 2x_1^2 - 2x_1x_2 + 2x_2^2$

s.t. $\quad x_1 + x_2 + x_3 = 2$

$and \ \ x_1, x_2 \geq 0.$

Choosing arbitrary x_3 as the basic variable, we have

$$X_B = (x_3) \ \ and \ \ X_{NB} = (x_1, x_2)$$

Expressing X_B and $f(x)$ in terms of X_{NB}, we have

$$x_3 = 2 - x_1 - x_2$$

and $\quad f = -6 + 6x_1 - 2x_1^2 + 2x_1x_2 - 2x_2^2$

Differentiate partially f with respect to x_1 and x_2 both sides, we get

$$\left(\frac{\partial f}{\partial x_1}\right)_{X_{NB}=0} = (6 - 4x_1 + 2x_2)_{\substack{x_1=0\\x_2=0}} = 6$$

and $\quad \left(\frac{\partial f}{\partial x_2}\right)_{X_{NB}=0} = (-2x_1 + 4x_2)_{\substack{x_1=0\\x_2=0}} = 0$

Here the most positive x_1 so x_1 enter the basis.

$min\left[\frac{\alpha_{30}}{|\alpha_{31}|}, \frac{\gamma_{10}}{|\gamma_{11}|}\right] = min\left[\frac{2}{|-1|}, \frac{6}{|-4|}\right] = \frac{3}{2}$ which corresponds to $\frac{\gamma_{10}}{|\gamma_{11}|}$ so x_3 not leave the basis.

Now we introduce a new non-basic free variables u_1, defined by

$$u_1 = \frac{1}{2}\frac{\partial f}{\partial x_1} = 3 - 2x_1 + x_2$$

Now $X_B = (x_3, x_1)$, $X_{NB} = (x_2, u_1)$

$$x_1 = \frac{3}{2} - \frac{1}{2}u_1 + \frac{1}{2}x_2$$

$$x_3 = 2 - x_1 - x_2 = 2 - \left(\frac{3}{2} - \frac{1}{2}u_1 + \frac{1}{2}x_2\right) - x_2 = \frac{1}{2} + \frac{1}{2}u_1 - \frac{3}{2}x_2$$

and $$f = -6 + 6x_1 - 2x_1^{\,2} + 2x_1x_2 - 2x_2^{\,2}$$

$$= -6 + 2x_1(3 - x_1 + x_2) - 2x_2^{\,2}$$

$$= -6 + 2\left(\frac{3}{2} - \frac{1}{2}u_1 + \frac{1}{2}x_2\right)\left(3 - \frac{3}{2} + \frac{1}{2}u_1 - \frac{1}{2}x_2 + x_2\right) - 2x_2^{\,2}$$

$$= -6 + (3 - u_1 + x_2)\left(\frac{3}{2} + \frac{1}{2}u_1 + \frac{1}{2}x_2\right) - 2x_2^{\,2}$$

$$= -6 + \frac{1}{2}(3 + x_2 - u_1)(3 + x_2 + u_1) - 2x_2^{\,2}$$

$$= -6 + \frac{1}{2}\left\{(3 + x_2)^2\right\} - 2x_2^{\,2}$$

$$= -6 + \frac{1}{2}\left\{9 + x_2 - 6x_2 - u_1^{\,2}\right\} - 2x_2^{\,2}$$

$$= -6 + \frac{9}{2} + \frac{x_2^{\,2}}{2} - \frac{6x_2}{2} - \frac{u_1^{\,2}}{2} - 2x_2^{\,2}$$

$$= \frac{-3}{2} - \frac{u_1^{\,2}}{2} - 3x_2 - \frac{3}{2}x_2^{\,2}$$

Differentiate partially above *f* with respect to x_2 and u_1 both sides, we get

$$\left(\frac{\partial f}{\partial x_2}\right)_{X_{NB}=0} = (3 - 3x_2)_{\substack{x_1=0\\u_2=0}} = 3$$

and $$\left(\frac{\partial f}{\partial u_1}\right)_{X_{NB}=0} = (-u_1)_{\substack{x_2=0\\u_1=0}} = 0$$

Here x_2 enter the basis.

$$min\left[\frac{\alpha_{30}}{|\alpha_{32}|}, \frac{\alpha_{10}}{|\alpha_{12}|}, \frac{\gamma_{20}}{|\gamma_{22}|}\right] = min\left[\frac{1/2}{|-3/2|}, \frac{3/2}{|1/2|}, \frac{3}{|-3|}\right] = \frac{1}{3}$$

This implies x_3 will leave the basis and new variables are

$$X_B = (x_1, x_2),\ X_{NB} = (u_1, x_3)$$

$$x_2 = \frac{1}{3} + \frac{1}{3}u_1 - \frac{2}{3}x_3$$

$$x_1 = \frac{3}{2} - \frac{1}{2}u_1 + \frac{1}{2}x_2 = \frac{3}{2} - \frac{1}{2}u_1 + \frac{1}{2}\left(\frac{1}{3} + \frac{1}{3}u_1 - \frac{2}{3}x_3\right) = \frac{3}{2} - \frac{1}{2}u_1 + \frac{1}{6} + \frac{1}{6}u_1 - \frac{1}{3}x_3$$

$$= \frac{5}{3} - \frac{1}{3}u_1 + \frac{1}{3}x_3$$

and $$f = -\frac{3}{2} - \frac{{u_1}^2}{2} + \frac{3}{2}x_2(2 - x_2)$$

$$= -\frac{3}{2} - \frac{{u_1}^2}{2} + \frac{3}{2}\left(\frac{1}{3} + \frac{1}{3}u_1 - \frac{2}{3}x_3\right)\left(2 - \frac{1}{3} - \frac{1}{3}u_1 + \frac{2}{3}x_3\right)$$

$$= -\frac{3}{2} - \frac{{u_1}^2}{2} + \frac{1}{6}(1 + u_1 - 2x_3)(5 - u_1 + 2x_3)$$

$$= -\frac{3}{2} - \frac{{u_1}^2}{2} + \frac{1}{6}\left(5 - u_1 + 2x_3 + 5u_1 - {u_1}^2 + 2x_3u_1 - 10x_3 + 2x_3u_1 - 4{x_3}^2\right)$$

$$= -\frac{3}{2} - \frac{{u_1}^2}{2} + \frac{5}{6} + \frac{4u_1}{6} - \frac{8}{6}x_3 - \frac{{u_1}^2}{6} + \frac{4}{6}x_3u_1 - \frac{4}{6}{x_3}^2$$

$$= -\frac{4}{6} - \frac{4}{6}{u_1}^2 + \frac{4}{6}u_1 - \frac{4}{3}x_3 - \frac{4}{6}{x_3}^2 + \frac{4}{6}x_3u_1$$

$$= -\frac{2}{3} + \frac{2}{3}u_1 - \frac{2}{3}{u_1}^2 - \frac{4}{3}x_3 + \frac{2}{3}x_3u_1 - \frac{2}{3}{x_3}^2$$

Differentiate partially above f with respect to x_3 and u_1 both sides, we get

$$\left(\frac{\partial f}{\partial x_3}\right)_{X_{NB}=0} = \left(-\frac{4}{3} - \frac{2}{3}u_1 - \frac{4}{3}x_3\right)_{\substack{x_3=0\\u_1=0}} = \frac{-4}{3}$$

and $$\left(\frac{\partial f}{\partial u_1}\right)_{X_{NB}=0} = \left(\frac{2}{3} - \frac{4}{3}u_1 + \frac{2}{3}x_3\right)_{\substack{x_3=0\\u_1=0}} = \frac{2}{3}$$

Since $\frac{\partial f}{\partial u_1} \neq 0$ so this solution can be further improved.

Now x_3 does not enter the basis. Thus we introduce another non basic free variable u_2 defined by

$$u_2 = \frac{1}{2}\frac{\partial f}{\partial u_1} = \frac{1}{3} - \frac{1}{3}u_1 + \frac{1}{3}x_3$$

$$X_B = (x_1, x_2, u_1),\ X_{NB} = (x_3, u_2)$$

$$u_1 = \frac{1}{2} - \frac{3}{2}u_2 + \frac{1}{2}x_3$$

$$x_2 = \frac{1}{3} + \frac{1}{3}u_1 - \frac{2}{3}x_3 = \frac{1}{3} + \frac{1}{3}\left(\frac{1}{2} - \frac{3}{2}u_2 + \frac{1}{2}x_3\right) - \frac{2}{3}x_3$$

$$= \frac{1}{2} - \frac{1}{2}u_2 - \frac{1}{2}x_3$$

$$x_1 = \frac{5}{3} - \frac{1}{3}u_1 - \frac{1}{3}x_3 = \frac{5}{3} - \frac{1}{3}\left(\frac{1}{2} - \frac{3}{2}u_2 + \frac{1}{2}x_3\right) - \frac{1}{3}x_3 = \frac{5}{3} - \frac{1}{6} + \frac{1}{2}u_2 - \frac{1}{6}x_3 - \frac{1}{3}x_3$$

$$= \frac{3}{2} + \frac{1}{2}u_2 - \frac{1}{2}x_3$$

$$f = -\frac{2}{3} + \frac{2}{3}u_1 - \frac{4}{3}x_3 + \frac{2}{3}x_3u_1 - \frac{2}{3}u_1^2 - \frac{2}{3}x_3^2$$

$$= -\frac{2}{3} + \frac{2}{3}u_1(1 + x_3 - u_1) - \frac{4}{3}x_3 - \frac{2}{3}x_3^2$$

$$= -\frac{2}{3} + \frac{2}{3}\left(\frac{1}{2} - \frac{3}{2}u_2 + \frac{1}{2}x_3\right)\left(1 + x_3 - \frac{1}{2} + \frac{3}{2}u_2 - \frac{1}{2}x_3\right) - \frac{4}{3}x_3 - \frac{2}{3}x_3^2$$

$$= -\frac{2}{3} + \frac{1}{6}(1 - 3u_2 + x_3)(2 + 2x_3 - 1 + 3u_2 - x_3) - \frac{4}{3}x_3 - \frac{2}{3}x_3^2$$

$$= -\frac{2}{3} + \frac{1}{6}(1 - 3u_2 + x_3)(1 + x_3 + 3u_2) - \frac{4}{3}x_3 - \frac{2}{3}x_3^2$$

$$= -\frac{2}{3} + \frac{1}{6}\left\{(1 + x_3)^2 - 9u_2^2\right\} - \frac{4}{3}x_3 - \frac{2}{3}x_3^2$$

$$= -\frac{2}{3} + \frac{1}{6}\left[1 + x_3^2 + 2x_3 - 9u_2^2\right] - \frac{4}{3}x_3 - \frac{2}{3}x_3^2$$

$$= -\frac{2}{3} + \frac{1}{6} + \frac{x_3^2}{6} + \frac{1}{3}x_3 - \frac{3}{2}u_2^2 - \frac{4}{3}x_3 - \frac{2}{3}x_3^2$$

$$f = -\frac{1}{2} - \frac{1}{2}x_3^2 - \frac{5}{3}x_3 - \frac{3}{2}u_2^2$$

Differentiate partially above f with respect to x_3 and u_2 both sides, we get

$$\left(\frac{\partial f}{\partial x_3}\right)_{X_{NB}=0} = \left(-x_3 - \frac{5}{3}\right) = -\frac{5}{3}$$

and $$\left(\frac{\partial f}{\partial u_2}\right)_{X_{NB}=0} = \left(-\frac{3}{2}u_2\right) = 0$$

This give the optimal solution. Ignoring the free variable u_1, the optimal solution is

$$x_1 = \frac{3}{2}, \quad x_2 = \frac{1}{2}$$

and $$Z = -f = -\left(-\frac{1}{2}\right) = \frac{1}{2}.$$

7.7 Summary

Quadratic programming problems can be solved using two important methods: Wolfe's modified simplex method and Beale's method. Quadratic programming itself is a type of math problem where you try to find the best values for certain things (called decision variables) while following specific rules (constraints). This is often used in finance, engineering, operations research, and machine learning. People apply it to various tasks, like figuring out the best investment portfolio or designing structures. Wolfe's method and Beale's method are two approaches to solve these kinds of problems effectively.

7.8 Terminal Questions

Q.1 Write a short note on quadratic programming problem.

Q.2. State Kuhn-Tucker Conditions. Use them to solve

$$\text{Min } f(x, y, z) = x^2+y^2+z^2+20x+10y$$

s.t. $x \geq 40,$

$x+y \geq 80,$

$x+y+z \geq 120.$

Q.3. Use Wolfe's modified simplex method to solve the Quadratic programming problem:

Max $Z = 2x_1 + x_2 - x_1^{\,2}$

s.t. $2x_1 + 3x_2 \le 6,$

$2x_1 + x_2 \le 4$

and $x_1, x_2 \ge 0$

Q.4. Use Wolfe's modified simplex method to solve the Quadratic programming problem:

Min $f = x_1^2 + 2x_2^2 - 2x_1x_2 - 6x_1$

s.t. $2x_1 - x_2 \le 16$

and $x_1, x_2 \ge 0$

Q.5. Use Beale's method to solve the Quadratic programming problem:

Max $f = 4x_1 + 6x_2 - x_1^2 - 3x_2^2$

s.t. $x_1 + 2x_2 \le 4$

and $x_1, x_2 \ge 0.$

Q.6. Use Beale's method to solve the Quadratic programming problem:

$$Max\ f(X) = \frac{1}{4}(2x_3 - x_1) - \frac{1}{2}(x_1^2 + x_2^2 + x_3^2)$$

s.t. $x_1 - x_2 + x_3 = 1$

and $x_1, x_2, x_3 \ge 0.$

Answer

2. At (40, 40, 40), f = 6000.

3. x_1=2/3, x_2=14/9, max Z=22/9.

4. x_1=0, x_2=1, min f =2.

5. x_1=2, x_2=1, max f =7.

6. x_1=1/8, x_2=0, x_3=7/8 max f =1/64.

UNIT-7: Separable Programming Problem

Structure

7.1 Introduction

In the present unit we shall discuss about the Separable programming problems and with its applications in detailed. Separable programming is a type of mathematical optimization technique that exploits the separability structure within the objective function. In separable programming, the objective function can be expressed as the sum of individual functions, each dependent on a subset of the decision variables. This separability property simplifies the optimization process, as the overall problem is decomposed into smaller, more manageable sub problems.

7.2 Objectives

After reading this unit the learner should be able to understand about:

- the Separable programming problem
- procedure for solving separable programming problem

7.3 Separable Programming

Separable programming is a mathematical programming in which the functions or terms involved in objection function and constraints are separable functions. A function of *n*-variables $f(x_1, x_2, \ldots\ldots x_n)$ is said to be *separable* if it can be written as the sum *n*-functions $f_1(x_1), f_2(x_2), \ldots\ldots\ldots f_n(x_n)$ each is a function of single variable $x_i, i = 1, \ldots\ldots, n.$

Thus $f(x_1, \ldots\ldots, x_n)$ is called a separable function if

$$f(x_1, \ldots\ldots, x_n) = f_1(x_1) + f_2(x_2) + \ldots\ldots\ldots + f_n(x_n) \qquad \ldots.(7.1)$$

Thus the general form of a separable programming is

$$Max(or\ min)Z = f(X) = f_1(x_1) + f_2(x_2) + \ldots\ldots\ldots + f_n(x_n)$$

s.t. $\quad g^1{}_1(x_1) + g^1{}_2(x_2) + \ldots\ldots + g^1{}_n(x_n) \leq = \geq b_1$

$$g^2{}_1(x_1) + g^2{}_2(x_2) + \ldots\ldots + g^2{}_n(x_n) \leq = \geq b_2$$

$$g^m{}_1(x_1) + g^m{}_2(x_2) + \ldots\ldots + g^m{}_n(x_n) \leq = \geq b_m$$

$$x_1, x_2, \ldots\ldots, x_n \geq 0$$

To use the simplex method is solving a separable programming, first the functions involved in the problem are approximated by piecewise linear functions. For this let us consider a function $f(x)$ which is continuous in the interval $[a_1, a_4]$ and $\hat{f}(x)$ be the function shown by dotted lines in the Fig (7.1).

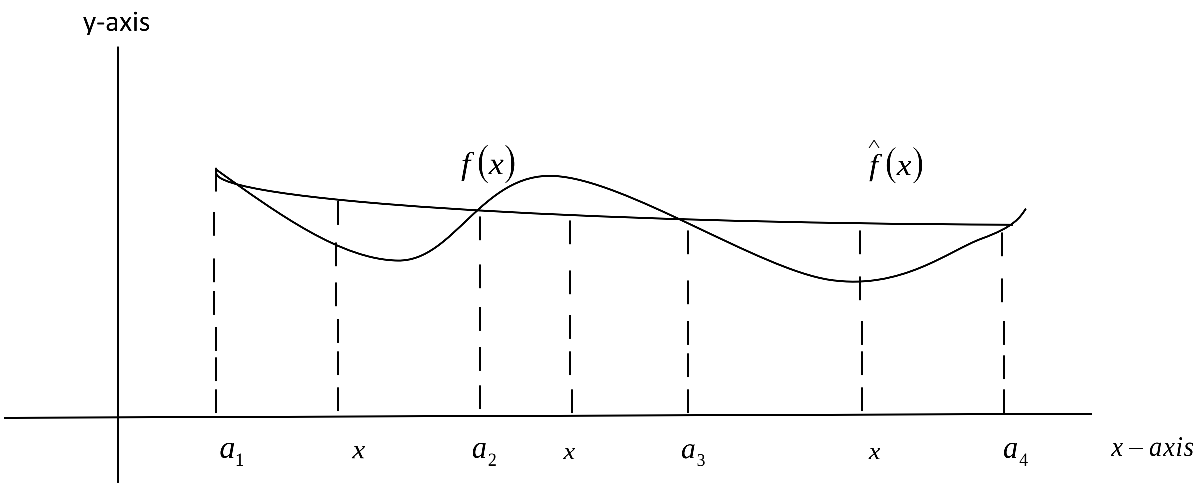


Fig. 7.1 showing the approximation of the function *f*(*x*) by continuous line segments $\hat{f}(x)$

$\hat{f}(x)$ in $[a_1, a_2]$ is a straight line segment whose equation is

$$\hat{f}(x) - f(a_1) = \frac{f(a_2) - f(a_1)}{a_2 - a_1}(x - a_1),\ a_1 \le x \le a_2$$

or $$\hat{f}(x) = f(a_1) + \frac{f(a_2) - f(a_1)}{a_2 - a_1}(x - a_1), \qquad a_1 \le x \le a_2 \qquad (7.2)$$

As $x \in [a_1, a_2]$, $x = \lambda_1 a_1 + (1 - \lambda_1) a_2$, (for same λ_1 such that $0 \le \lambda_1 \le 1$)

or $$x = \lambda_1 a_1 + \lambda_2 a_2, \lambda_1 + \lambda_2 = 1,\ \lambda_1, \lambda_2 \ge 0 \qquad (7.3)$$

or $$x - a_1 = \lambda_2 (a_2 - a_1) \qquad .....(7.4)$$

Now using equation (7.4) from (7.2), we get

$$\hat{f}(x) = \lambda_1 f(a_1) + \lambda_2 f(a_2) \qquad ....(7.5)$$

Similarly for interval $[a_2, a_3]$, we have

$$\hat{f}(x) = \lambda_2 f(a_2) + \lambda_3 f(a_3),\ x \in [a_2, a_3] \qquad (7.6)$$

or $$x = \lambda_2 a_2 + \lambda_3 a_3 \qquad ....(7.7)$$

$$\hat{f}(x) = \lambda_3 f(a_3) + \lambda_4 f(a_4) \qquad ....(7.8)$$

$$x = \lambda_3 a_3 + \lambda_4 a_4,\ (x \in [a_3, a_4]) \qquad ....(7.9)$$

Combining Equation (7.3) and Equations (7.5) to (7.8), we have

$$\left.\begin{aligned} &\hat{f}(x) = \lambda_1 f(a_1) + \lambda_2 f(a_2) + \lambda_3 f(a_3) + \lambda_4 f(a_4) \\ &x = \lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 + \lambda_4 a_4, \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1 \ and\ \lambda_i's \ge 0,\ i = 1,2,3,4 \end{aligned}\right\} \quad .... (7.10)$$

With restrictions (i) at most two $\lambda_i's > 0$ with rest zero

and (ii) only adjacent $\lambda_i's$ are allowed to be positive. .... (7.11)

Conditions in (7.11) are called restricted basis conditions.

In above derivation $\hat{f}(x)$ is the piecewise linear approximation of a non-linear function $f(x)$ in the interval $[a_1, a_4]$ and $\lambda_i's$ are called the weights and $a_i's$ are called breaking points of the i^{th} interval. Thus to solve an NLPP, we shall denote the breaking points of the i^{th} variable x_i be a_i^1, a_i^2 …… and the weight associated with K^{th} breaking points of the i^{th} variable by λ_i^K.

Note.1. Separable programming gives the approximate solution of the problem

Note.2. Greater the number of breaking points, better will be the approximation.

7.4 Procedure for Solving Separable Programming Problem

In solving a separable programming problem we use the following steps:

Step-1. Identify the terms which are not linear in the objective function and the constraints.

Step-2. Find the range of the variable x_i in which terms are not linear and divide it into integer breaking points (for simplified calculation) a_i^K's (k^{th} breaking points) and associate weights λ_i^K's with these points.

Step-3. Prepare a table of calculated values of the functions at the breaking points.

Step-4. Write the given problem into the form an approximated LPP.

Step-5. Solve by simplex method the approximated LPP using restricted basis condition till optimality.

Examples

Example.1. Using separable programming to solve the following NLPP

$$Max.\ Z = x_1 + x_2^4$$

s.t. $3x_1 + 2x_2^2 \le 9$ ...(7.12)

and $x_1, x_2 \ge 0$

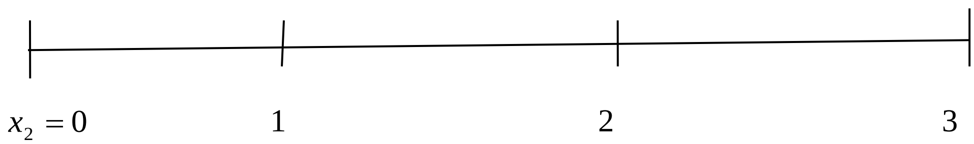
Solution: Here $f_1(x_1) = x_1,\ f_2(x_2) = x_2^4,\ g_1^1(x_1) = 3x_1,\ g_2^2(x_2) = 2x_2^2$

Here $f_2(x_2)$ and $g_2^2(x_2)$ are not linear.

From constraint (7.12), we have

$$x_2 < 3.$$

Let us linearize $f_2(x_2)$ and $g_2^2(x_2)$ in $[0,\ 3]$.



k	a_2^K	$f_2(a_2^K)$	$g_2^2(a_2^K)$
1	0	0	0
2	1	1	2
3	2	16	8
4	3	81	18

Thus we have

$$f_2(x_2) \approx \hat{f}(x_2) = \lambda_2^1 f_2(a_2^1) + \lambda_2^2 f_2(a_2^2) + \lambda_2^3 f(a_2^3) + \lambda_2^4 f(a_2^4)$$

$$= \lambda_2^2 + 16\lambda_2^3 + 81\lambda_2^4$$

and $g_2^1(x_2) \approx \hat{g}_2^1(x_2) = 2\lambda_2^2 + 8\lambda_2^3 + 18\lambda_2^4$

Thus the given problem is approximated to LPP

Max $Z = f_1(x_1) + f_2(x_2) \approx x_1 + \lambda_2^2 + 16\lambda_2^3 + 81\lambda_2^4$

$s.t\ 3x_1 + 2\lambda_2^2 + 8\lambda_2^3 + 18\lambda_2^4 \le 9$

$$\lambda_2^1 + \lambda_2^2 + \lambda_2^3 + \lambda_2^4 = 1, x_1 \ge 0,\ \lambda_2^K \ge 0,\ K = 1,2,3,4$$

and at most two consecutive $\lambda_2^K > 0$ (restricted basis conditions).

The standard form of above LPP is

$Max\ Z = x_1 + \lambda_2^2 + 16\lambda_2^3 + 81\lambda_2^4$

$s.t\ \ 3x_1 + 2\lambda_2^2 + 8\lambda_2^3 + 18\lambda_2^4 + x_3 = 9$

$$\lambda_2^{\,1}+\lambda_2^{\,2}+\lambda_2^{\,3}+\lambda_2^{\,4}=1$$

$x_1 \geq 0, \lambda_2^{\,K} \geq 0, K=1,2,3,4,\;\; x_3 \geq 0$ is slack variable.

Now $\lambda_2^{\,1}$ can enter, then $\lambda_2^{\,3}$ leaves, but $\lambda_2^{\,1}$ and $\lambda_2^{\,4}$ are not consecutive so $\lambda_2^{\,1}$ cannot enter. Next promising candidate is $\lambda_2^{\,2}$ to enter.

But if $\lambda_2^{\,2}$ enters, then again $\lambda_2^{\,3}$ leaves which is not possible as $\lambda_2^{\,2}$ and $\lambda_2^{\,4}$ are not consecutive.

So table 3 gives the approximate optimal solution of the problem and is given by

$$\lambda_2^{\,3}=\frac{9}{10}, \lambda_2^{\,4}=\frac{1}{10} \;\; and \;\; MaxZ = 22.5.$$

or $x_1 = 0, x_2^{\,4} \approx \lambda_2^{\,2}+16\lambda_2^{\,3}+81\lambda_2^{\,4} = 0+\frac{144}{10}+\frac{81}{10}=\frac{255}{10}=22.5$

Cost →				1	1	16	81	0	0	Minimum Ratio
Variable →				x_1	$\lambda_2^{\,2}$	$\lambda_2^{\,3}$	$\lambda_2^{\,4}$	x_3	$\lambda_2^{\,1}$	$\frac{X_{B_i}}{\alpha_i^j}$
Table No.	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^5	α^4	α^6	
1	0	x_3	9	3	2	8	18	1	0	$\frac{9}{8}$
	0	$\leftarrow \lambda_2^{\,1}$	1	0	1	[1]	1	0	1	$\frac{1}{1}=1 \rightarrow$
	$c_j - Z_j$			1	1	16↑	81	0	0	
2	0	$\leftarrow x_3$	1	3	-6	0	[10]	1	-8	$\frac{1}{10} \rightarrow$
	0	$\lambda_2^{\,3}$	1	0	1	1	1	0	1	$\frac{1}{1}=1$
	$c_j - Z_j$			1	-15	0	65↑	0	-16	

3	0	$\lambda_2^{\,4}$	1/10	3/10	-6/10	0	1	1/10	-8/10	
	0	$\lambda_2^{\,3}$	9/10	-3/10	16/10	1	0	-1/10	18/10	
	c_j-Z_j			-18.5	24	0	0	-65/10	36	

Note: If a variable in any separable programming is non-linear even at one place in objective function or constraint(s), then it has also to be linearized at a place where it is linear.

Example.2 Using separable programming to solve the following NLPP

$$Max\ Z = x_1 + x_2^{\,4}$$

$$s.t\ 3x_1^{\,2} + 2x_2^{\,2} \le 9, \qquad x_1, x_2 \ge 0$$

Solution: Here

$$f_1(x_1) = x_1 \qquad f_2(x_2) = x_2^{\,4},\ \ g_1^{\,1}(x_1) = 3x_1, \qquad g_2^{\,1}(x_2) = 2x_2^{\,2} \qquad \text{.... (7.13)}$$

From the constraint $3x_1^{\,2} + 2x_2^{\,2} \le 9,$ we have $x_1 < 2 \quad and \quad x_2 < 3.$

Let $a_1^{\,1} = 0,\ \ a_1^{\,2} = 1,\ \ a_1^{\,3} = 2,\ \ and\ \ a_2^{\,1} = 0,\ \ a_2^{\,2} = 1,\ \ a_2^{\,3} = 2\ \ and\ \ a_2^{\,4} = 3$ be the breaking points for variables $x_1\, and\ x_2$ respectively.

We calculate the values of $f_i's\ and\ \ g_i's$ as per the following tables

K	$a_1^{\,K}$	$f_1(a_1^{\,K})$	$g_1^{\,1}(a_1^{\,K})$
1	0	0	0
2	1	1	3
3	2	2	12

Table-7.1

Using calculated values in table 7.1 & 7.2 from Equation (7.13), we have

$$f_1(x_1) \approx \lambda_1^{\,1} \times 0 + \lambda_1^{\,2} \times 1 + \lambda_1^{\,3} \times 2 = \lambda_1^{\,2} + 2\lambda_1^{\,3}$$

$$f_2(x_2) = \lambda_2^{\,2} + 16\lambda_2^{\,3} + 81\lambda_2^{\,4}$$

$$g_1^{\,1}(x_1) = 3\lambda_1^{\,2} + 12\lambda_1^{\,3}$$

$$g_2^{\,1}(x_2) = 2\lambda_2^{\,2} + 8\lambda_2^{\,3} + 18\lambda_2^{\,4}$$

K	$a_2^{\,K}$	$f_2(a_2^{\,K})$	$g_2^{\,1}(a_2^{\,K})$
1	0	0	0
2	1	1	2
3	2	16	8
4	3	81	18

Table-7.2

Thus the given problem approximated in the form of linear programming problem is

$$Max \;\; Z = \lambda_1^{\,2} + 2\lambda_1^{\,3} + \lambda_2^{\,2} + 16\lambda_2^{\,3} + 81\lambda_3^{\,4}$$

$$s.t \qquad 3\lambda_1^{\,2} + 12\lambda_1^{\,3} + 2\lambda_2^{\,2} + 8\lambda_2^{\,3} + 18\lambda_2^{\,4} \le 9$$

$$\lambda_1^{\,1} + \lambda_1^{\,2} + \lambda_1^{\,3} = 1$$

$$\lambda_2^{\,1} + \lambda_2^{\,2} + \lambda_2^{\,3} + \lambda_2^{\,4} = 1$$

and the restricted basis conditions all $\lambda_i^{\,j}{'s} \ge 0 \;\; i = 1,2; \; j = 1,2,3,4$ they should be consecutive. The standard form is

$$Max \;\; Z = \lambda_1^{\,2} + 2\lambda_1^{\,3} + \lambda_2^{\,2} + 16\lambda_2^{\,3} + 81\lambda_3^{\,4}$$

$$s.t \; 3\lambda_1^{\,2} + 12\lambda_1^{\,3} + 2\lambda_2^{\,2} + 8\lambda_2^{\,3} + 18\lambda_2^{\,4} = 9$$

$$\lambda_1^{\,1} + \lambda_1^{\,2} + \lambda_1^{\,3} = 1$$

$$\lambda_2^{\,1} + \lambda_2^{\,2} + \lambda_2^{\,3} + \lambda_2^{\,4} = 1$$

Almost two $\lambda_i^{j}{}'s > 0$ *and* $\lambda_2^{j}{}'s > 0$ and they should be consecutive.

Table No.	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^7	α^8	Minimum Ratio $\frac{X_{B_i}}{\alpha_i^j}$
		Cost →		1	2	1	16	81	0	0	0	
		Variable →		$\lambda^2{}_1$	$\lambda^3{}_1$	$\lambda^2{}_2$	$\lambda^3{}_2$	$\lambda^4{}_2$	x_3	$\lambda^1{}_1$	$\lambda^1{}_2$	
1	0	x_3	9	3	12	2	8	18	1	0	0	$\frac{9}{8}$
	0	$\lambda^1{}_1$	1	1	1	0	0	0	0	1	0	…
	0	$\leftarrow \lambda^1{}_2$	1	0	0	1	[1]	1	0	0	1	$\frac{1}{1}=1$ →
		c_j-Z_j		1	2	1	16 ↑	81	0	0	0	
2	0	$\leftarrow x_3$	1	3	12	-6	0	[10]	1	0	-8	$\frac{1}{10}$ →
	0	$\lambda^1{}_1$	1	1	1	0	0	0	0	1	0	-----
	16	$\lambda^3{}_2$	1	0	0	1	1	1	0	0	1	$1/1=1$
		c_j-Z_j		1	2	-15	0	65↑	0	0	-16	
3	81	$\lambda^4{}_2$	1/10	3/10	12/10	-6/10	0	1	1/10	0	-8/10	
	0	$\lambda^1{}_1$	1	1	1	0	0	0	0	1	0	
	16	$\lambda^3{}_2$	9/10	-3/10	-12/10	16/10	1	0	-1/10	0	18/10	
		c_j-Z_j		-37/2	-76	24	0	0	-13/2	0	36	

In table 1, λ_2^{4} is the candidate to enter the basis but it cannot enter because of restriction basis conditions.

Next promising candidate to enter the basis is $\lambda_2^{\,3}$, when $\lambda_2^{\,3}$ enters the basis $\lambda_2^{\,1}$ leaves. In table 2, $\lambda_2^{\,4}$ enter and x_3 leaves.

In table 3, $\lambda_2^{\,1}$ and $\lambda_2^{\,2}$ are the candidates to enter but they cannot enter as in both the cases restriction basis conditions are not satisfied.

Hence the approximate optimal solution is

$$\lambda_1^{\,1} = 1,\ \lambda_2^{\,3} = \frac{9}{10},\ \lambda_2^{\,4} = \frac{1}{10}\ .$$

$$x_1 = 0\lambda_1^{\,1} + 1\lambda_1^{\,2} + 2\lambda_1^{\,3} = 0$$

$$x_2 = 0\lambda_2^{\,1} + 1.\lambda_2^{\,2} + 2\lambda_2^{\,3} + 3\lambda_2^{\,4} = 2\times\frac{9}{10} + 3\times\frac{1}{10} = \frac{18+3}{10} = 2.1$$

i.e., $x_1 = 0, x_2 = 2.1$ and $Max\ \ Z = \frac{45}{2}.$

7.4 Summary

Separable programming is a mathematical programming in which the functions or terms involved is objection function and constraints are separable functions. Separable programming is particularly beneficial in certain applications, such as machine learning, where the optimization of complex models involves managing numerous variables and functions. By exploiting the separability of the objective function, optimization algorithms can efficiently navigate the solution space.

7.5 Terminal Questions

Q.1. Use Separable programming to solve the Quadratic programming problem:

$Max\ f(X) = 3x_1 + 2x_2$

s.t. $\quad 4x_1^{\,2} + x_2^{\,2} \le 16$

and $x_1, x_2 \ge 0.$

Q.2. Use Separable programming to solve the Quadratic programming problem:

$Max\ f(X) = 16 - 2(x_1 - 3)^2 - (x_2 - 7)^2$

s.t. $x_1^{\ 2} + x_2 \le 16$

and $x_1, x_2 \ge 0.$

Answer

1. $x_1 = 1$, $x_2 = 24/7$, max $f = 69/7$.

2. $x_1 = 3$, $x_2 = 7$, max $f = 16$.

Master of Science
PGMM -102N
Classical Optimization Techniques

**U. P. Rajarshi Tandon
Open University**

Block

3 Dynamic Programming Problems

Unit- 8

Introduction to Dynamic Programming

Unit- 9

Applications of Dynamic Programming

Block-3

Dynamic Programming Problems

Dynamic programming is a method for efficiently solving a broad range of search and optimization problems that exhibit the property of overlapping subproblems and optimal substructure. In dynamic programming, the key idea is to break down a complex problem into simpler, overlapping subproblems and solve each subproblem only once, storing the solutions to subproblems in a table to avoid redundant computations. One classic example of a dynamic programming problem is the Fibonacci sequence calculation. The recursive approach to calculating Fibonacci numbers has exponential time complexity due to redundant computations. However, dynamic programming allows us to store the results of previously solved subproblems and reuse them to compute the solution for larger subproblems.

Dynamic programming is applicable to various problems beyond Fibonacci, including shortest path problems, sequence alignment, and many optimization problems. It provides an efficient way to solve problems by breaking them down into smaller, solvable subproblems and storing the results for reuse. Bellman's Principle of Optimality is a key concept in dynamic programming, providing a foundation for solving complex problems by iteratively breaking them down into simpler subproblems. This principle is widely used in various fields, including operations research, control theory, and artificial intelligence, where dynamic programming techniques are applied to find optimal solutions to problems with overlapping subproblems and optimal substructure.

In the eighth unit , we shall discussed about the dynamic programming, Bellman’s principle of optimality, dynamic programming solution procedure using forward and backward techniques. Solution of linear programming problem using dynamic programming and applications of dynamic programming problem are also discussed in details in unit ninth.

UNIT-8: Introduction to Dynamic Programming

Structure

8.1 Introduction

In this chapter we shall be considering the problems which will be solved by breaking the problem into different parts called stages. Assuming that it is possible to split the problem into different stages and the problem of each stage is solvable easily. For this we shall consider problems in the form of various models. Using Bellman's principle of optimality we have obtained the solution of these models. As applications of dynamic programming in solving travelling salesman problem, have been discussed in details.

The term "programming" in dynamic programming doesn't refer to computer programming but is used in the mathematical sense, indicating a plan or strategy. Dynamic programming is a problem-solving technique used to tackle complex problems by breaking them down into simpler subproblems. The key idea is to solve each subproblem only once and store the solutions, avoiding redundant computations.

The technique is particularly effective for optimization problems, where the goal is to find the best solution among a set of feasible solutions. Dynamic programming is often applied to problems in which the same subproblems are solved multiple times, and the solutions to these subproblems can be reused to solve the overall problem more efficiently.

8.2 Objectives

After reading this unit the learner should be able to understand about the:

- Dynamic programming
- Bellman’s principle of optimality
- dynamic programming solution procedure using forward and backward techniques
- shortest route problem using dynamic programming

8.3 Dynamic Programming

There arise many occasions in business and industry where many decisions are to be taken to tackle a problem. It is not possible to take all these decisions simultaneously but are to be taken one by one at different stages. The problems where such situations of taking decision at different stages arise are called multistage decision problems. A multistage decision problem can be solved by a mathematical technique called “Dynamic Programming” based on the principle developed by Richard Bellman in early 1950s called Bellman’s principle of optimality.

When we solve a problem by dynamic programming the following terms will be used very frequently.

Stage

In dynamic programming, the problem is divided into various sub problems. Each sub problem is known as a stage. At each stage various decision alternates are available and out of these the optimum decision is selected.

State

At each stage, again a number of choices are available, known as states. For each such choice alternatives are there, known as state variables. By analyzing the effect of these state variables the optimum decision is taken for its use in the next stage.

Optimal Decision Rule

This rule specifies the decision to be made as a function of state variable and the stage number. This is also a policy that transforms the state at a given stage into a state associated with the next stage.

Optimal Policy

The rule that optimizes the value of the objective function at a particular stage is known as optimal policy.

8.4 Bellman’s Principle of Optimality

Bellman's Principle of Optimality is a fundamental concept in dynamic programming, named after mathematician and computer scientist Richard Bellman. The principle states that an optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision. In simpler terms, the principle asserts that an optimal solution to a problem can be constructed by breaking it down into smaller subproblems, and the optimal solution to each subproblem contributes to the overall optimal solution.

"Whatever the initial state and the initial decisions are the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision". As seen in the light of dynamic programming, a multi-stage process is one in which a number of single-stage processes considered as function of single variable are connected so that the optimum solution of one stage is used in the succeeding stage.

Thus according to Bellman's principle of optimality, the problem of decision making must be split up in different stages and the optimal decisions be taken sequentially. The various iterations of decision making process are shown below:

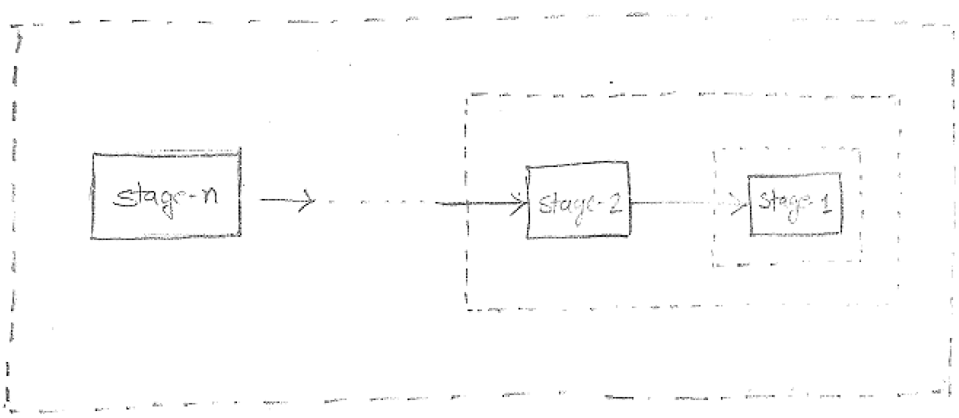


Fig.8.1 Division of Optimization Problem in Stages

Dividing the problem of optimize Z=f(X) in a sequential manner as

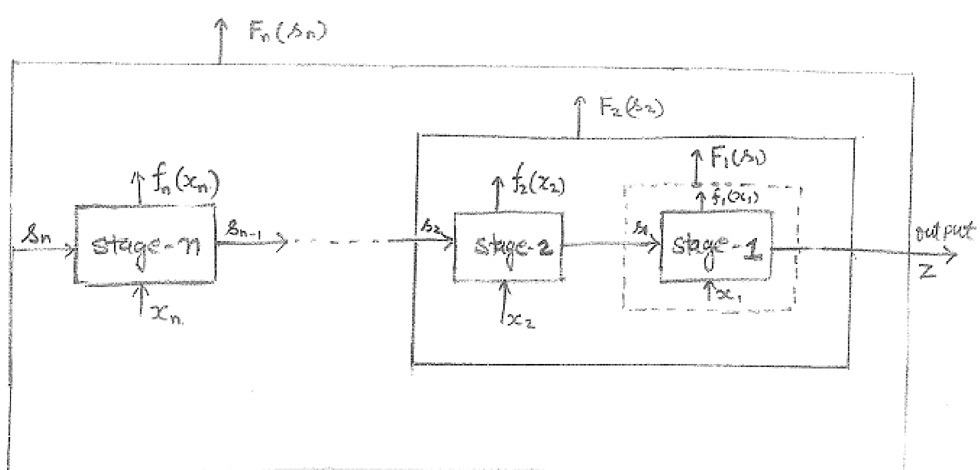


Fig.8.2 Optimization Problem in Stages with Optimum returns

Procedure discussed above is called backward recursion, because the stage transformation function is of the type $s_i = T_i(s_{i+1}, x_{i+1})$. This is convenient when s_n is specified. When s_1 is specified, then it would be convenient to reverse the direction. This is called forward recursion.

8.5 Dynamic Programming Algorithm for solving Shortest Route Problem

The Shortest Route Problem is a classic optimization problem, and dynamic programming provides an effective algorithmic approach to solving it. One common application of dynamic programming for this problem is the Bellman-Ford algorithm.

A variety of problems can be solved by dynamic programming using the principle of optimality. Any problem which can be broken into stages and decision can be taken stage wise using the principle of optimality, can be solved by dynamic programming. Computations in dynamic programming while solving shortest route problem, mathematically can be expressed as follows. Let $f_i(x_i)$ be the shortest distance to node x_i at the stage i.

Define $d(x_{i-1}, x_i)$ = distance from node x_{i-1} to node x_i. Compute f_i from f_{i-1}. Using the following recursive equation

$$f_i(x_i) = \underset{all\ \ feasible\ \ routes\ (x_{i-1}, x_i)}{Min} \left[d\left(x_{i-1}, x_i\right) + f_{i-1}\left(x_{i-1}\right) \right]$$

Where i =1, 2, 3, …, n (n= number of stages)

This is known as forward dynamic programming. The backward recursive equation is

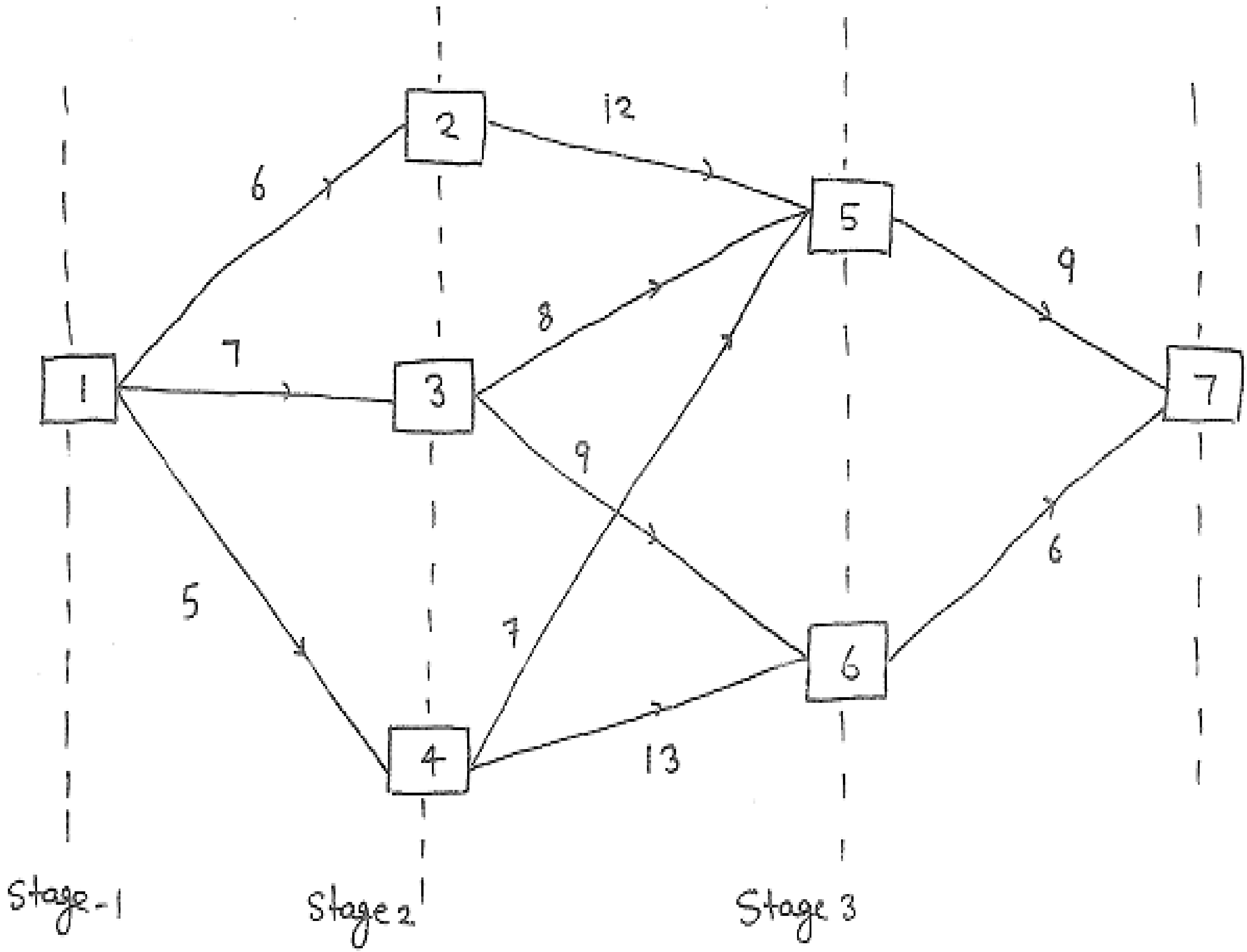
$$f_i(x_i) = \underset{all\ \ feasible\ \ routes\ (x_i, x_{i+1})}{Min} \left[d\left(x_i, x_{i+1}\right) + f_{i+1}\left(x_{i+1}\right) \right]$$

Where i =1, 2, 3, …, n (n= number of stages).

The Bellman-Ford algorithm is a dynamic programming approach used to solve the Shortest Route Problem in graphs. This algorithm efficiently finds the shortest paths from a designated source vertex to all other vertices in a weighted graph, even when negative-weight edges are present, as long as there are no negative cycles.

Examples

Example.1. A travelling salesman has to go from city 1 to city 7. The distances between various cities are given in the following diagram:



Using dynamic programming, find the minimum distance covered by the salesman. Also find the optimal path.

Solution. To solve this problem by dynamic programming, first we decompose it into stages. Then we carry out the computation for each stage separately. The optimum result at each stage is used in finding the optimum at the succeeding stage.

Here we are solving this problem by backward recursive equation which is given below:

$$f_i(x_i) = \underset{all\ feasible\ routes\,(x_i,\, x_{i+1})}{Min} \left[d(x_i,\, x_{i+1}) + f_{i+1}(x_{i+1}) \right], \quad i = 1, 2, 3, \ldots\ldots\ldots, n$$

(n is the number of stages).

We have $f_4(x_4)=0$, for $x_4=7$. The associated order of computation is $f_3 \rightarrow f_2 \rightarrow f_1$.

Stage-3. Node 7 (x_4=7) is connected to node 5 and 6 (x_3 =5 or 6) with exactly one route each. Results can be summarized as

	$d(x_3, x_4)$	Optimum Solution	
x_3	x_4=7	$f_3(x_3)$	x_4^*
5	9	9	7
6	6	6	7

Stage-2. Route (2, 6) is not a feasible alternative as it does not exist. The optimum solution at stage 2 reads as follows:

	$d(x_2, x_3)+ f_3(x_3)$		Optimum Solution	
x_2	x_3=5	x_3=6	$f_2(x_2)$	x_3^*
2	12+9=21	---	21	5
3	8+9=17	9+6=15	15	6
4	7+9=16	13+6=19	16	5

If you are in cities 2 or 4 the shortest route passes through city 5 and if you are in city 3, the shortest route passes through city 6.

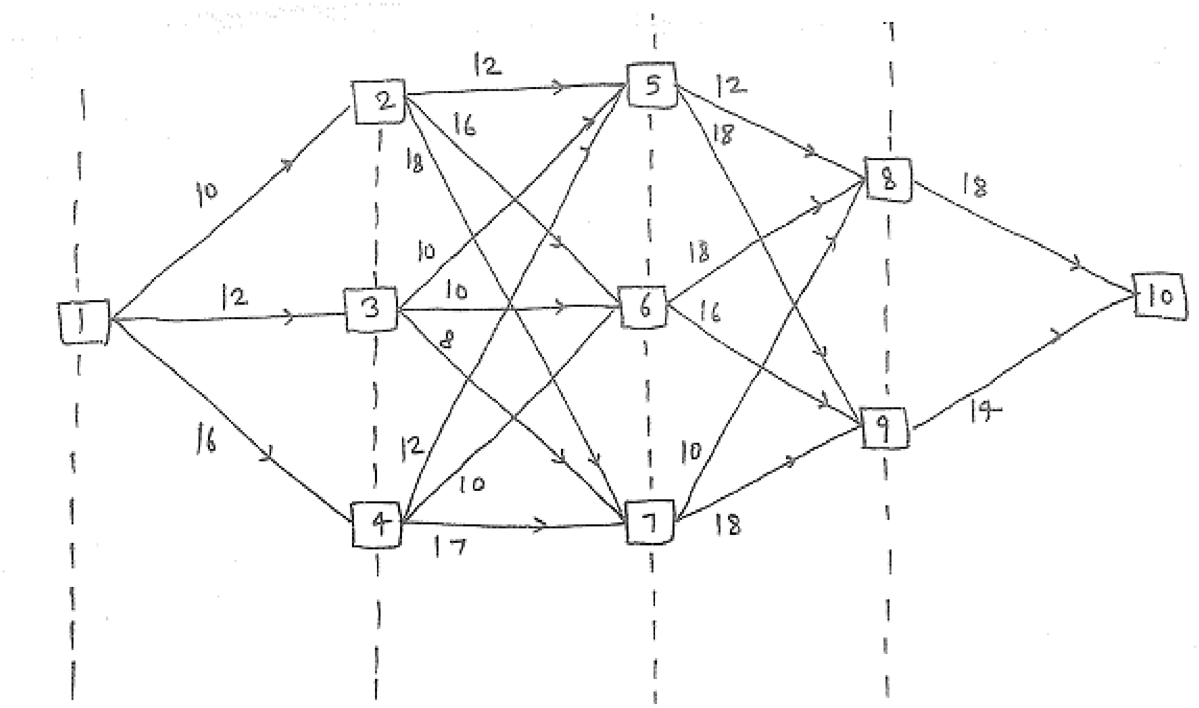
Stage-1. From node 1, we have three alternative routes (1, 2), (1, 3) and (1, 4). Using $f_2(x_2)$ from stage 2, we can compute the following table:

	$d(x_1, x_2)+ f_2(x_2)$			Optimum Solution	
x_1	x_2=2	x_2=3	x_2=4	$f_2(x_2)$	x_2^*
1	6+21=27	7+15=22	5+16=21	21	4

The optimum distance is 21 and optimum route is $1 \rightarrow 4 \rightarrow 5 \rightarrow 7$.

Note. We can solve this problem by exhaustively enumerating all the routes between nodes 1 to 7 (there are five such routes). However in a large network, exhaustive enumeration is not efficient computationally.

Example.2. A travelling salesman has to go from city 1 to city 10. The distances between various cities are given in the following diagram:

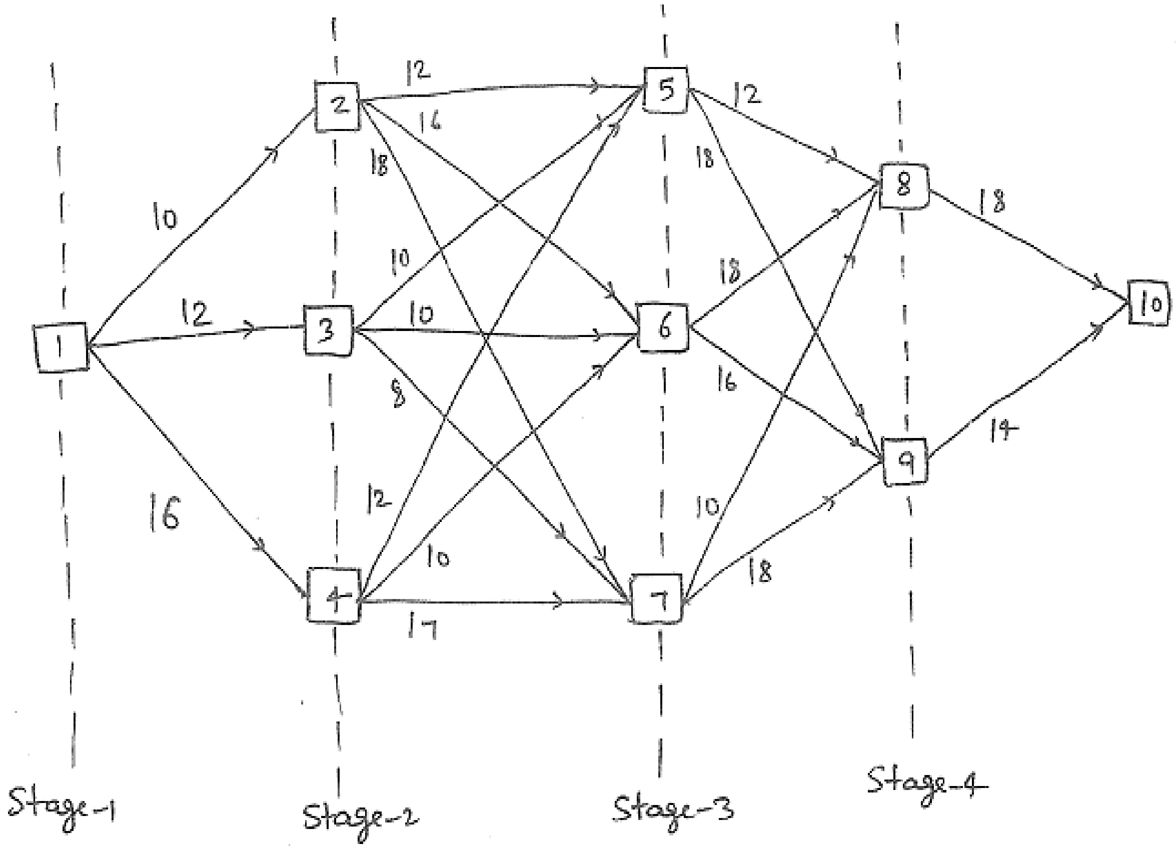


Using dynamic programming, find the minimum distance covered by the salesman. Also find the optimal path.

Solution. We can solve the given dynamic programming problem either using forward dynamic programming or backward dynamic programming.

We solve the problem using backward dynamic programming.

First we divide the problem into four stages as per the following diagram:



Stage-4. In the stage 4 there are two states 8 and 9, and decision variables are 18 and 14.

To State → / From State ↓	Decision 10	Optimum Decision	Optimum Distance
8	18	10	18
9	14	10	14

Stage-3. In the stage 3 there are three states 5, 6 and 7, and decision variables are (12, 18), (18, 16) and (10, 18) respectively.

To State → / From State ↓	Decision		Optimum Decision	Optimum Distance
	8	9		
5	12+18=30	18+14=32	8	30
6	18+18=36	16+14=30	9	30
7	10+18=28	18+14=32	8	28

Stage-2. In the stage 2 there are three states 2, 3 and 4, and decision variables are (12, 16, 18), (10, 10, 8) and (12, 10, 17) respectively.

To State → / From State ↓	Decision			Optimum Decision	Optimum Distance
	5	6	7		
2	12+30=42	16+30=46	18+28=46	5	42
3	10+30=40	10+30=40	8+28=36	7	36
4	12+30=42	10+30=40	17+28=45	6	40

Stage-1. In the stage 1 there is one state 1 and decision variables is (8, 12, 14).

To State → / From State ↓	Decision			Optimum Decision	Optimum Distance
	2	3	4		
1	10+42=52	12+36=48	16+40=56	3	48

Hence optimum (shortest) route is 1→3→7→8→10 and the optimum (minimum) distance is 48 units.

8.6 Dynamic Programming using Calculus Method

Here we shall discuss the problems which can be solved using dynamic programming techniques through calculus method. We illustrate the following cases:

Case-1. When both the objective function and the constraint are in additive form.

Let us consider the problem

$$\text{Max (or Min) } Z = f_1(x_1)+f_2(x_2)+\ldots\ldots\ldots+f_n(x_n) = \sum_{i=1}^{n} f_i x_i$$

$$\text{s.t.} \quad a_1 x_1 + a_2 x_2 + a_3 x_3 + \ldots\ldots\ldots + a_n x_n \geq \text{ or } = \text{ or } \leq b$$

$$\text{or } \sum_{i=1}^{n} a_i x_i \geq or = or \leq b$$

$$\text{and } (a_1, a_2, a_3, \ldots, a_n) \geq 0, (x_1, x_2, x_3, \ldots, x_n) \geq 0 \text{ and } b>0$$

$$\text{or } a_i \geq 0, x_i \geq 0 \text{ and } b>0.$$

To solving this type of problem, first we define the state variables s_i's as

$$s_n = a_1 x_1 + a_2 x_2 + a_3 x_3 + \ldots\ldots\ldots + a_n x_n \geq \text{ or } = \text{ or } \leq b$$

$$s_{n-1} = a_1 x_1 + a_2 x_2 + a_3 x_3 + \ldots\ldots\ldots + a_{n-1} x_{n-1} = s_n - a_n x_n = T_{n-1}(s_n, x_n)$$

$$s_{n-2} = a_1 x_1 + a_2 x_2 + a_3 x_3 + \ldots\ldots\ldots + a_{n-2} x_{n-2} = s_{n-1} - a_{n-1} x_{n-1} = T_{n-2}(s_{n-1}, x_{n-1})$$

… … … … … … …

… … … … … … …

$$s_i = a_1 x_1 + a_2 x_2 + a_3 x_3 + \ldots\ldots\ldots + a_i x_i = s_{i+1} - a_{i+1} x_{i+1} = T_i(s_{i+1}, x_{i+1})$$

… … … … … … …

… … … … … … …

$$s_2 = a_1 x_1 + a_2 x_2 = s_3 - a_3 x_3 = T_2(s_3, x_3)$$

$$s_1 = a_1 x_1 = s_2 - a_2 x_2 = T_1(s_2, x_2)$$

Where $s_i = T_i(s_{i+1}, x_{i+1})$, i=1, 2, 3,…., n-1 are called state transformation functions connecting different states. The recursive relations are

$$F_i(s_i) = \text{Max (or Min)} [f_i (x_i)+ F_{i-1}(s_{i-1})]; \; i=1, 2, 3,\ldots,n.$$

and $F_1(s_1)=f_1(x_1)$.

Now we can start with $F_1(s_1)$ and recursively optimize to get $F_2(s_2)$, $F_3(s_3)$, …………, $F_n(s_n)$.

This is forward dynamic programming. Also we can solve this type of problem by using backward dynamic programming technique.

Case-2. When the objective function is in additive form and the constraint is in multiplicative form.

Case-3. When the objective function is in multiplicative form and the constraint is in additive form.

Examples

Example.3. Solve the following problem using dynamic programming:

$$\textbf{Min } Z= x_1^2+x_2^2+x_3^2$$

$$\textbf{s.t.} \quad x_1 +x_2 +x_3 \geq 15$$

$$\textbf{and} \quad x_1, x_2, x_3 \geq 0.$$

Solution. We can solve it either using forward dynamic programming or backward dynamic programming. We solve it by using backward dynamic programming.

Backward Dynamic Programming

As there are three variables in the problem, we will have three stages to solve the problem using dynamic programming.

Now we define the state variables as:

$$s_1= x_1 +x_2 +x_3 \geq 15$$

$$s_2= x_2 +x_3= s_1- x_1$$

$$s_3= x_3= s_2- x_2$$

The recursive relations are

$$F_3(s_3) = x_3^2= (s_2- x_2)^2$$

$$F_2(s_2) = \min [x_2^2+ F_3(s_3)]$$

$= \min (x_2^2+x_3^2)$

$= \min [x_2^2+(s_2- x_2)^2]$ with respect to x_2

$F_1(s_1) = \min [x_1^2+ F_2(s_2)]$ with respect to x_1

For minimum of F_2, we have

$$\frac{dF_2}{dx_2} = 0$$

gives $2x_2-2(s_2- x_2) = 0$

$\Rightarrow \quad x_2= \frac{s_2}{2}.$

Also $\frac{d^2F_2}{dx_2^2} = 2+4=4$ which is positive.

Therefore F_2 in minimum at $x_2 = \frac{s_2}{2}$.

$$F_2(s_2) = \left[\left(\frac{s_2^2}{2}\right)^2 + \left(s_2 - \frac{s_2}{2}\right)^2\right] = \frac{s_2^2}{2}.$$

Now $F_1(s_1)= \min\left[x_1^2 + \frac{s_2^2}{2}\right] = \min\left[x_1^2 + \frac{(s_1 - x_1)^2}{2}\right]$

For minimum of F_1, we have

$$\frac{dF_1}{dx_1} = 0$$

gives $2x_1- (s_1- x_1) = 0$

$\Rightarrow \quad x_1= \frac{s_1}{3}.$

Also $\frac{d^2F_1}{dx_1^2} = 2+1=3$ which is positive.

Therefore F_1 in minimum at $x_1=\frac{s_1}{3}$.

Thus $F_1(s_1)= \frac{s_1^2}{3}, \quad s_1 \geq 15.$

Now Min Z= min $[x_1^2+x_2^2+x_3^2]$

= min $F_1(s_1)$

$$= \frac{(15)^2}{3} = 75 \quad for \quad s_1 = 15.$$

Here $x_1 = (s_1/3) = 5$.

Min Z=75, for $x_1=5$, $s_2= s_1- x_1=10$, $x_2= \frac{s_2}{2} = 5$, $s_3= s_2- x_2=5$, $x_3=s_3=5$.

Hence Min Z=75, for $x_1=5$, $x_2= 5$, $x_3=5$.

Example.4. Solve the following problem using dynamic programming:

Min Z= $x_1^2+x_2^2+x_3^2+x_4^2$

s. t. $x_1 x_2 x_3 x_4 =16$

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

Solution. We define the state variables as:

$s_1= x_1$

$s_2= x_1 x_2$

$s_3= x_1 x_2 x_3$

$s_4= x_1 x_2 x_3 x_4$

This gives $s_1=x_1=\frac{s_2}{x_2}$, $s_2 = \frac{s_3}{x_3}$, $s_3 = \frac{s_4}{x_4}$.

As there are four decision variables in the problem, we shall have four stages in the form of following four recursive relations:

$F_1(s_1)= x_1^2= s_1^2$

$F_2(s_2)$ = min $[x_2^2+ F_1(s_1)]$ = $\min\left[x_2^2 + \frac{s_2^2}{x_2^2}\right]$ with respect to x_2

$F_3(s_3)$= min $[x_3^2 + F_2(s_2)]$ with respect to x_3

$F_4(s_4)$= min $[x_4^2 + F_3(s_3)]$ with respect to x_4

For minimum of F_2, we have

$$\frac{dF_2}{dx_2} = 0$$

gives $2x_2 - 2\frac{s_2^2}{x_2^3} = 0$

$\Rightarrow \quad x_2 = \sqrt{s_2}.$

Also $\frac{d^2F_2}{dx_2^2} = 2+6\frac{s_2^2}{x_2^4}$ which is positive.

Therefore F_2 in minimum at $x_2 = \sqrt{s_2}.$

$$F_2(s_2) = \left[\left(\sqrt{s_2}\right)^2 + \frac{s_2^2}{\left(\sqrt{s_2}\right)^2}\right] = 2s_2.$$

Now $\quad F_3(s_3)$= min $[x_3^2 + F_2(s_2)]$

$F_3(s_3) = \min\left[x_3^2 + 2s_2\right]$

For minimum of F_3, we have

$$\frac{dF_3}{dx_3} = 0$$

gives $2x_3 - 2\frac{s_3}{x_3^2} = 0$

$\Rightarrow \quad x_3 = s_3^{1/3}. \qquad \left(s_2 = \frac{s_3}{x_3}\right)$

Also $\frac{d^2F_3}{dx_3^2} = 2+4\frac{s_3}{x_3^3}$ which is positive.

Therefore F_3 in minimum at $x_3 = s_3^{1/3}.$

Thus $F_3(s_3) = 3\, s_3^{\,2/3}$.

Now $F_4(s_4) = \min\, [x_4^2 + F_3(s_3)]$

$$= \min\left[x_4^2 + 3s_3^{\,2/3}\right] = \min\left[x_4^2 + 3\frac{s_4^{\,2/3}}{x_4^{2/3}}\right]$$

For minimum of F_4, we have

$$\frac{dF_4}{dx_4} = 0$$

gives $2x_4 - 2\dfrac{s_4^{2/3}}{x_4^{5/3}} = 0$

$\Rightarrow \quad x_4 = s_4^{\,1/4}$.

Also $\dfrac{d^2F_4}{dx_4^2} = 2 + \dfrac{10}{3}\dfrac{s_4^{2/3}}{x_4^{8/3}}$ which is positive.

Therefore F_4 in minimum at $x_4 = s_4^{\,1/4}$.

Thus $F_4(s_4) = 4\, s_4^{\,1/2}$.

$$x_4 = s_4^{\,1/4} = (16)^{1/4} = 2.$$

Now we have

$$s_3 = \frac{s_4}{x_4} = \frac{16}{2} = 8,$$

$$x_3 = s_3^{\,1/3} = (8)^{1/3} = 2,$$

$$s_2 = \frac{s_3}{x_3} = \frac{8}{2} = 4,$$

$$x_2 = \sqrt{s_2} = \sqrt{4} = 2,$$

$s_1 = \frac{s_2}{x_2} = \frac{4}{2} = 2 = x_1.$

Hence Min Z=16, for $x_1 = x_2 = x_3 = x_4 = 2$.

Example.5. Solve the following problem using dynamic programming:

Max $Z = x_1 x_2 x_3$

s.t. **$x_1 + x_2 + x_3 = 5$**

$x_1, x_2, x_3 \geq 0$.

Solution. We first define the state variables as:

$s_1 = x_1$

$s_2 = x_1 + x_2$

$s_3 = x_1 + x_2 + x_3 = 5$

This gives

$s_1 = x_1 = s_2 - x_2$

$s_2 = x_1 + x_2 = s_3 - x_3$

$s_3 = x_1 + x_2 + x_3 = 5$

Recursive relations are

$F_1(s_1) = x_1 = s_1$

$F_2(s_2) = \max [x_2 F_1(s_1)] = \max (x_2 . s_1) = \max [x_2 . (s_2 - x_2)]$ with respect to x_2

$F_3(s_3) = \max [x_3 . F_2(s_2)]$ with respect to x_3

For maximum of F_2, $\frac{dF_2}{dx_2} = 0$

gives $s_2 - 2x_2 = 0$

$\Rightarrow \quad x_2 = \frac{s_2}{2}.$

Also $\frac{d^2F_2}{dx_2^2} = -2x_2$ which is negative.

Therefore F_2 in maximum at $x_2 = \frac{s_2}{2}$.

Thus $F_2(s_2) = \frac{s_2^2}{4}$.

Now $F_3(s_3) = \max\left[x_3.\frac{s_2^2}{4}\right] = \max\left[x_3.\frac{(s_3 - x_3)^2}{4}\right]$

For minimum of F_3, we have

$$\frac{dF_3}{dx_3} = 0$$

gives $(s_3-x_3)^2-2x_3(s_3-x_3) = 0$

$\Rightarrow$ $(s_3-x_3)(s_3-3x_3) = 0$

or $(s_3-3x_3) = 0$ [$s_3-x_3 \neq 0$ because $x_2 = s_3-x_3>0$, $x_2 \neq 0$ because if $x_2=0$, then max $Z=0$.]

i.e., $x_3 = \frac{s_3}{3}$.

Also $\frac{d^2F_3}{dx_3^2} = 6x_3-4s_3 = -2s_3 < 0$ which is negative.

Therefore F_3 in maximum at $x_3 = \frac{s_3}{3}$.

Thus $F_3(s_3) = \frac{s_3^3}{27} = \frac{(5)^3}{27} = \frac{125}{27}$ *for* $s_3 = 5$.

Now we have

$x_3 = \frac{s_3}{3} = \frac{5}{3}$, $s_2 = s_3 - x_3 = 5 - \frac{5}{3} = \frac{10}{3}$, $x_2 = \frac{s_2}{2} = \frac{10}{3}.\frac{1}{2} = \frac{5}{3}$, $s_1 = s_2 - x_2 = \frac{10}{3} - \frac{5}{3} = \frac{5}{3}$, $x_1 = s_1 = \frac{5}{3}$.

Hence Max $Z = \frac{125}{27}$, for $x_1 = x_2 = x_3 = \frac{5}{3}$.

8.7 Summary

Dynamic programming is like a problem-solving superhero that tackles complex issues in a smart and efficient way. Imagine you have a big problem, and dynamic programming helps by breaking it down into smaller, more manageable parts. The cool trick is that it solves these smaller problems only once and stores the solutions so it doesn't have to repeat itself. The technique is particularly effective for optimization problems, where the goal is to find the best solution among a set of feasible solutions.

Dynamic programming is often applied to problems in which the same subproblems are solved multiple times, and the solutions to these subproblems can be reused to solve the overall problem more efficiently.

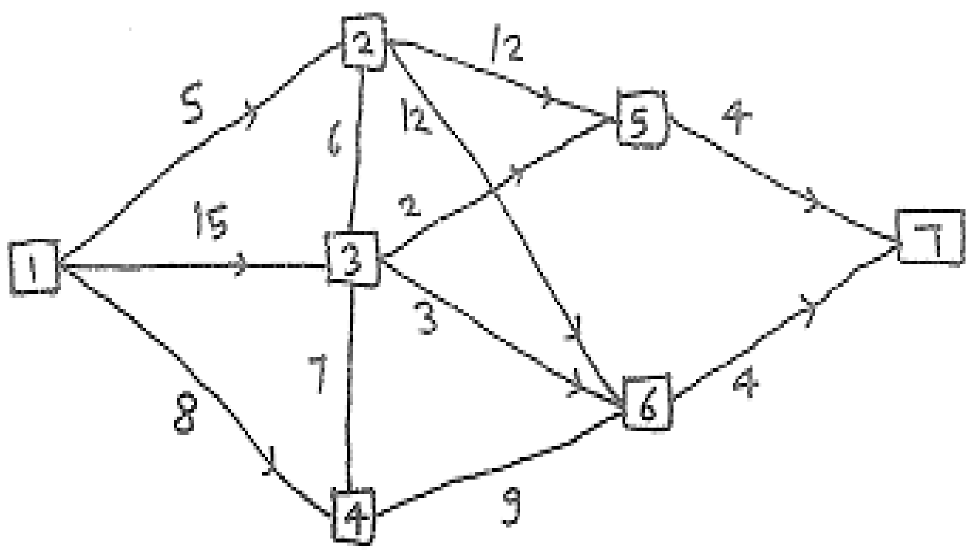
8.8 Terminal Questions

Q.1. Write a short note on Dynamic programming problem:

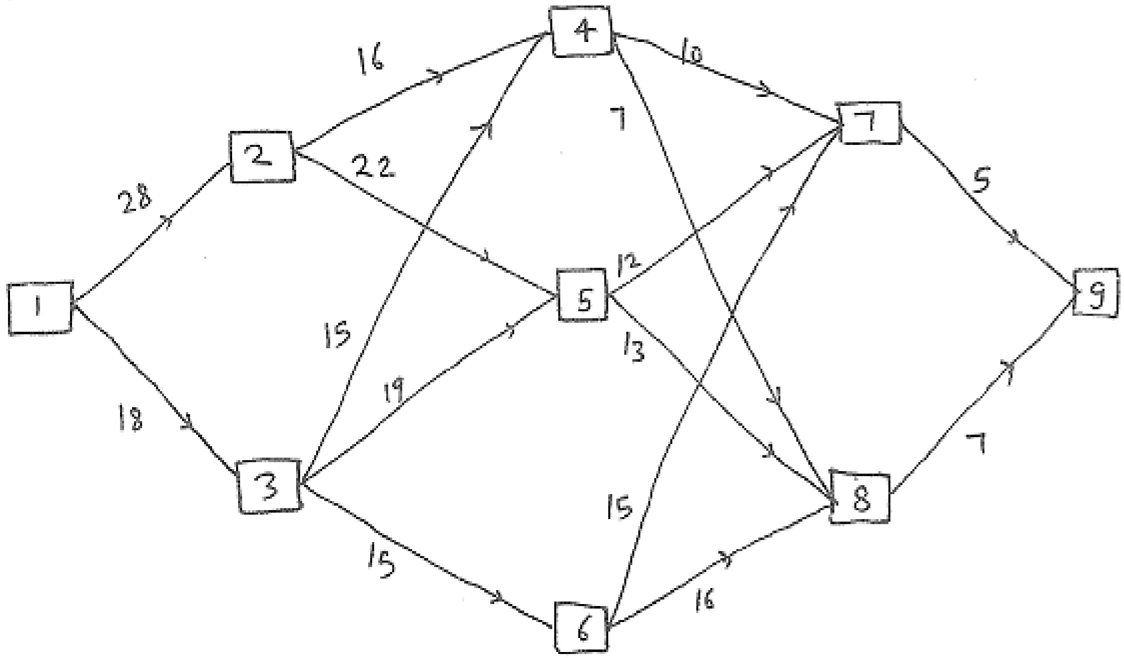
Q.2. State the Bellman's Principle of Optimality.

Q.3. What do you mean by stage and state?

Q.4. Find the shortest route of the following problem:



Q.5. Find the shortest path for the following travelling salesman problem:



Q.6. Solve the following problems using dynamic programming:

Min $Z = x_1^2 + x_2^2 + 2x_3^2$

s.t. $x_1 + x_2 + 2x_3 = 12$

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

Q.7. Solve the following problems using dynamic programming:

Min $Z = x_1^2 + x_2^2 + x_3^2$

s.t. $x_1 x_2 x_3 = 27$

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

Q.8. Solve the following problems using dynamic programming:

Max $Z = x_1 x_2 x_3$

s.t. $x_1 + x_2 + x_3 = 12$

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

Answers

4. 1→2→3→5→7; 17 units.

5. $1 \to 3 \to 4 \to 8 \to 9$; 47 units.

6. $x_1 = x_2 = x_3 = 3$, Min. Z=36.

7. $x_1 = x_2 = x_3 = 3$, Min. Z=27.

8. $x_1 = x_2 = x_3 = 4$, Max. Z=64.

UNIT-9: Applications of Dynamic Programming

Structure

9.1 Introduction

In this unit we shall discuss about the solution of linear programming problem using dynamic programming problem and application of dynamic programming problem. Dynamic programming is a powerful optimization technique that is used to solve problems by breaking them down into simpler, overlapping subproblems and solving each subproblem only once, storing the solutions to subproblems in a table to avoid redundant computations. Dynamic programming is widely used in various fields due to its ability to efficiently solve problems with optimal substructure and overlapping subproblems.

Dynamic programming can be applied to a wide range of problems in different domains where optimal solutions can be found by solving subproblems efficiently and avoiding redundant computations. Dynamic programming is used in algorithms like Dijkstra's and Floyd-Warshall for finding the shortest paths in a graph. In diverse algorithmic applications, dynamic programming plays a pivotal role, exemplified by its utilization in algorithms like Dijkstra's and Floyd-Warshall. The versatility of dynamic programming extends beyond graph-related challenges to encompass problem-solving in finance, bioinformatics, natural language processing, and more. The hallmark of dynamic programming lies in its capacity to disassemble intricate problems into simpler components, fostering optimal solutions through the systematic resolution of subproblems. This systematic approach and the storage of solutions for subsequent reuse contribute to the widespread adoption of dynamic programming across a spectrum of problem domains.

9.2 Objectives

After reading this unit the learner should be able to understand about the :

- solving a Linear Programming Problem Using Dynamic Programming
- applications of Dynamic Programming

9.3 Solving a Linear Programming Problem Using Dynamic Programming

Suppose a linear programming problem with *n* decision variables and *m* constraints. This problem can be considered as n-stage dynamic programming problem with m state variables because there are n-decision variables in the problem. However it becomes complicated if we take n ≥ 3. So we consider a linear programming problem with two decision variables and m-constraints as:

$$\text{Max } Z = c_1 x_1 + c_2 x_2$$

s. t.

$a_{11} x_1$	+	$a_{12} x_2$	≤	b_1
$a_{21} x_1$	+	$a_{22} x_2$	≤	b_2
…	…	…		…
…	…	…		…
…	…	…		…
$a_{m1} x_1$	+	$a_{m2} x_2$	≤	b_n
Stage-1		Stage-2		Resources

We develop an algorithm for solving this problem by forward dynamic programming as:

Stage-1.

$$F_1(x_1) = \underset{\left\{0 \le x_1 \le \left(\frac{b_1 - a_{12}x_2}{a_{11}}, \frac{b_2 - a_{22}x_2}{a_{21}}, \ldots\ldots\ldots, \frac{b_m - a_{m2}x_2}{a_{m1}}\right)\right\}}{Max\, c_1 x_1}$$

$$= \underset{0 \le x_1 = Min\left(\frac{b_1 - a_{12}x_2}{a_{11}}, \frac{b_2 - a_{22}x_2}{a_{21}}, \ldots\ldots\ldots, \frac{b_m - a_{m2}x_2}{a_{m1}}\right)}{Max\, c_1 x_1}$$

Here $x_1^* = \text{Min}\left(\frac{b_1 - a_{12}x_2}{a_{11}}, \frac{b_2 - a_{22}x_2}{a_{21}}, \ldots\ldots\ldots, \frac{b_m - a_{m2}x_2}{a_{m1}}\right)$ ..… (9.1)

$$F_1(x_2) = \text{Max } c_1 \left(\frac{b_1 - a_{12}x_2}{a_{11}}, \frac{b_2 - a_{22}x_2}{a_{21}}, \ldots\ldots\ldots, \frac{b_m - a_{m2}x_2}{a_{m1}}\right)$$

Stage-2. We have

$$F_2(x_2) = \text{Max}\left[c_1 \, Min\left(\frac{b_1 - a_{12}x_2}{a_{11}}, \frac{b_2 - a_{22}x_2}{a_{21}}, ..., \frac{b_m - a_{m2}x_2}{a_{m1}} \right) + c_2 x_2 \right] \quad ..... (9.2)$$

s.t. $$x_2 \le Min\left\{ \frac{b_1}{a_{12}}, \frac{b_2}{a_{22}}, ..., \frac{b_m}{a_{m2}} \right\} \quad ..... (9.3)$$

Explain (9.3) gives upper bound on x_2.

Now to calculate $Min\left(\frac{b_1 - a_{12}x_2}{a_{11}}, \frac{b_2 - a_{22}x_2}{a_{21}}, ..., \frac{b_m - a_{m2}x_2}{a_{m1}} \right)$.

We proceed as follows:

(i) We consider $\frac{b_1 - a_{12}x_2}{a_{11}}$ to be less than or equal to remaining $\frac{b_2 - a_{22}x_2}{a_{21}}, \frac{b_3 - a_{32}x_2}{a_{31}}$, ….., $\frac{b_m - a_{m2}x_2}{a_{m1}}$ one by one and find a condition on x_2. Thus getting (m-1) conditions on x_2 and then take intersection of all these to get the maximum x_2 provided $a_{ij} \ge 0$ for *i* and *j*. Then we find x_1 from (9.1) and Max Z from (9.2). This will give one set of solution.

(ii) Now we consider $\frac{b_2 - a_{22}x_2}{a_{21}}$ to be less than or equal to (i.e., minimum of remaining) $\frac{b_1 - a_{12}x_2}{a_{11}}$, $\frac{b_3 - a_{32}x_2}{a_{31}}$, ….., $\frac{b_m - a_{m2}x_2}{a_{m1}}$ and proceeding as in (i),we will get another set of solution.

Repeating this process we will get at most m such solutions. Among these solutions consider that one which gives the optimum solution of the problem.

Examples

Example.1. Solve the following problem using dynamic programming:

Max Z=500x_1 +800x_2

s.t. **5x_1 +6$x_2 \le$ 60**

x_1 +2$x_2 \le$ 16

$x_1 \le$ 8

$x_2 \le$ 6

and $\quad x_1, x_2 \geq 0.$

Solution. The given problem can be written as:

$$\text{Max } Z = 500x_1 + 800x_2$$

s.t. $\quad 5x_1 + 6x_2 \leq 60$

$x_1 + 2x_2 \leq 16$

$x_1 + 0.x_2 \leq 8$

$0.x_1 + x_2 \leq 6$

and $\quad x_1, x_2 \geq 0.$ ….. (9.4)

As there are two variables in the problem, we shall have two stages to solve the problem using dynamic programming. We can solve it either using forward dynamic programming or backward dynamic programming. Let us first solve it by using backward dynamic programming.

Stage-1. $F_1(x_1) = \underset{\left\{0 \leq x_1 \leq \left(\frac{60-6x_2}{5},\, 16-2x_2,\, 8\right)\right\}}{Max}\ 500\ x_1$

$$= \underset{0 \leq x_1 = Min\left(\frac{60-6x_2}{5},\, 16-2x_2,\, 8\right)}{Max}\ 500\ x_1$$

Here $x_1^* = \text{Min}\left(\frac{60-6x_2}{5}, 16-2x_2, 8\right)$

$$F_1(x_2) = \text{Max } 500\left(\frac{60-6x_2}{5}, 16-2x_2, 8\right)$$

Stage-2. We have

$$F_2(x_2) = \text{Max}\left\{500\ Min\left(\frac{60-6x_2}{5}, 16-2x_2, 8\right) + 800\,x_2\right\} \quad ..… (9.5)$$

s.t. $\quad x_2 \leq Min\left\{\frac{60}{6}, \frac{16}{2}, 6\right\} = 6 \quad ..… (9.6)$

Now to calculate $Min\left(\frac{60-6x_2}{5}, 16-2x_2, 8\right)$. We proceed as follows:

Case-I. Assuming 8 to be minimum, we have

$$8 \le \frac{60-6x_2}{5} \quad \Rightarrow \quad x_2 \le \frac{10}{3}$$

Also $\quad 8 \le 16-2x_2 \quad \Rightarrow \quad x_2 \le 4,\ x_2 \ge 0.$

Thus we have $0 \le x_2 \le \frac{10}{3}$.

Using equation (9.5), we get

$$\text{Max } Z = \text{Max } [500\times 8+800\, x_2], \qquad \left(0 \le x_2 \le \frac{10}{3}\right)$$

$$= \text{Max } [4000+800\times\frac{10}{3}]$$

$$= \frac{20000}{3} \quad \text{when } x_2=\frac{10}{3} \text{ and } x_1^*=8. \qquad \text{…. (9.7)}$$

Case-II. Let $Min\left(\frac{60-6x_2}{5}, 16-2x_2, 8\right)=\frac{60-6x_2}{5}$. Then we have

$$\frac{60-6x_2}{5} \le 8 \quad \Rightarrow \quad x_2 \ge \frac{10}{3}$$

Also $\quad \frac{60-6x_2}{5} \le 16-2x_2 \quad \Rightarrow \quad x_2 \le 5,\ x_2 \ge 0.$

Thus we have $\frac{10}{3} \le x_2 \le 5$.

Using equation (9.5), we get

$$\text{Max } Z= \text{Max } \left[500\left(\frac{60-6x_2}{5}\right)+800x_2\right], \qquad \left(\frac{10}{3} \le x_2 \le 5\right)$$

$$= \text{Max } [6000+200\, x_2]$$

$$= 7000 \qquad \text{when } x_2=5 \text{ and } x_1^*=6. \qquad \text{….. (9.8)}$$

Case-III. Let $Min\left(\frac{60-6x_2}{5}, 16-2x_2, 8\right)=16-2x_2$. Then we have

$$16-2x_2 \le 8 \quad \Rightarrow \quad x_2 \ge 4.$$

Also $16-2x_2 \le \dfrac{60-6x_2}{5}$ $\Rightarrow$ $x_2 \ge 5$, $x_2 \ge 0$.

Thus we have

$x_2 \ge 5$.

From equation (9.6), we have

$x_2 \le 6$.

Therefore we have

$5 \le x_2 \le 6$.

Now using equation (9.5), we have

$$\text{Max } Z = \text{Max }[500(16-2x_2)+8x_2], \quad (5 \le x_2 \le 6)$$

$$= \text{Max }[8000-200x_2]$$

$$= 7000 \qquad \text{when } x_2=5 \text{ and } x_1^*=6 \qquad \ldots.. (9.9)$$

From equations (9.7), (9.8) and (9.9), the optimal solution is

$$\text{Max } Z = \text{Max}\left[\frac{20000}{3}, 7000, 7000\right]$$

$$= 7000 \text{ which is for } x_1^*=5 \text{ and } x_2=6.$$

Example.2. Solve the following problem using dynamic programming:

$$\textbf{Max } Z=10x_1+8x_2$$

s.t. $2x_1+x_2 \le 25$

$3x_1+2x_2 \le 45$

$x_2 \le 10$

and $x_1, x_2 \ge 0$.

Solution. The given problem can be written as:

$$\text{Max } Z=10x_1+8x_2$$

s.t. $2x_1+x_2 \le 25$

$$3x_1 + 2x_2 \le 45$$

$$0.x_1 + x_2 \le 10$$

$$\text{and} \quad x_1, x_2 \ge 0. \qquad \text{... (9.10)}$$

As there are two variables in the problem, we shall have two stages to solve the problem using dynamic programming. We can solve it either using forward dynamic programming or backward dynamic programming. Let us first solve it by using backward dynamic programming.

Stage-2. $F_2(x_2) = \underset{\left\{0 \le x_2 \le \left(25-2x_1, \frac{45-3x_1}{2}, 10\right)\right\}}{Max}\ 8\,x_2$

$$= \underset{0 \le x_2 = Min\left(25-2x_1, \frac{45-3x_1}{2}, 10\right)}{Max}\ 8\,x_2$$

Here $x_2^* = \text{Min}\left(25-2x_1, \frac{45-3x_1}{2}, 10\right)$

$$F_2(x_1) = \text{Max } 8\left(25-2x_1, \frac{45-3x_1}{2}, 10\right)$$

Stage-1. We have

$$F_1(x_1) = \text{Max}\left[10x_1 + 8\left\{Min\left(25-2x_1, \frac{45-3x_1}{2}, 10\right)\right\}\right] \qquad \text{..... (9.11)}$$

s.t. $\qquad x_1 \le Min\left\{\frac{25}{2}, \frac{45}{3}\right\} = \frac{25}{2}$ $\qquad$ ..... (9.12)

Now to calculate $Min\left(25-2x_1, \frac{45-3x_1}{2}, 10\right)$.

We proceed as follows:

Case-I. Assuming 10 to be minimum, we have

$$10 \le 25-2x_1$$

$$\Rightarrow \qquad x_1 \le \frac{15}{2}$$

Also $\qquad 10 \le \frac{45-3x_1}{2}$

$\Rightarrow \qquad x_1 \le \frac{25}{3}, x_1 \ge 0.$

Thus we have

$$0 \le x_1 \le \frac{15}{2}.$$

Using equation (9.11), we get

$$\text{Max Z} = \text{Max } [10\, x_1 + 8\times 10], \quad \left(0 \le x_1 \le \frac{15}{2}\right)$$

$$= \text{Max } [10\times\frac{15}{2}+80]$$

$$= 155 \text{ when } x_1 = \frac{15}{2} \text{ and } x_2^* = 10. \qquad \text{…. (9.13)}$$

Case-II. Let $Min\left(25-2x_1, \frac{45-3x_1}{2}, 10\right) = 25\text{-}2x_1$.

Then we have

$$25 - 2x_1 \le \frac{45-3x_1}{2}$$

$\Rightarrow \qquad x_1 \ge 5$

Also $\qquad 25 - 2x_1 \le 10$

$\Rightarrow \qquad x_1 \ge \frac{15}{2}, x_1 \ge 0.$

Thus we have

$$x_1 \ge \frac{15}{2}.$$

From equation (9.12) we have

$$x_1 \le \frac{25}{2}$$

Therefore we have

$$\frac{15}{2} \le x_1 \le \frac{25}{2}.$$

Using equation (9.11), we get

$$\text{Max } Z = \text{Max } [10\, x_1 + 8(25-2x_1)], \quad \left(\frac{15}{2} \le x_1 \le \frac{25}{2}\right)$$

$$= \text{Max } [200-6x_1]$$

$$= 155 \quad \text{when } x_1 = \frac{15}{2} \text{ and } x_2^* = 10. \qquad \text{.... (9.14)}$$

Case-III. Let $Min\left(25-2x_1, \frac{45-3x_1}{2}, 10\right) = \frac{45-3x_1}{2}$.

Then we have

$$\frac{45-3x_1}{2} \le 25 - 2x_1$$

$\Rightarrow$ $x_1 \le 5$ (not applicable)

Also $\frac{45-3x_1}{2} \le 10$

$\Rightarrow$ $x_1 \ge \frac{25}{3}$, $x_1 \ge 0$.

Thus we have

$$x_1 \ge \frac{25}{3}.$$

From equation (9.12) we have

$$x_1 \le \frac{25}{2}$$

Therefore we have

$$\frac{25}{3} \le x_1 \le \frac{25}{2}.$$

Using equation (9.11), we get

$$\text{Max } Z = \text{Max } \left[10\, x_1 + 8\left(\frac{45-3x_1}{2}\right)\right], \quad \left(\frac{25}{3} \le x_1 \le \frac{25}{2}\right)$$

$$= \text{Max } [10\, x_1 + 180 - 12x_1]$$

$= \text{Max } [180 - 2x_1]$

$= \dfrac{490}{3}$ when $x_1 = \dfrac{25}{3}$ and $x_2^* = 10.$ …. (9.15)

From equations (9.13), (9.14) and (9.15), the optimal solution is

$$\text{Max } Z = \text{Max}\left[155, 155, \frac{490}{3}\right]$$

$= \dfrac{490}{3}$ which is for $x_1 = \dfrac{25}{3}$ and $x_2 = 10.$

9.4 Applications of Dynamic Programming

Example.3. A vessel which can carry a load of at most 7 units is loaded with three items. The weight per unit of different items and their values are given below. How many units of each item be loaded in the vessel so as to maximize the value of the loaded items.

Item	Weight/Unit	Value/Unit
1	1	20
2	3	90
3	2	70

Solution. Suppose x_1, x_2 and x_3 be the number of units of items 1, 2 and 3 loaded respectively. The given problem can be formulated as:

$$\text{Max } Z = 20x_1 + 90x_2 + 70x_3$$

s.t. $x_1 + 3x_2 + 2x_3 \le 7$

$x_1, x_2, x_3 \ge 0$ and integers.

The decision variables are discrete here, so we cannot use calculus method and solve the problem in tabular form. Now we define the state variables as:

$s_1 = x_1$ …. (9.16)

$s_2 = x_1 + 3x_2$ …. (9.17)

$s_3 = x_1 + 3x_2 + 2x_3 \leq 7$ …. (9.18)

This gives

$s_1 = x_1 = s_2 - 3x_2$

$s_2 = s_3 - 2x_3$

$s_3 = x_1 + x_2 + x_3 \leq 7$

Recursive relations are

$F_1(s_1) = f_1(x_1) = 20\ x_1 = 28\ s_1$

$F_2(s_2) = \text{Max}\ [f_2(x_2) + F_1(s_1)]$

$F_3(s_3) = \text{Max}\ [f_3(x_3) + F_2(s_2)]$

From equation (9.18), we have

x_1=0,1, 2, 3, 4, 5, 6, 7;

x_2=0, 1, 2;

x_3= 0, 1, 2, 3. The state transformations are shown in following tables 1 and 2:

x_3 \ s_3	0	1	2	3
0	0	---	---	---
1	1	---	---	---
2	2	0	---	---
3	3	1	---	---
4	4	2	0	---
5	5	3	1	---
6	6	4	2	0
7	7	5	3	1

Table-2

x_2 \ s_2	0	1	2
0	0	---	---
1	1	---	---
2	2	---	---
3	3	0	---
4	4	1	---
5	5	2	---
6	6	3	0
7	7	4	1

Table-1

In table 1 the entries are $s_2= s_3-2x_3$. (Table 1 gives possible values of s_2) Then we prepare table 2, in which entries are $s_1=s_2-3x_2$ values (that is s_1 can assume for various combinations of s_2 and x_2 values). Now we prepare Table 3, 4 and 5 for using recursive relations:

$x_1=s_1$	0	1	2	3	4	5	6	7
$F_1(s_1)=f_1(x_1)=20\ x_1$	0^{**}	20*	40	60	80	100	120	140

Table-3

	$f_2(x_2)=90\ x_2$	$F_1(s_1)$	$f_2(x_2)=90\ x_2$	$F_2(s_2)$

x_2 \ s_2	0	1	2	0	1	2	0	1	2	Max of rows
0	0	---	---	0	---	---	0	---	---	0
1	0	---	---	20^*	---	---	20^*	---	---	20^*
2	0	---	---	40	---	---	40	---	---	40
3	0	90	---	60	0^{**}	---	60	90^*	---	90^{**}
4	0	90	---	80	20	---	80	110	---	110
5	0	90	---	100	40	---	100	130	---	130
6	0	90	180	120	0	0	120	150	180	180
7	0	90	180	140	80	20	140	170	200	200

Table-4

In table-4:

(a) We enter values in possible position. The possible positions are dictated by Table 1.

(b) To get the $F_1(s_1)$ matrix, we read x_1 from Table 2 and then read $F_1(s_1)$ from Table 3.

(c) The last column gives the maximum with respect to x_2 for a fixed s_2.

The above is the solution of second sub-problem. Now we proceed to calculate $F_3(s_3)$, which will give the optimal value for this problem. To get $F_3(s_3)$, we construct Table 5.

	$f_3(x_3)=70\ x_3$	$F_2(x_2)$	$f_3(x_3)+F_2(x_2)$	$F_3(x_3)$ Max of

													Rows
x_3 / s_3	0	1	2	3	0	1	2	3	0	1	2	3	
0	0	---	---	---	0	---	---	---	0	---	---	---	0
1	0	---	---	---	20	---	---	---	20	---	---	---	20
2	0	70	---	---	40	0	---	---	40	70	---	---	70
3	0	70	---	---	90	20	---	---	90	90	---	---	90
4	0	70	140	---	110	40	---	---	110	110	140	---	140
5	0	70	140	---	130	90	20	---	130	180	180	---	160
6	0	70	140	210	180	110	40	0	180	180	180	210	210
7	0	70	140	210	200	130	90^{**}	20^{*}	200	200	230^{**}	230^{*}	230^{***}

Table 5

In Table 5, to get matrix $F_2(s_2)$,we read s_2 from Table 1 and then read corresponding $F_2(s_2)$ from Table 4. From the last column of Table 5, we see that largest $F_3(s_3)$=230 is for s_3 and x_3.now we trace back the entries which gave this largest value. These entries give the optimal solution. This is shown in table by marking the optimal entries by (*). From Table 4 and Table 3, we have s_2=1, x_2=0 and s_1=x_1=1 respectively. The optimal values are x_1=1, x_2=0, x_3=3 and maximum value is 230.

In this problem, Table 5 shows that there is an alternative optimal solution. This solution, we have shown by marking entries with (**). The alternate optimal solution is x_3=2, x_2=1, x_1=0 and maximum value is 230.

9.4 Summary

Dynamic programming serves as a potent optimization technique, employed to address problems by decomposing them into manageable, overlapping subproblems. The approach entails solving each subproblem only once and storing solutions in a table to prevent redundant computations. This methodology is extensively utilized across various domains due to its efficacy in efficiently tackling problems characterized by optimal substructure and overlapping subproblems.

9.5 Terminal Questions

Q.1. Write a short note on applications of dynamic programming.

Q.2. Solve the following problems using dynamic programming:

$$\text{Min } Z = x_1^2 + x_2^2 + x_3^2$$

$$\text{s.t. } x_1 + x_2 + x_3 \geq 10$$

$$\text{and } x_1, x_2, x_3 \geq 0.$$

Q.3. A vessel which can carry a load of at most 4 tons is loaded with three items. The following table gives the unit weight, w_i, in tons and the unit revenue, r_i, in thousand Rs. for items *i*. How many units of each item be loaded in the vessel so as to maximize the value of the loaded items.

Item *i*	w_i	r_i
1	2	31
2	3	47
3	1	14

Answers

2. $x_1 = x_2 = x_3 = 10/3$. Min. $Z = 100/3$.

3. $x_1 = 2$, $x_2 = 0$, $x_3 = 0$. Maximum value is 62,000.

Master of Science
PGMM -102N
Classical Optimization Techniques

U. P. Rajarshi Tandon Open University

Block

4 Advanced Optimization Techniques

Unit- 10

Networking

Unit- 11

Game Theory

Unit- 12

Goal Programming

Unit- 13

Integer Programming-I

Unit- 14

Integer Programming-II

Block-4

Advanced Optimization Techniques

In the context of operations research, networking typically refers to the study and optimization of networks, where a network is a collection of interconnected nodes and links. Operations research utilizes various mathematical and analytical methods to model, analyze, and optimize the flow of resources, information, or activities within these networks. Networking in operations research provides valuable tools for decision-makers to optimize resource allocation, improve efficiency, and make informed decisions in various domains where interconnected systems play a crucial role.

Game theory is an important branch of mathematics and economics that studies the strategic interactions between rational decision-makers, known as players, in situations where the outcome of each player's choice depends on the choices of others. It is a framework for analyzing and understanding the behavior of individuals, organizations, or countries in competitive or cooperative situations.

Goal programming is a special type of mathematical optimization technique used to solve decision-making problems where multiple, often conflicting, objectives need to be considered simultaneously. It is employed in situations where there is a need to achieve several goals, and these goals may have different priorities or importance levels. The primary aim of goal programming is to find a solution that minimizes the deviations from the specified goals.

Integer programming problems are generally more challenging to solve than linear programming problems without integer constraints. Traditional optimization techniques may involve exploring a large solution space. Branch and bound, cutting-plane methods, and specialized algorithms like branch and cut are common approaches used to solve integer programming problems. Applications of integer programming can be found in various fields, including manufacturing, logistics, finance, and project management, where decisions involve discrete choices or whole quantities.

UNIT-10: Networking

Structure

10.1 Introduction

Network analysis holds significant importance in electrical theory and communication systems, encompassing applications in circuits, pipelines, mobile networks, roads, transportation, railways, airlines, blood vessels, production, resource management, distribution, planning, scheduling, and control of research and development projects, among others. Transportation and assignment problems, both categorized as Linear Programming Problems (LPP), are examples of network optimization models.

In this chapter, we will delve into the discussion of specific network optimization problems, including the shortest route problem, minimum spanning tree problem, and maximum flow problems.

10.2 Objectives

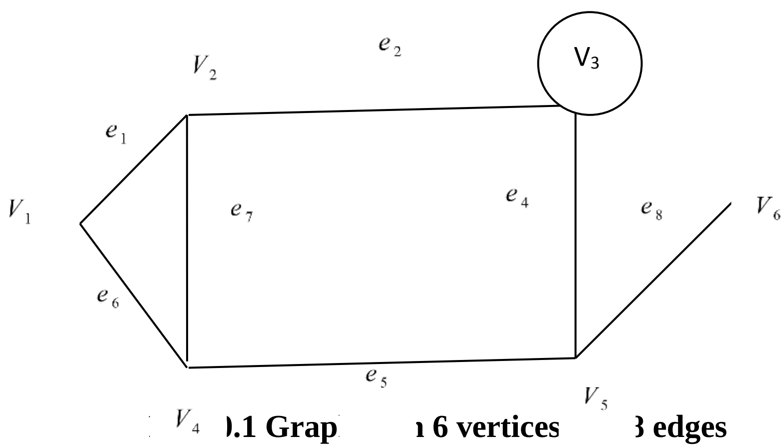
After reading this unit the learner should be able to understand about:

- the terminologies used in networking
- the networking

- the shortest route problem
- the minimum spanning tree problem
- the maximum flow problem
-

10.3 Terminologies Used in Networking

A **graph** G is a collection of vertices (V) that may or may not be connected to each other by edges (E). In other words, a **graph** G is a set of points (known as vertices denoted by V) connected by lines (called Edges denoted by E). For example, the graph with 6 vertices, $V = (V_1, V_2, V_3, V_4, V_5, V_6)$ and 8 edges, $E = (e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8)$ are given below (Fig. 10.1):



.1 Grap 6 vertices 8 edges

A graph G is said to be a ***weighted graph*** in which weight are assigned with each edge. For example, Fig 10.2 is a weighted graph.

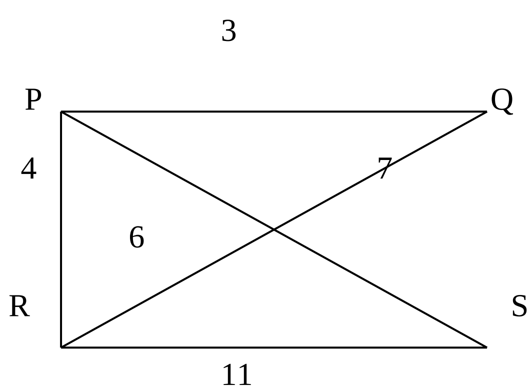


Fig 10.2 Weighted Graph

A graph G is said to be ***labeled graph*** in which each vertex is assigned a unique name or label. Fig 10.3 shown a labeled graph. A graph G is said to be ***complete*** in which every vertex is connected to every other vertex i.e. at least one edge exists between every pair of vertices. It is denoted by K*n*. Fig. 10.4 shown a K_4 graph (complete graph).

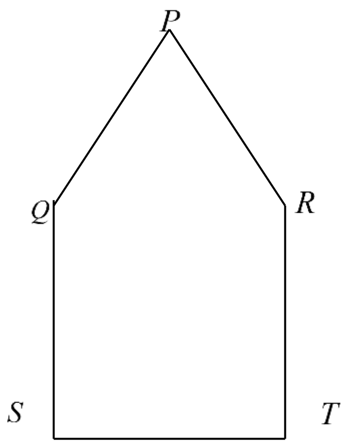


Fig. 10.3 Labeled Graph

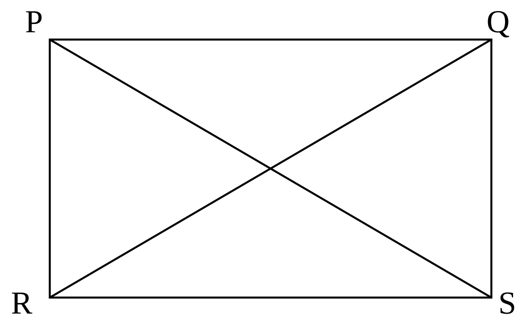
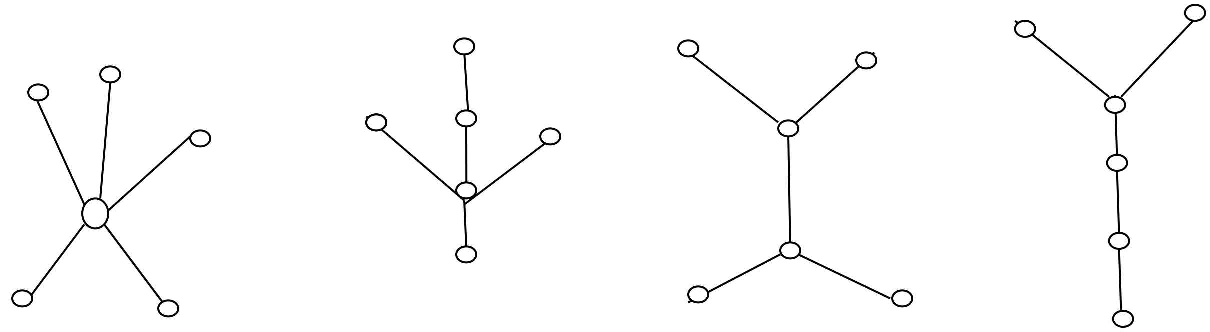


Fig. 10.4 Complete Graph K_4

A graph is said to be ***finite*** which have finite no of vertices as well as finite no of edges otherwise, it is called ***infinite graph***. An edge is ***incident*** with a vertex if the edge is joined to the vertex. If there is an edge joining a pair of vertices, those vertices are said to be ***adjacent*** otherwise they are ***non-adjacent***. The number of edges which are connected to a given vertex is called the ***degree*** of that vertex. It is denoted by *d*(V). The degree of self -loops counted twice. In Fig. 10.2, $d(v_1)=2,\ d(v_2)=3,\ d(v_3)=4,\ d(v_5)=3,\ d(v_4)=3,\ d(v_6)=1.$ A vertex is said to be ***Isolated*** if a vertex having no incident edge. A vertex have a zero degree is called ***isolated vertex.*** A vertex is said to be pendant vertex if a vertex having one degree.is ***pendant vertex.***

A ***walk*** is defined of a graph G is an alternating sequence of finite vertices of edges. Which is beginning and ending with vertices. No edge traversed more than once in a walk whenever a vertex may appear more than once. A walk to begin and end at the distinct vertex is called open walk. A walk to begin and end at the same vertex is called closed walk. An open walk in which on vertex appears more than once is called a path and also called a simple or elementary path. In a path, total no of edge is called the length of a path. A self-loop can be included in a walk but not in a path. In fig. 10.1, $v_6e_8v_5e_4v_3e_2v_2e_1v_1e_6v_4$ is a path. A closed walk in which no vertex (except the beginning and ending vertex) appears more than once is called a ***circuit*** and also called a cycle, elementary cycle, circular path and polygon. Every circuit is not a self-loop but every self-loop is circuit. In Fig. 10.1, $v_1e_1v_2e_2v_3e_4v_5e_5v_4e_6v_1$ is a circuit. A graph G is said to be connected if there is at least one edge between every pair of vertices in G. otherwise, we can say the graph is disconnected but a disconnected graph consists of two or more connected graphs each of these connected graph is known as a component.

A ***tree*** is a connected graph without any circuit. By the definition of the tree it is clear that tree is connected simple and a cyclic graph. The trees are represented by symbol T. The graph show in Fig 10.5 are the examples of tree.



T_1 T_2 T_3 T_4

Fig 10.5 The trees of order 6

A path in any graph is tree because there is no circuit in path. A tree with only one vertex is called a trivial tree and all other trees are called non trivial trees. A tree with *n* vertices has exactly $(n - 1)$ edges. A connected graph G is said to be ***minimally connected*** if deletion of any edge from G, then the graph G is disconnected. The graph shown in Fig. 10.6 is minimally connected graph.

Fig. 10.6 Minimally connected graph

The distance between two vertices of a tree is easy because there is no circuit and therefore there is one and only one path between every pair of vertices. The distance between the vertices of tree T shown in Fig. 10.7 is:

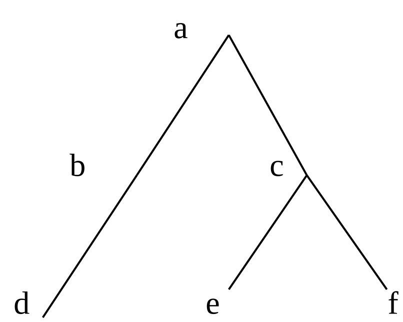


Fig 10.7 Tree

Here $d(a, b) = 1$, $d(a, d) = 2$, $d(a, c) = 1$, $(a, e) = 2$, $d(a, f) = 2$, $d(c, d) = 2$, $d(c, d) = 3$, $d(e, d) = 4$, $d(f, d) = 4$ and so on. Two special types of trees called the rooted and binary tree. A tree in which one vertex is distinct from all other vertices is called a ***rooted tree.***

A tree in which there is exactly one vertex of degree two and each of the remaining vertices of degree one or three is called a ***binary tree***. Binary tree is also a rooted tree because the vertex of degree 2 is distinct from all other vertices, hence the vertices with degree 2 is root of the binary tree.

Binary trees are widely used in computer applications such as searching methods known as binary search, sorting methods know as heap sort. The vertex with degree one in a binary tree are called external vertices or terminal vertices and all the other vertices are called internal vertices. The number of vertices in binary tree is always odd. The number of external vertices in binary tree is $\frac{(n+1)}{2}$. The number of internal vertices in binary tree is $\frac{(n-1)}{2}$. Source is the starting event of a project. It is an event with only succeeding but no proceeding activity. Sink is the last event showing the end of a project. It is an event with only preceding but no succeeding activity.

10.4 Networking

A network can be characterized as a collection of points or nodes linked by connections or arrows. Essentially, it is a graph where a flow can traverse the branches. In a network, the branches intersect exclusively at nodes. The challenge within network analysis involves identifying a course of action that minimizes certain performance metrics, such as time, distance, cost, and more.

10.5 Shortest Route Problem

The objective of determining the most concise path from an origin to a destination within a network is termed the shortest route problem, alternatively recognized as the minimum path problem. To address the challenge of identifying the minimum path, we will explore the method known as Dijkstra's method, pioneered by F.W. Dijkstra in 1959. Let u_i be the shortest path from node 1 to node i and d_{ij} is an arc length (i, j). The label for node j is defined as $[u_j, i] = [u_i+d_{ij}, i]$, $d_{ij} \geq 0$.

In Dijkstra algorithm, there are two types of node labels: temporary and permanent. A temporary label can be replaced with another label, if a shortest route to the same node can be found. If no shortest route can be found then it becomes node with permanent label. The procedure to determine the shortest route using Dijkstra's algorithm consists of the following steps:

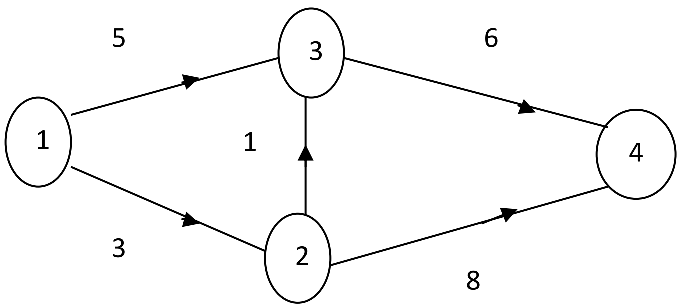
Step-1: First label the node 1, with permanent label [0, —], and set $j = 1$.

Step-2: Find the temporary labels $[u_i+d_{ij}, i]$ for each node j that can be reached from node i. If the node j has already labeled $[u_j, r]$ through another node r, such that $u_i+d_{ij} < u_j$, then replace $[u_j, r]$ with $[u_i+d_{ij}, i]$. Otherwise $[u_i+d_{ij}, i]$ is the permanent label of node j.

Step-3: For $j<n$, set j= next j, to reach permanent label. If all the nodes have permanent labels, then stop.

Examples

Example.1. Determine the shortest route from the following network:



Sol. It is given that the network has 4 nodes. The calculations for these nodes are as follows:

Node.1: [0,—].

Node.2: [0+3, 1] = [3, 1].

Node.3: [0+5, 1] = [5, 1] and [3+1, 2] = [4, 2].

Node.4: Reached from nodes 2 and 3: [3+8, 2] = [11, 2] and [4+6, 3] = [10, 3].

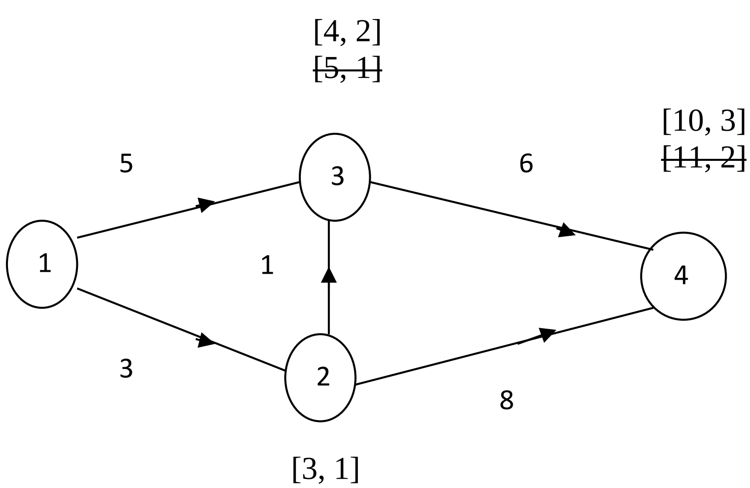
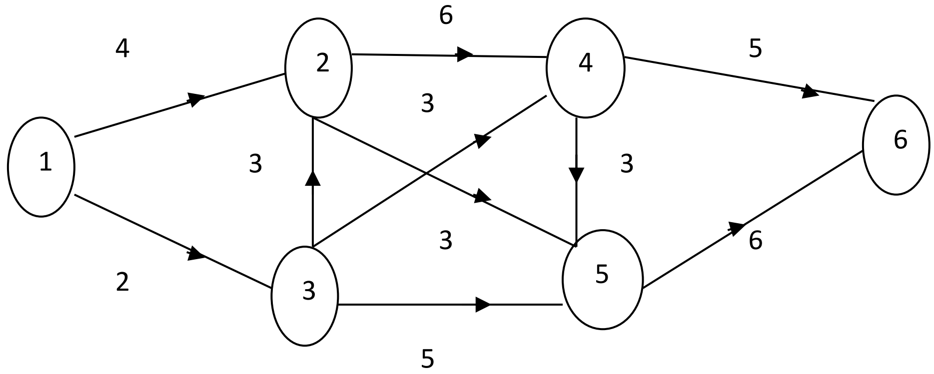


Fig. 10.8 Showing Shortest Route

Thus the shortest route covers a distance of 10 units. Now back tracking from node 4 (Fig. 10.8), by checking the second elements of the labels, we find the possible route as 1→2→3→4.

Example.2: Find the shortest route from the following network:



Sol. It is given that the network has 6 nodes. The calculations for the nodes are as follows:

Node.1: [0, —].

Node.3: [0+2, 1]=[2, 1].

Node.2: [0+4, 1]=[4, 1] and [2+3, 3]=[5, 3].

Node.4: [4+6, 2]=[10, 2] and [2+3, 3]=[5, 3].

Node.5: [4+3, 2]=[7, 2], [2+5, 3]=[7, 3] and [5+3, 4]=[8, 4].

Node.6: [5+5, 4]=[10, 4] and [7+6, 5]=[13, 5].

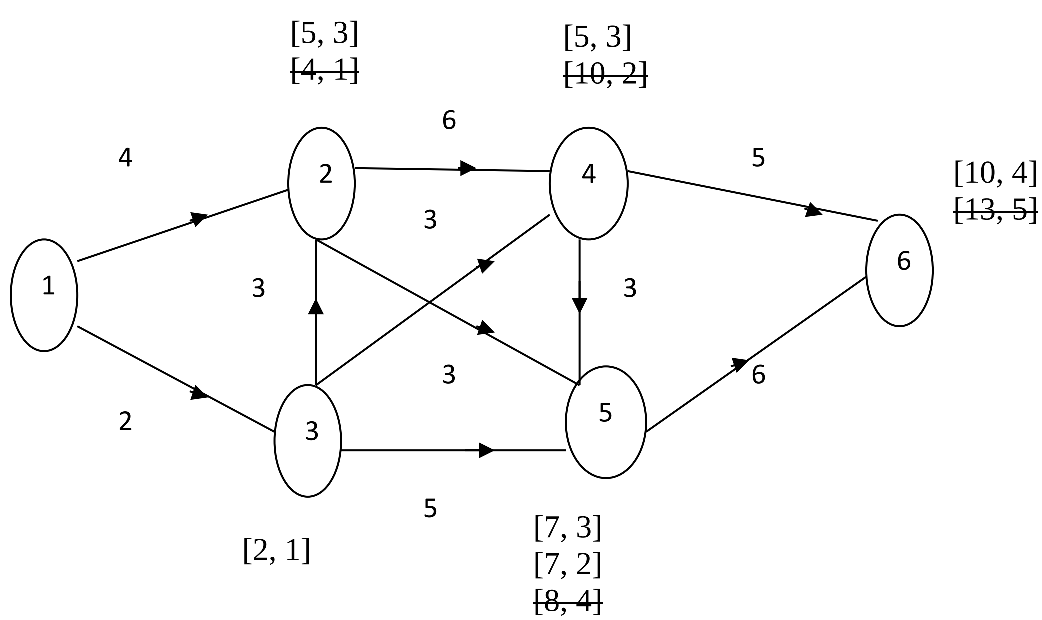
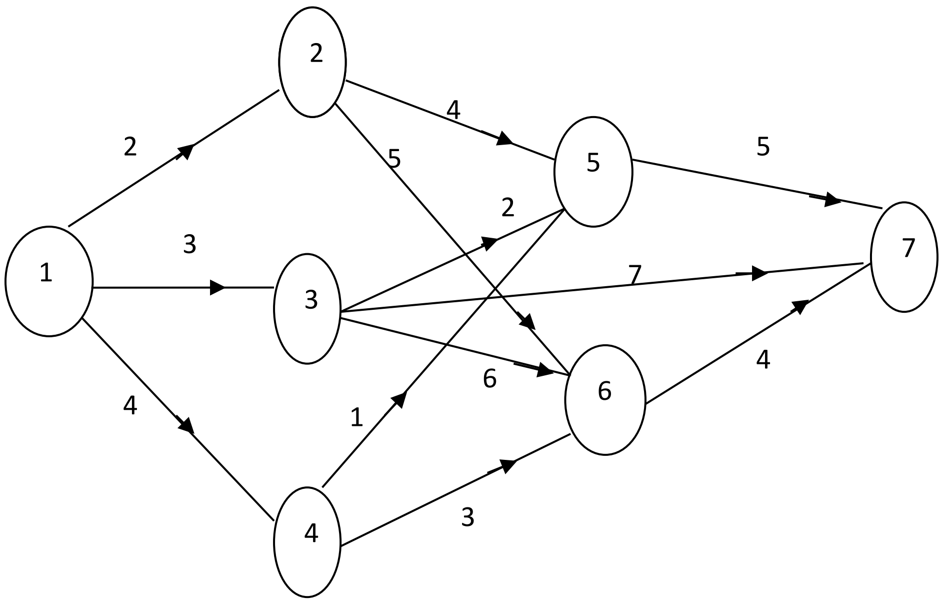


Fig. 10.9 Showing Shortest Route

Thus the shortest route covers a distance of 10 units. Now back tracking from node 6 (Fig. 10.9), by checking the second elements of the labels, we find the possible route as 1→3→4→6.

Example.10.3: Find the shortest route from the following network:



Sol. It is given that the network has 7 nodes. The calculations for the nodes are as follows:

Node.1: [0, —].

Node.2: [0+2, 1]=[2, 1].

Node.3: [0+3, 1]=[3, 1].

Node.4: [0+4, 1]=[4, 1].

Node.5: [2+4, 2]=[6, 2], [3+2, 3]=[5, 3] and [4+1, 4]=[5, 4].

Node.6: [2+5, 2]=[7, 2], [3+6, 3]=[9, 3] and [4+3, 4]=[7, 4].

Node.7: [3+7, 3]=[10, 3], [5+5, 5]=[10, 5] and [7+4, 6]=[11, 6].

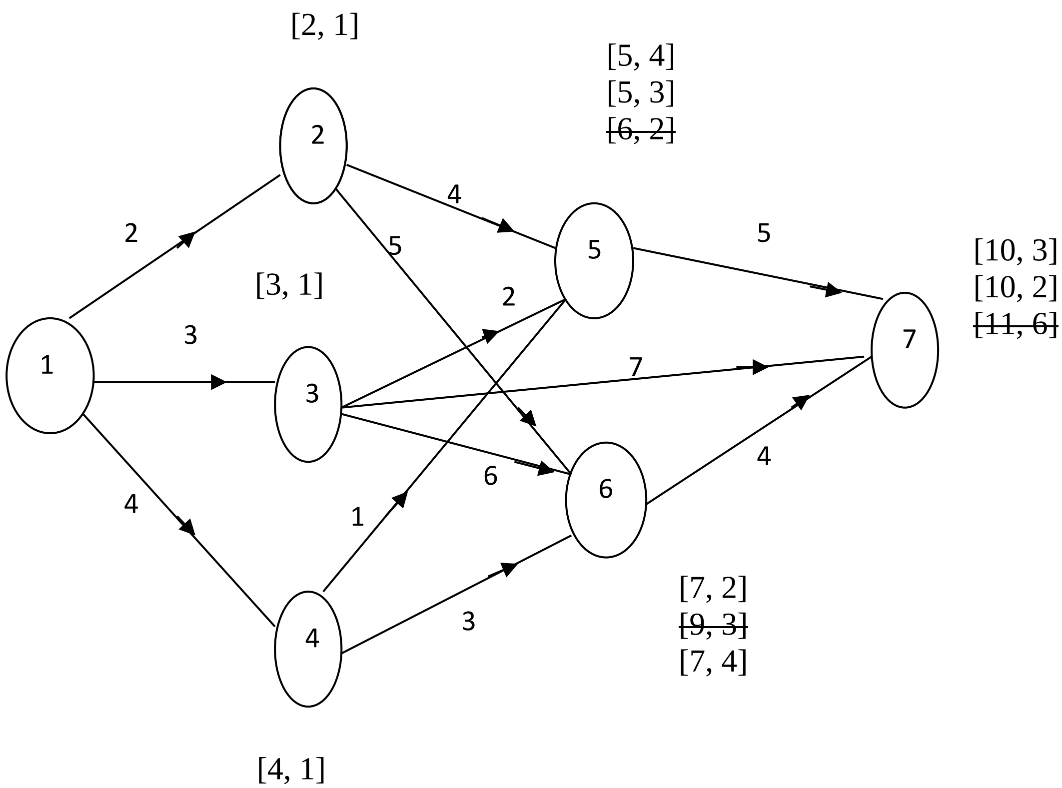


Fig. 10.10 Showing Shortest Route

Thus the shortest route covers a distance of 10 units.

Now back tracking from node 6 (Fig. 10.10), by checking the second elements of the labels, we find the possible routes are

$1\to3\to5\to7$ or $1\to4\to5\to7$ or $1\to3\to7$.

10.6 Minimum Spanning Tree Problem

To determine the minimum spanning tree using a straightforward method initially introduced by J.W. Kruskal in 1956, the approach involves selecting the smallest edge length to build the growing minimum spanning tree, ensuring it remains loop-free.

Consider nodes labeled 1, 2, 3, ..., *n*. The procedure for finding the minimum spanning tree comprises the following steps:

Step-I: First of all write the arc length in the increasing order of magnitude. Take S, the set of nodes of minimum spanning tree.

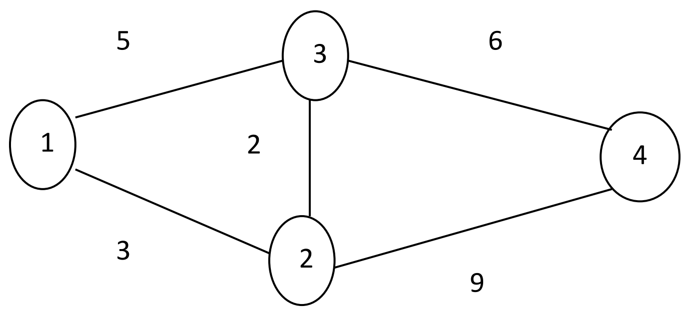
Step-II: Take the minimum arc length (i, j), $i \neq j$, both i, j not belonging to S (chose arbitrarily in case of tie).

Step-III: Now set S=S∪{i, j}. If S connecting all nodes then stop and check if S consists of unconnected branches then go to step IV otherwise go to step II.

Step-IV: Connect the unconnected branches by minimum arcs and stop.

Examples

Example.10.4. Find the minimum spanning tree from the following network:



Sol. First write the arc length in the increasing order of magnitude: (2, 3), (1, 2), (1, 3), (3, 4) and (2, 4).

Iteration-1: Take the minimum length arc (2, 3), therefore we have

S = (2, 3)

Iteration-2: Now minimum arc length (1, 2), therefore we have

S = {1, 2, 3}

Iteration-3: Next minimum arc length (1, 3), but 1, 3 are in S so neglect (1, 3).

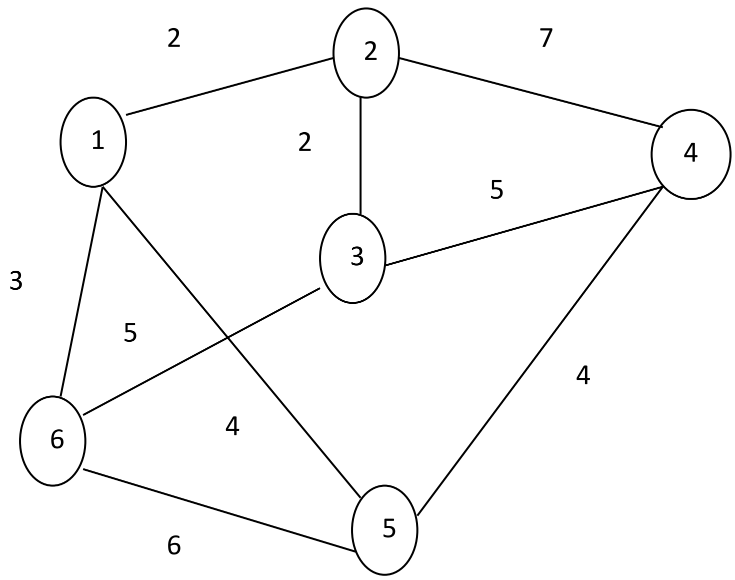
Iteration-4: Now minimum length arc (3, 4), therefore we have

S = {1, 2, 3, 4}.

All nodes are now connected.

Hence the length of the given tree is 11.

Example.18.5. Determine the minimum spanning tree from the following network:



Sol. Write the arc length in the increasing order of magnitude:

(1, 2), (2, 3), (1, 6), (1, 5), (4, 5), (3, 4), (3, 6), (5, 6) and (2, 4).

Iteration-1: Take the minimum length arc (1, 2), S = {1, 2}

Iteration-2: Now minimum arc length (2, 3), therefore we have

S = (1, 2, 3}

Iteration-3: Next minimum arc length (1, 6), therefore we have

S = {6, 1, 2, 3}.

Iteration-4: Next minimum arc length (4, 5) or also be chose (1, 5), therefore we have

S = {6, 1, 2, 3, 4, 5}.

Iteration-5: Next minimum arc length (1, 5), therefore we have

S = {6, 1, 2, 3, 4, 5}.

All nodes are now connected.

Hence the length of the given tree is 15.

10.7 Maximum Flow Problem

Numerous situations in our daily lives involve the concept of flow rates, such as the flow of oil in pipelines and traffic flow. In these scenarios, the objective is to maximize the flow, a problem that can be formulated as a Linear Programming (LP) problem. Maximum flow problems typically revolve around directing flow through a connected network, starting from a designated node called the source and terminating at another node known as the sink. All other nodes in the network are considered transshipment nodes.

Flow through an arc is permitted only in the direction indicated by the arrowhead, with the maximum flow determined by the capacity of that specific arc. At the source, all arcs point away from the node, while at the sink, all arcs point into the node. The primary goal is to maximize the total flow from the source to the sink, measured in two equivalent ways—either as the amount leaving the source or the amount entering the sink.

The procedure for determining the maximum flow in these problems consists of the following steps:

Step-I: First find a path from source to sink that can be accommodating a positive flow of the material.

Step-II: Obtain the maximum flow that can be shipped along the finding path and denoted by it Z (say).

Step-III: Now decrease the capacity of each branch in the direction of flow of the Z units in this path and increase the reverse capacity by Z, and add Z units to the amount delivered to the sink. Again find a path from source to sink. If none of exists then go to step IV otherwise go to step II.

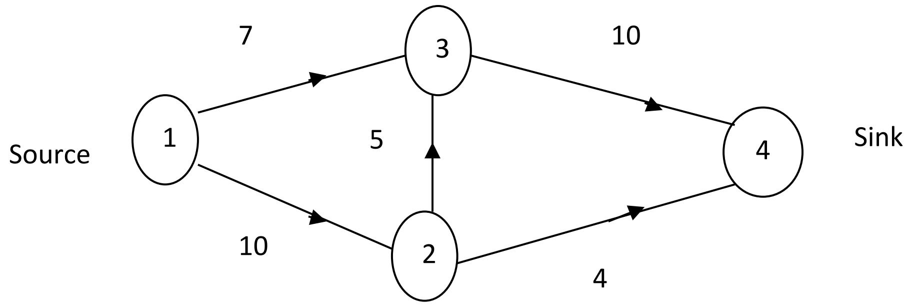
Step-IV: Calculate the maximum flow of the problem is the amount of material delivered to the sink.

Examples

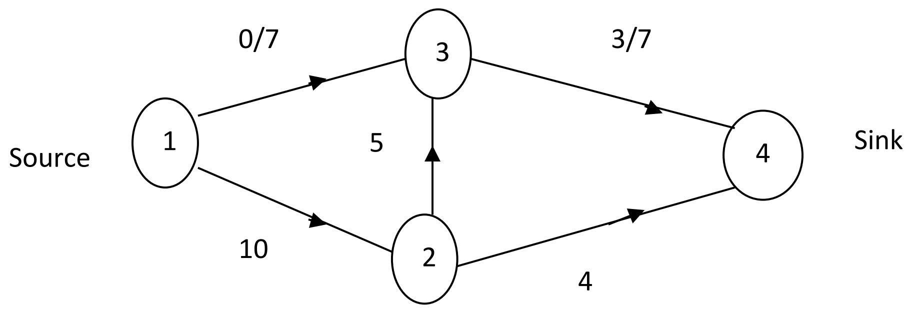
Example.6. Find the maximum flow for which the arcs and capacities are given in the following:

Arc	Capacity
1-2	10
2-3	5
1-3	7
2-4	4
3-4	10

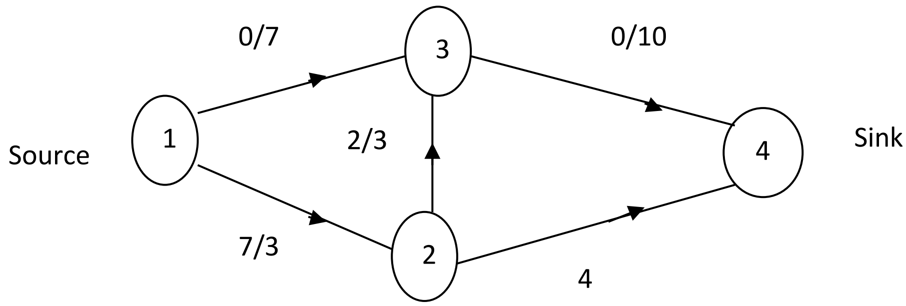
Sol. The network for the given maximum flow problem is



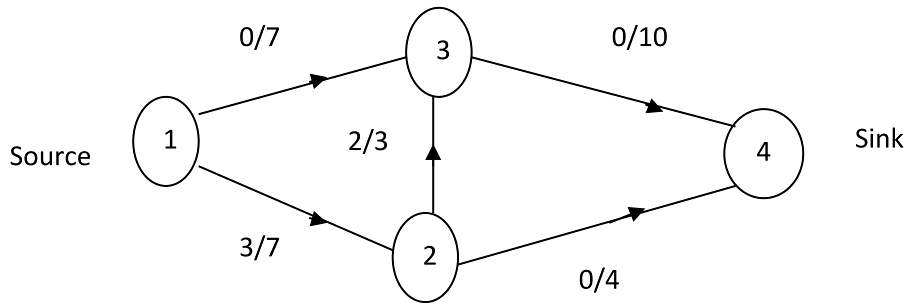
First find a path from source to sink (1-3-4) that can be accommodating a positive flow of the material. Obtain the maximum flow (*i.e.*, 7) that can be shipped along the finding path. Now decrease the capacity of each branch in the direction of flow of the 7 units in this path (1-3-4) and increase the reverse capacity by 7, and add 7 units to the amount delivered to the sink.



Now find a path from source to sink (1-2-3-4). Obtain the maximum flow (*i.e.*, 3) that can be shipped along the finding path. Now decrease the capacity of each branch in the direction of flow of the 3 units in this path (1-2-3-4) and increase the reverse capacity by 3, and add 3 units to the amount delivered to the sink.



Again find a path from source to sink (1-2-4). Obtain the maximum flow (*i.e.*, 4) that can be shipped along the finding path. Now decrease the capacity of each branch in the direction of flow of the 4 units in this path (1-2-4) and increase the reverse capacity by 4, and add 4 units to the amount delivered to the sink.



Here none of path is exists from source to sink that can be accommodating a positive flow of the material.

Calculate the maximum flow (*i.e.,* 14) of the problem is the amount of material delivered to the sink.

10.8 Summary

To find the shortest route from the origin to the destination in a network, we encounter what is known as the shortest route problem. In the Dijkstra algorithm, nodes are classified into two types: temporary and permanent. A temporary label can be replaced with another label if a shorter route to the same node is discovered. If no shorter route is found, the label becomes permanent. When aiming to obtain the minimum spanning tree, the first step involves arranging the arc lengths in ascending order of magnitude.

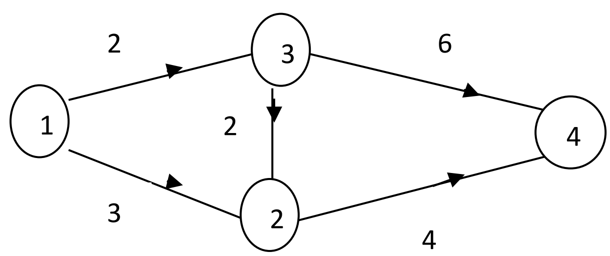
Let S represent the set of nodes forming the minimum spanning tree. Now choose the minimum arc length (i, j), $i \neq j$, both i, j not belonging to S (chose arbitrarily in case of tie). In problems related to maximum flow, the primary goal is to maximize the overall flow volume from the source to the sink. This flow quantity can be assessed in two equivalent manners—either as the amount departing from the source or as the amount entering the sink.

10.9 Terminal Questions

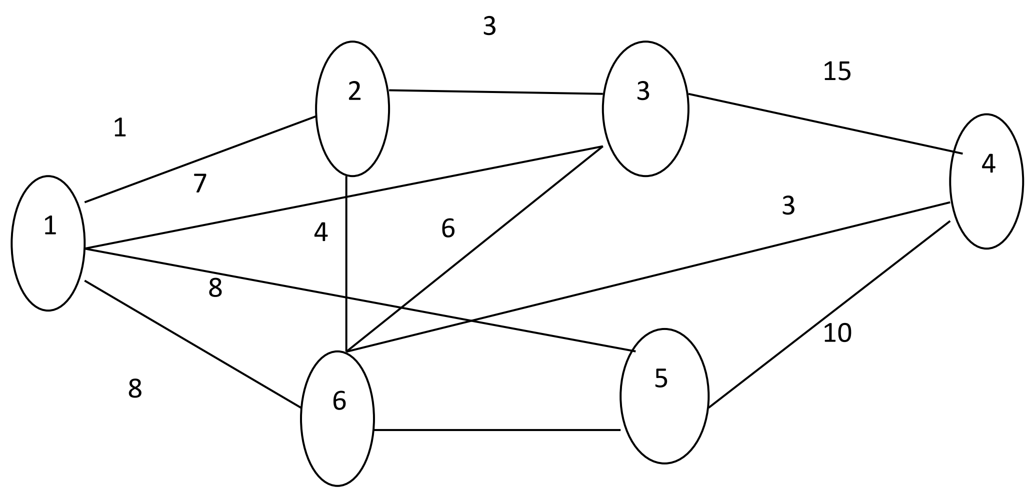
Q.1. Explain the Networking.

Q.2. Write the procedure of minimum spanning tree problem.

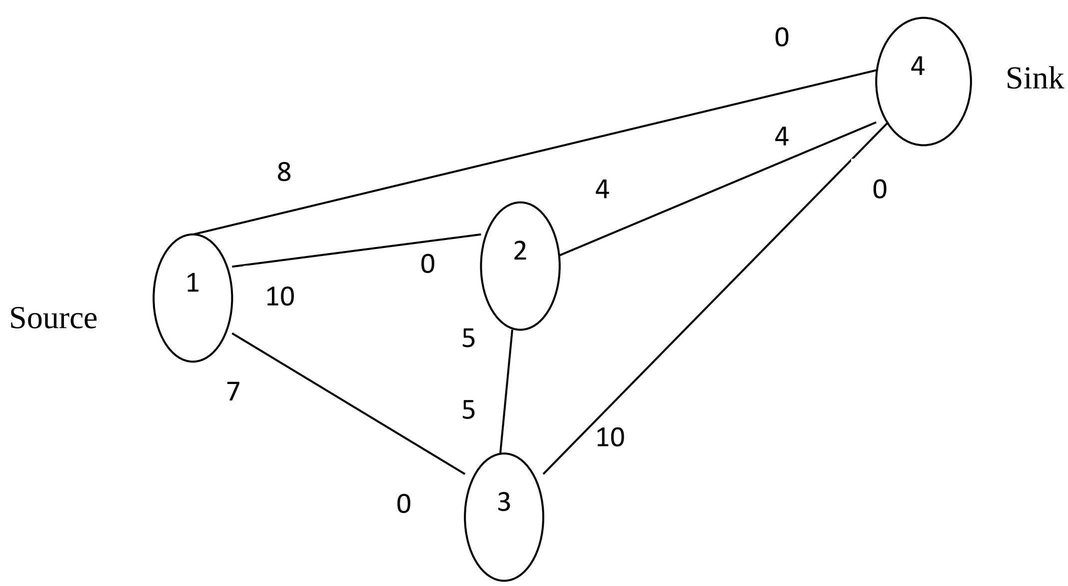
Q.3. Determine the shortest roué from the following network:



Q.4.Find the minimum spanning tree from the following network:



Q.5. Find the maximum flow from the following network:



Answer

3. 7 units, 1→2→4.

4. [(1, 2), (2, 3), (4, 6), (5, 6), (2, 6)]

5. 22 units.

UNIT-11: Game Theory

Structure

11.1 Introduction

In practical problem-solving scenarios, such as those in economics, military strategy, and other fields, analysts often encounter situations where multiple parties are pursuing conflicting objectives. The outcome of each party's actions depends on the choices made by others involved, leading to what is commonly referred to as "conflict situations." Examples of such situations include those arising during military operations or various economic scenarios, especially those involving free competition among entities like firms and industrial enterprises. Conflicting situations in real life are typically intricate.

However, our objective here is to focus on simpler situations and create a formalized model of the scenario. This model, termed a "game," is distinguished from real conflict situations by the presence of well-defined rules governing its play. Essentially, a game is defined as a set of rules that guide the interactions within it.

11.2 Objectives

After reading this unit the learner should be able to understand about:

- the game theory
- two-person zero-sum (rectangular) games
- the payoff matrix and value of the game
- saddle point
- strategies
- the lower and upper value of the game
- procedure to find saddle point
- the minimax principle with pure strategies
- the two person zero sum game with mixed strategies

11.3 Game

A game can be characterized as a conflict of interests between two or more opponents or parties. In the first scenario, it is termed a two-person game, and in the latter, a multi-person game. While both types exist, two-person games are often more prominent in practical applications, and our focus will be primarily on these. In the context of the game, participants are commonly referred to as "players," and the result of an encounter is labeled as the parties' "payoff" or gain.

A fundamental concept in game theory is that of a strategy. For a player, a strategy entails a set of clear rules that dictate the selection of each individual move based on the unfolding situation within the game. The term "finite game" is used when either player possesses only a finite number of strategies.

B / A	B_1	B_2	…	B_1	…	B_m
A_1	a_{11}	a_{12}	…	a_{1j}	…	a_{1m}

A_2	a_{21}	a_{22}	…	a_{2j}	..	a_{2m}
…	…	…	…	…		…
A_i	a_{i1}	a_{i2}	…	a_{ij}	..	a_{2m}
…	…	…	…	…		…
A_n	a_{n1}	a_{n2}	…	a_{nj}		a_{nm}

Table 11.1 Payoff Matrix

In a finite game, where player A has m strategies say $A_1, A_2, \ldots\ldots\ldots, A_m$ and player B has n strategies say $B_1, B_2, \ldots\ldots\ldots, B_n$ is called a $m \times n$ game.

The choice of strategies A_i by player A and B_j by player B determines the payoff (gain) to player A, denoted as a_{ij}. The values a_{ij}'s can be put in matrix form as given in above.

This matrix is known as payoff (gain) matrix of player A. Entry a_{ij} may be positive (gain to player A) or negative (loss to player A) or zero (no loss, no gain to A).

The matrix of a $m \times n$ game is of the form given above will be denoted as $(a_{ij})_{m \times n}$. The strategy (A_i, B_j) which when repeated a number of times assumes to a player the maximum possible gain (or minimum possible loss) is called an optimal strategy and the corresponding value a_{ij} is called the game value denoted by α.

11.4 Two-Person Zero-Sum (Rectangular) Games

A game featuring only two players, denoted as player A and player B, is termed a two-person zero-sum game when the gain of one player (say, A) is equivalent to the loss incurred by the other player (say, B). This equality ensures that the sum of their net gains is zero. Two-person zero-sum games are also referred to as rectangular games, as they can be effectively represented by a payoff matrix in a rectangular format.

11.5 Payoff Matrix

Suppose the player A has m activities (strategies) $A_1, A_2, \ldots\ldots\ldots, A_m$ and player B has n activities $B_1, B_2,$

………, B_n, then a payoff matrix can be formed by adopting the following rules:

(i) The row designations of each matrix are activities $A_1, A_2, \ldots\ldots, A_i, \ldots, A_m$ available to player A.

(ii) The column designations of each matrix are activities $B_1, B_2, \ldots\ldots, B_i, \ldots, B_n$ available to player B.

(iii) The cell entry "a_{ij}" in the payment to player A in A's payoff matrix when A chooses the activity A_i and B chooses the activity B_j.

(iv) In a zero-sum two person game the cell entry in B's payoff matrix will be negative of the corresponding cell entry "a_{ij}" in the player A's payoff matrix, so that the sum of the payoff matrices of player A and player B is ultimately zero.

		B_1	B_2	……	B_j	……	B_n
	A_1	a_{11}	a_{12}	……	a_{1j}	……	a_{1n}
	A_2	a_{21}	a_{22}	……	a_{2j}	……	a_{2n}
	⋮	.					
Player A	A_i	a_{i1}	a_{i2}	……	a_{ij}	……	a_{in}
	⋮	⋮	⋮		⋮		⋮
	A_m	a_{m1}	a_{m2}	……	a_{mj}	……	a_{mn}

Table 11.2 Representing A's payoff matrix

		B_1	B_2	……	B_j	……	B_n
	A_1	$-a_{11}$	$-a_{12}$	……	$-a_{1j}$	……	$-a_{1n}$
	A_2	$-a_{21}$	$-a_{22}$	……	$-a_{2j}$	……	$-a_{2n}$
	⋮	.					
Player B	A_i	$-a_{i1}$	$-a_{i2}$	……	$-a_{ij}$	……	$-a_{in}$
	⋮	⋮ ⋯	⋮	⋯	⋮	⋯	⋮
	A_m	$-a_{m1}$	$-a_{m2}$	……	$-a_{mj}$	……	$-a_{mn}$

Table 11.3 Representing B's payoff matrix

Note: In practical problems, there is no need of writing B's payoff matrix as it is just the negative of A's payoff matrix. Thus if 'a_{ij}' is the gain to player A, then $-a_{ij}$ is the gain to player B, so that net gain is zero.

11.6 Value of the Game

The payoff a_{rs} at the saddle point (A_r, B_s) is called the value of the game.

11.7 Saddle Point

A saddle point (A_r, B_s) of a payoff matrix is the position of such an element in the payoff matrix which is minimum in its row and maximum in its column. Mathematically if the payoff matrix (a_{ij}) is such that $\max_i\left(\min_j\left(a_{ij}\right)\right)=\min_j\left(\max_i\left(a_{ij}\right)\right)=a_{rs}$ then matrix is said to have a saddle point (A_r, B_s)

11.8 Strategies

When two players play a game then they have different alternatives at their disposal to go ahead with the game. These alternatives used in the game are called strategies.

Optimal Strategy

If the payoff matrix (a_{ij}) has saddle point (r, s) then A_r and B_s are called the optimal strategies of player A and B respectively.

Pure Strategy

A single alternative when played for certain, then it is called pure strategy.

Mixed Strategy

If several alternatives with different choices are used to play the game then we call it a game with mixed strategy. When saddle point does not exist, mixed strategies are used to find the value of the game.

11.9 The Lower and Upper Value of the Game

Consider the payoff matrix $(a_{ij})_{m\times n}$ of the player A (see table 11.1). If player A chooses the strategy A_i, then he is sure to get $\min_j a_{ij}$, j varies from 1 to n. Thus A would like to choose that strategy A_i, i=1, 2, …., m for which $\min_j a_{ij}$ is maximized to get $\max_i\left(\min_j a_{ij}\right)$. Denote it by $\underline{a}$, called the **lower value** of the game.

Thus $\max_i \left(\min_j a_{ij} \right) = \underline{a}$ ….(11.1)

On the other hand if player B chosen strategy B_j, then he is sure that player A will not get more than $\max_i a_{ij}$. Thus B would like to choose that strategy B_j which minimizes the maximum gain to player A i.e., to have $\min_j \left(\max_i a_{ij} \right)$. Denote it by $\bar{a}$, called the **upper value** of the game.

Thus $\min_j \left(\max_i a_{ij} \right) = \bar{a}$ ….(11.2)

Equations in (11.1) and (11.2) are called maximin and minimax criteria respectively. If $\underline{a} = \bar{a} = \alpha$ (say), then the game is said to have a saddle point (i.e., solution to the game exists) and α is the game value (payoff to player A). If $\alpha = a_{ij}$ (see table 11.1), then we say that the optimal strategy of player A is A_i and that of B is B_j.

Note: A game is said to be **fair** if both the lower and upper values of the game are equal to zero.

Theorem.11.1: If $\underline{a} = \max_i \left(\min_j a_{ij} \right)$ is the lower value and $\bar{a} = \min_j \left(\max_i a_{ij} \right)$ is the upper value of the game, then lower value is always less than or equal to the upper value of the game i.e., $\underline{a} = \max_i \left(\min_j a_{ij} \right) \le \min_j \left(\max_i a_{ij} \right) = \bar{a}$.

Proof.: It is evident that $\min_j a_{ij} \le a_{ij}$ for any j and i fixed ….(11.3)

Also $\max_i a_{ij} \ge a_{ij}$ for any i and j fixed ….(11.4)

Let $\min_j a_{ij} = a_{ir}$ for $j=r$ ….(11.5)

and $\max_i a_{ij} = a_{sj}$ for $i=s$ ….(11.6)

Thus from equations (11.5) and (11.6), we have $a_{ir} \le a_{ij} \le a_{sj}$ for all i and j.

and hence $\max_i a_{ir} \le a_{ij} \le \min_j a_{sj}$ ….(11.7)

Using equations (11.5), (11.6) and (11.7), we have

$$\max_i \min_j a_{ij} \leq a_{ij} \leq \min_j \max_i a_{ij}$$

or $$\max_i \min_j a_{ij} \leq \min_j \max_i a_{ij}$$

11.10 Procedure to find the Saddle point

The following steps are used to determine the game value of the given game problem:

Step-I: First choose the minimum element of each row *i* (α_i's) of the payoff matrix and write it on the extreme right of that row *i*.

Step-II: Choose the greatest element of each column *j* (β_j's) of the payoff matrix and write it against (below) that column *j*.

Step-III: If maximum of α_i's equal to minimum of β_j's then the common value is the game value otherwise game value does not exist in pure strategies.

Step-IV: When (saddle point) does not exist, use mixed strategies to find the value of the game.

Examples

Example.1. How to devise strategies by two warring country. The aim of both the country is to hit the other country and devise a way to remain unit. This type of problem is explained with the help of the following device. Consider a problem of military operation between two countries A and B. Country A has three kinds of weapons A_1, A_2 and A_3. Country B has three kinds of weapons B_1, B_2 and B_3. Country A's goal is to hit its enemy's aircraft while the other party's goal is to avoid being hit. When armament A_1 is used aircrafts B_1, B_2 and B_3 are hit with probabilities 0.9, 0.4 and 0.2 respectively, when armament A_2 is used, they are hit with probabilities 0.3, 0.6 and 0.8; and when armament A_3 is used they are hit with probabilities 0.5, 0.7 and 0.2. Formulate the problem in terms of game theory. Also find the lower value and upper value of the game?

Solution. The formulation of the game theory is

B / A	B_1	B_2	B_3	α_i
A_1	0.9	0.4	0.2	0.2

A_2	0.3	0.6	0.8	0.3
A_3	0.5	0.7	0.2	0.2
β_j	0.9	0.7	0.8	

Here $\alpha_1 = 0.2$, $\alpha_2 = 0.3$, $\alpha_3 = 0.2$ and $\beta_1 = 0.9$, $\beta_2 = 0.7$, $\beta_3 = 0.8$.

The lower value of the game

$$\underline{a} = \max(0.2, 0.3, 0.2) = 0.3.$$

The upper value of the game

$$\bar{a} = \min(0.9, 0.7, 0.8) = 0.7.$$

Example.2. Two players A and B, each write down simultaneously and independent by of each other, one of the four numbers 1, 2, 3 or 4. If the sum of the numbers they have written down is even, then B pays that sum to A; if the sum is odd, then A pays that sum to B.

(i) Construct the game matrix.

(ii) Find the lower and upper value of the game.

(iii) Find the maximin and minimax strategies for players A and B.

Solution. **(i)** There are four strategies for player A: writing down 1, A_1; writing down 2, A_2; writing down 3, A_3; and writing down 4, A_4. Opponent B also has the same four strategies: B_1, B_2, B_3 and B_4. This is a 4×4 game with game matrix (payoff matrix of player A) given below:

B / A	B_1	B_2	B_3	B_4	α_i (row maximum)
A_1	2	-3	4	-5	-5
A_2	-3	4	-5	6	-5

A_3	4	-5	6	-7	-7
A_4	-5	6	-7	8	-7
β_j (column maximum)	4	6	6	8	

where α_i = minimun of i^{th} row

and β_j = maximum of j^{th} column.

Evidently B can respond to any strategy chosen in the way which is worst for A. Indeed if A chooses strategy A_1, for instance, B will always counter it with strategy B_4. Strategy A_2 will always be countered by B_3. Strategy A_3 will be countered by B_4 and A_4 will be countered by B_3.

(ii) The lower value of the game

$\underline{a}$ = max (-5, -5, -7, -7) =-5.

The upper value of the game

$\bar{a}$ = min (4, 6, 6, 8) = 4.

(iii) Player A's, maximin strategy is either of the strategies A_1 or A_2; employing them systematically he can in any case guarantee that his gain is not less than -5 (a loss not greater than 5).

B's minimax strategy is B_1; employing it systematically he can guarantee that his loss is not greater than 4.

Note: 1. If any of the players A or B deviates from these maximin or minimax strategies respectively then their gain is adversely affected.

Note: 2. In the above examples 1 and 2 we have come across the situation in which both players employ their minimax strategies is unstable $(\underline{a} \neq \bar{a})$and may be distributed by information received about the opposite party's strategy.

However there are some games, for which minimax strategies are stable $(\underline{a} = \bar{a})$. These are the games for which the lower value of the game is equal to the upper one *i.e.*, $\underline{a} = \bar{a} = \alpha$ (*i.e.*, saddle point and hence game value exists).

Example.3. Find the lower and upper value of the game for the game matrix (payoff matrix of player A) given below:

$$\begin{array}{cc} & Player\,B \\ Player\,A & \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \end{array}$$ Does the saddle point exist?

Solution. The game matrix (payoff matrix of player A) is given below:

B \ A	B_1	B_2	α_i
A_1	4	0	0
A_2	0	2	0
β_j	4	2	

Where α_i = minimun of i^{th} row

and β_j = maximum of j^{th} column.

Max $\{\alpha_i\} = 0$, min $\{\beta_j\} = 2$.

Here strategy of player A is a maximin strategy and strategy of player B is a minimax strategy. Lower value of the game is 0 and upper game value is 2.

Hence the lower game value and upper game value are not equal, so we can say that the value of the saddle point does not exist.

Example.4. Find the lower and upper value of the game for the game matrix (payoff matrix of player A) given below:

$$\begin{array}{cc} & Player\,B \\ Player\,A & \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} \end{array}$$

Does the saddle point exist?

Solution. The game matrix (payoff matrix of player A) is given below:

B \ A	B_1	B_2	α_i
A_1	1	-1	-1
A_2	2	3	2
β_j	2	3	

where α_i = minimun of i^{th} row

and β_j = maximum of j^{th} column.

Max $\{\alpha_i\} = 2$, min $\{\beta_j\} = 2$.

Here strategy of player A is a maximin strategy and strategy of player B is a minimax strategy. Lower value of the game is 2 and upper game value is 2.

Hence the lower game value and upper game value are equal, so we can say that the saddle point exist.

11.11 Two person Zero-sum Game with Mixed Strategies

While solving a game problem, saddle point need not exist always. In such cases we use mixed strategies to find the value of the game. Every two person zero-sum game has a solution in mixed strategies. Let player A selects the strategy A_i with probability p_i and player B selects the strategy B_j with probability q_j, $i = 1,2........m;\ j = 1,2,.....n.$ A_i and B_j can be considered events with probabilities p_i and q_j *i.e.*, $P(A_c) = p_i$ and $p(B_j) = q_j$

$$
\begin{array}{cc|cccccc|}
 & & q_1 & q_2 & \cdots\cdots & q_j & \cdots\cdots & q_n \\
 & & B_1 & B_2 & \cdots\cdots & B_j & \cdots\cdots & B_n \\
\hline
p_1 & A_1 & a_{11} & a_{12} & \cdots\cdots & a_{1j} & \cdots\cdots & a_{1n} \\
p_2 & A_2 & a_{21} & a_{22} & \cdots\cdots & a_{2j} & \cdots\cdots & a_{2n} \\
\vdots & \vdots & . & & & & & \\
p_i & A_i & a_{i1} & a_{i2} & \cdots\cdots & a_{ij} & \cdots\cdots & a_{in} \\
\vdots & \vdots & \vdots & \vdots & & \vdots & & \vdots \\
p_m & A_m & a_{m1} & a_{m2} & \cdots\cdots & a_{mj} & \cdots\cdots & a_{mn} \\
\hline
\end{array}
$$

Fig 11.4- Matrix representing payoff to player A in mixed strategies

To find the value of the game; we have to find the value of $p_i's$ and $q_j's$. We do this by using minimax (maximin) criteria. If B selects pure strategy B_j then the expected payoff to player A is

$$a_{ij}p_1 + a_{2j}.p_2 + ... + a_{mj}\ p_m = \sum_{i=1}^{m} a_{ij}p_i$$

Player B can select only of the pure strategies $B_j\,(j = 1,......,n)$, and hence player A would like to select $p_i's$ in such a way that it maximizes its smallest expected payoff.

Thus $A's$ problem is

$$\underset{p_i,p_2,...p_m}{Max}\left[\min\left\{\sum_{i=1}^{m} a_{i1}p_i, \sum_{i=1}^{m} a_{i2}p_i, ..., \sum_{i=1}^{m} a_{in}p_i\right\}\right] \qquad ....(11.8)$$

Subject to the condition that $\sum_{i=1}^{m} p_i = 1, p_i's \geq 0$

Similarly player B would payoff like to select $q_j's$ which minimize the largest expected payoff to A.

Thus $B's$ problem is

$$\underset{q_i,q_2,...q_n}{Min}\left[\max\left\{\sum_{j=1}^{n} a_{1j}q_j, \sum_{j=1}^{n} a_{2j}q_j ...., \sum_{j=1}^{n} a_{mj}q_j\right\}\right] \qquad .....(11.9)$$

Subject to the condition $\sum_{j=1}^{n} q_j = 1, q_j's \geq 0\ for\ \forall_j$.

The value in (11.8) and (11.9) are maximin and minimax expected payoff's to A respectively by Let these be denoted by $\underline{a}$ and $\overline{a}$ respectively.

$$\underline{a} = Min\left\{\sum_{i=1}^{m} a_{i1}p_i, \sum_{i=1}^{m} a_{i2}p_i, ..., \sum_{i=1}^{m} a_{in}p_i\right\}$$

Then A'S problem given in (11.8) reduces to

$$Max\ Z_A = \underline{a} \qquad ....(11.10)$$

Subject to $\sum_{i=1}^{m} a_{ij}p_i \geq \underline{a}, j = 1,....,n, \sum_{i=1}^{m} p_i = 1,\ p_i's \geq 0$

In this case we do not know the value of $\underline{a}$, it may be positive or non-positive.

We shall assume that $\underline{a}$ is equal to some positive number. But to have $\underline{a}$ positive it is evidently sufficient for all elements a_{ij} of the payoff matrix to be non-negative.

This can always be attained by adding to elements a_{ij}, a sufficiently large positive number C, then the value of the game will be increased by C while the solution will remain unchanged.

The optimum value of the objective function is obtained by subtracting the added constant C.

Thus without any loss of generality, assume $\underline{a}$ to be positive.

Divide constraints in (11.10) by $\underline{a}$ and let $x_i = \frac{p_i}{\underline{a}}, i = 1,...m;$ then problem in equation (11.10) reduces to

$$Max\ Z_A = \underline{a} \qquad \text{......}(11.11)$$

s.t $\sum_{i=1}^{m} a_{ij} x_i \geq 1,\ j = 1,2,...,n.$

Also $\sum_{i=1}^{m} x_i = \sum_{i=1}^{m} \frac{p_i}{\underline{a}} = \frac{1}{\underline{a}} \sum_{i=1}^{m} p_i = \frac{1}{\underline{a}}$

Thus we have

$$\sum_{i=1}^{m} x_i = x_1 + x_2 + \text{......} + x_m = \frac{1}{\underline{a}} \qquad \text{.....}(11.12)$$

Max $Z_A = \underline{a}$ is equivalent to min $Z'_A = \frac{1}{\underline{a}}$ .....(11.13)

Thus A's problem can be written in the form,

$$Min\ Z_A' = x_1 + x_2 +, \text{.....} + x_m$$

s.t $\sum_{i=1}^{m} a_{ij} x_i \geq 1$ ....(11.14)

$$x_i's \geq 0,\ i = 1,2,\text{.....}m.$$

Next we come to B's problem, let $\bar{a} > 0.$

This can always be adjusted by adding a constant C to make all entries a_{ij} in the payoff matrix positive.

Thus proceeding exactly in the same way as in reducing A's problem to LPP form, the B's problem in (11.9) can be reduced in following LPP form,

$$Max\ Z_B = y_1 + y_2 + ... + y_n$$

$$s.t \quad \sum_{r=1}^{n} a_{ij} y_j \leq 1, i = 1,....,m \qquad ......(11.15)$$

$$y_j 's \geq 0, \; j = 1,2,...., n$$

It is important to observe that problems (11.14) and (11.15) are the dual of each other.

There are feasible $p_i 's$ and $q_j 's(i = 1,....m; \; j = 1,....n)$ because each player can use pure strategies.

Consequently there exist feasible $x_i 's$ and $y_j 's$ as $\underline{a}$ and $\bar{a}$ are positive.

As a result problem (11.14) and (11.15) have optimal solutions. If one problem has a solution, then other problem, also has a solution.

We prefer to solve that problem which has lesser number of constraints.

Examples

Example.5. Solve the following game the method of LP:

$$\begin{array}{cc} & B \\ & \begin{array}{ccc} B_1 & B_2 & B_3 \end{array} \\ A \begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} & \begin{bmatrix} 1 & 1 & 1 \\ 2 & -2 & 2 \\ 3 & 3 & -3 \end{bmatrix} \end{array}$$

Solution. The game matrix (payoff matrix of player A) is given below:

B \ A	B_1	B_2	B_3	α_i
A_1	1	1	1	1
A_2	2	-2	2	-2

A_3	3	3	-3	-3
β_j	3	3	2	

The lower value of the game $\underline{a}$ = max (1, -2, -3) = 1.

The upper value of the game $\bar{a}$ = min (3, 3, 2) =2. The value of the game lies between 1 and 2.

Let player A uses strategies A_i's with probabilities p_i, $\sum p_i = 1, i = 1, 2, 3$ and player B uses strategies B_j's with probabilities q_j's, $\sum q_j = 1, j = 1, 2, 3.$

$$\begin{array}{ccc} & y_1 & y_2 & y_3 \\ & q_1 & q_2 & q_3 \\ & B_1 & B_2 & B_3 \end{array}$$

$$\begin{array}{ccc} x_1 & p_1 & A_1 \\ x_2 & p_2 & A_2 \\ x_3 & p_3 & A_3 \end{array} \begin{bmatrix} 1 & 1 & 1 \\ 2 & -2 & 2 \\ 3 & 3 & -3 \end{bmatrix}$$

Player B's problem is

$$\text{Max } Z_B = y_1 + y_2 + y_3 = \frac{1}{V}$$

s.t. $y_1 + y_2 + y_3 \le 1$

$2y_1 - 2y_2 + 2y_3 \le 1$

$3y_1 + 3y_2 - 3y_3 \le 1$

$$y_j \ge 0, y_j = \frac{q_j}{V}, j = 1, 2, 3.$$

Introducing slack variables y_4, y_5, y_6, the standard form is

$$\text{Max } Z_B = \frac{1}{V} = y_1 + y_2 + y_3 + 0.y_4 + 0.y_5 + 0.y_6$$

s.t. $y_1 + y_2 + y_3 + y_4 = 1$

$2y_1 - 2y_2 + 2y_3 + y_5 = 1$

$3y_1 + 3y_2 - 3y_3 + y_6 = 1$

$$y_j \geq 0.$$

Using simplex method, we get

Cost →				1	1	1	0	0	0	Minimum Ratio $\frac{X_{B_i}}{\alpha_i^j}$
Variable →				y_1	y_2	y_3	y_4	y_5	y_6	
T. No.	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	
1	0	y_4	1	1	1	1	1	0	0	1
	0	y_5	1	2	-2	[2]	0	1	0	$\frac{1}{2}$ →
	0	y_6	1	3	3	-3	0	0	1	…
	c_j-Z_j			1	1	1 ↑	0	0	0	
2	0	y_4	1/2	0	[2]	0	1	-1/2	0	$\frac{1}{4}$ →
	1	y_3	1/2	1	-1	1	0	1/2	0	….
	0	y_6	5/2	6	0	0	0	3/2	1	….
	c_j-Z_j			0	2 ↑	0	0	-1/2	0	
3	1	y_2	1/4	0	1	0	1/2	-1/4	0	
	1	y_3	3/4	1	0	1	1/2	1/4	0	
	0	y_6	5/2	6	0	0	0	3/2	1	
	c_j-Z_j			0	0	0	-1	0	0	

This is the optimal table. Hence $y_1=0$, $y_2=1/4$, $y_3=3/4$.

Now we have $\frac{1}{V}=y_1+y_2+y_3=0+\frac{1}{4}+\frac{3}{4}=1 \Rightarrow V=1$.

The actual value is 1.

Using $y_j=\frac{q_j}{V}$, we have $q_1=0\times 1=0, \quad q_2=\frac{1}{4}\times 1=\frac{1}{4}, \quad q_3=\frac{3}{4}\times 1=\frac{3}{4}$.

Thus player B should strategies B_1, B_2 and B_3 with probabilities $q_1=0, \quad q_2=\frac{1}{4}, \quad q_3=\frac{3}{4}$.

A_i's best strategies appear in c_j-Z_j row under slack variables y_4, y_5 and y_6. Thus $x_1=1, \quad x_2=0, \quad x_3=0$.

Using $x_i=\frac{p_i}{V}$, we have $p_1=1\times 1=1, \quad p_2=0\times 1=0, \quad p_3=0\times 1=0$.

11.12 Summary

A game can be viewed as a clash of interests involving two or more opponents, referred to as a two-person or multi-person game. Two-person zero-sum games are alternatively known as rectangular games, as they are depicted by a payoff matrix in rectangular form. A saddle point in a payoff matrix is a position occupied by an element that is the minimum in its row and the maximum in its column.

When a single alternative is played with certainty, it is termed a pure strategy. In contrast, if multiple alternatives with different choices are employed during the game, it is referred to as a game with a mixed strategy. A game is deemed fair when both the lower and upper values of the game are equal to zero.

11.13 Terminal Questions

Q.1. Define two person zero sum game.

Q.2. Define saddle point and strategies.

Q.3. Find the lower and upper value of the game for the game matrix (payoff matrix of player A) given below:

$$\begin{array}{cc} & \text{Player } B \\ \text{Player } A & \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \end{array}$$

Does the saddle point exist?

Q.4. Find the lower and upper value of the game for the game matrix (payoff matrix of player A) given below:

$$\begin{array}{cc} & Player\,B \\ Player\,A & \begin{bmatrix} 7 & 10 \\ 5 & 6 \end{bmatrix} \end{array}$$

Does the saddle point exist?

Q.5. Solve the following game the method of LP:

$$\begin{array}{ccc} & & Player\ B \\ & & \begin{array}{cc} B_1 & B_2 \end{array} \\ Player\ A & \begin{array}{c} A_1 \\ A_2 \end{array} & \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \end{array}$$

Answer

3. Lower value = 0 and upper value =2, saddle point does not exists.

4. Lower value = 7 and upper value =7, saddle point exists.

5. $\begin{bmatrix} A_1 & A_2 \\ 1/2 & 1/2 \end{bmatrix}, \begin{bmatrix} B_1 & B_2 \\ 1/2 & 1/2 \end{bmatrix}$ and game value $= 0.$

UNIT-12: Goal Programming

Structure

12.1 Introduction

Goal programming serves as an extension of linear programming (LP), venturing into the realm of multi-objective programming where multiple objectives come into play. In this framework, there exists more than one objective, and the primary aim is to minimize the discrepancies between the actual outcomes and desired goals, prioritized according to assigned priorities for different goals.

The conceptualization of the goal programming model traces back to its initial presentation by Charnes and Cooper, who introduced it as an extension of the LP model, albeit without explicitly labeling it as the Goal Programming (GP) model.

12.2 Objectives

After reading this unit the learner should be able to understand about:

- formulation of goal programming
- single goal objective programming

- Goal Programming Algorithm
- multiple goal model
- multiple goal model with equal or no priorities

12.3 Formulation of Goal Programming (GP)

Goal programming (GP) involves the formulation of mathematical models to address problems with multiple, often conflicting, objectives. If there are m goals and p resource constraints in a problem then the most general GP model can be written in the following form:

Min $Z = \sum_{i=1}^{m} p_i \left(w_i^+ d_i^+ + w_i^- d_i^- \right)$

s.t. $\sum_{j=1}^{n} a_{ij} x_j + d_i^- - d_i^+ = b_i, \; i = 1, \ldots\ldots, m.$

$$\sum_{j=1}^{n} a_{ij} x_j \leq = \geq b_i, \; i = m+1, \ldots\ldots, m+p.$$

Where $x_j, j = 1, \ldots\ldots, n$ are *n* decision variables;

$p_i, i = 1, \ldots\ldots, m$ are the priorities associated with *m* goals

w_i^+ is the relative weight of d_i^+ in *i*th priority level

w_i^- is the relative weight of d_i^- in *i*th priority level

$d_i^- =$ derivational variable representing under achievements in *i*th goal

$d_i^+ =$ derivational variable representing over achievements in *i*th goal

Here priorities p_i's are not assigned any values, but this is simply a convenient way of indicating more importance of one goal over another. Thus if p_i is given more importance than p_j that means $p_i > n\, p_j$ however large *n* may be.

We also indicate this by writing $p_i >>> p_j$. At the same priority level p_i the deviational variables d_i^- and d_i^+ may be given different weight in the objective function.

The person formulating the GP model has to analyze each of *m* goals carefully. If over-achievements is acceptable d_i^+ (called surplus variable in LP) is removed from the objective function and if under-achievement is acceptable, d_i^- (called slack variable in LP formulation) is removed from the objective function.

If the exact achievement of the i^{th} goal is desired, then both d_i^- and d_i^+ must be included in the objective function and ranked to their priority order. In this manner the higher priority order goals are considered before the lower priority order goals.

The following some important definitions and terms are useful for goal programming problem:

Goal equation

A goal expressed in the form of an equation by using variables d_i^- (under-achievement) and d_i^+ (over-achievement) is known as goal equation. These variables d_i^- and d_i^+ are known as deviational variables.

Priorities in GP model

The coefficients of deviational variables in the objective function of GP are called priorities. Magnitude of priorities reflects the preference order of a goal.

Multiple goals with priorities and weights

A multiple goal models in which different weights are employed in one or more priority levels to distinguish the goal preference are called multiple goals with priorities and weights.

Goal Programming (GP) simplex method

The modified simplex method used in solving a Goal Programming problem is called Goal Programming simplex method.

Resource (structure) constraints

The constraints that do not directly related to the goals of the problem are called structural or resource constraints. Deviational variables are not in corporated into these constraints. These constraints have to be satisfied by the solution.

Trade-off function

The objective function of a Goal Programming is known as trade-off function. This may be linear or non-linear.

Example

Example.1. A company produces two types of products A and B. These products are produced in two different departments D_1 and D_2 of the company. Product A requires 2 hrs in department D_1 and 3 hrs in department D_2. Product B requires 2 hrs in department D_1 and 5 hrs in department D_2. Production time is limited in department D_1 to 60 hrs and in department D_2 to 80 hrs. The profit per unit of the products A and B is Rs. 3 and Rs. 6 respectively. To maximize the profit the company has set a high profit goal of Rs 1500. The management of the company desires to produce a least 30 units of each product A and B. The management is considering this second goal equally to be as important as the first goal which is maximizing the profit. Formulate the given problem as a goal programming problem.

Solution:

Profit constraint:

Let $x_1 =$ number of product A to be produced.

$x_2 =$ number of product B to be produced

Therefore, the profit goal is $3x_1 + 6x_2 \approx 1500$

The Profit goal equation is

$$3x_1 + 6x_2 + d_1^- - d_1^+ = 1500$$

Where $d_1^- =$ amount of under-achievement of profit.

$d_1^+ =$ amount of over-achievement of profit.

Resource (structure) constraints:

$$2x_1 + 3x_2 \leq 60$$

or $\quad 2x_1 + 3x_2 + x_3 = 60$

and $\quad 2x_1 + 5x_2 \leq 80$

or $\quad 2x_1 + 5x_2 + x_4 = 80$

Production constraint:

Since at least 30 units of each product A and B are desired to be produced, the production constraints can be written as:

$x_1 + d_2^- - d_2^+ = 30$ and $x_2 + d_3^- - d_3^+ = 30$

Here $x_1 \geq 30,\ x_2 \geq 30$ thus d_2^- and d_3^- are the deviational variables and are to be included in the objective function.

Objective function:

With priority ranked goals one objective function have to be formulated for each goal in the goal programming. The management have given equal priority to all the goals. Thus the complete GP model is formulated as

Min $Z = d_1^- + d_2^- + d_3^-$ (objective function)

s.t. $3x_1 + 6x_2 + d_1^- - d_1^+ = 1500$ (goal 1)

$x_1 + d_2^- - d_2^+ = 30$ (goal 2)

$x_2 + d_3^- - d_3^+ = 30$ (goal 3)

$2x_1 + 3x_2 + x_3 = 60$

$2x_1 + 5x_2 + x_4 = 80$ Resource constraints

Where $x_1, x_2, x_3, x_4, d_1^-, d_1^+, d_2^-, d_2^+, d_3^-, d_3^+ \geq 0;\ x_3, x_4$ are the slack variables d_1^+, d_2^+, d_3^+ are over-achievements for goals 1, 2 and 3 respectively.

12.4 Single Goal Programming

To solve a single goal programming problems we use the standard simplex method. To have clear cut understanding of the Goal Programming-Linear Programming relationship consider the following example:

Example

Example.2. A manufacturer products two models p_1 *and* p_2 which have to go through two machines m_1 and m_2 before getting them in final forms. The machine time available, time required by each product on two machines and the profit on each product is given as her the following table:

Product→ / Machine↓	Product		Available time (hours)
	p_1	p_2	
M_1	2	3	60
M_2	2	5	80
Profit	3	5	

Find how the manufacturer can earn maximum profit. Further if the manufacturer fives the target of achieving maximum profit of Rs. 200 then formulate the problem in GP model.

Solution: The LP formulation of the problem is

$$\left.\begin{aligned} &\text{Max } Z = 3x_1 + 5x_2 \\ &\quad \text{s.t } 2x_1 + 3x_2 \le 60 \\ &\quad 2x_1 + 5x_2 \le 80 \\ &\quad x_1, x_2 \ge 0 \end{aligned}\right\} \quad \ldots\ldots (12.1)$$

Where $x_1 =$ number produced of p_1 product and $x_2 =$ number produced of p_2 product

So the manufacturer objective is to find $x_1\ and\ x_2$ so that profit is maximum. If we solve it by simplex method the maximum profit $Z = 95,\ for\ x_1 = 15\ and\ x_2 = 10$.

Now in the above problem (12.1), suppose manufacturer fixes the target of achieving maximum profit of Rs.200, then the goal programming formulation of the problem is:

$$\left.\begin{aligned} &\text{Min } Z = d_1^- \text{ or Max } -Z = -d_1^- \\ &\text{s.t.} \quad 2x_1 + 3x_2 \le 60 \\ &\quad 2x_1 + 5x_2 \le 80 \\ &3x_1 + 5x_2 + d_1^- - d_1^+ = 200 \end{aligned}\right\} \quad \ldots\ldots (12.2)$$

$x_1, x_2, d_1^-, d_1^+ \geq 0$.

Here $x_i's$ are the decision variables and d_i^- (under achievement) and d_1^+ (over achievement) are the derivational variables. Taking x_3 *and* x_4 as slack variables added to constraints of Equation (12.2), we solve the GP models as follows:

Costs→				0	0	0	0	0	-1	Min ratio
T.No.	C_B	Basic Variables	X_B	x_1	x_2	d_1^+	x_3	x_4	d_1^-	
1	0	x_3	60	2	3	0	1	0	0	60/3 = 20
	0	x_4	80	2	[5]	0	0	1	0	80/5=16→
	-1	d_1^-	200	3	5	-1	0	0	1	200/5 = 40
	$c_j - Z_j$			3	5↑	-1	0	0	0	
2	0	x_3	12	[4/5]	0	0	1	-3/5	0	60/4 =15→
	0	x_2	16	2/5	1	0	0	1/5	0	16/(2/5) = 40
	-1	d_1^-	120	1	0	-1	0	-1	1	120
	$c_j - Z_j$			1↑	0	-1	0	-1	0	
	0	x_1	15	1	0	0	5/4	-3/4	0	
	0		10	0	1	0	-1/2	1/2	0	

3	0	x_2 d_1^-	105	0	0	-1	-5/4	-1/4	0	
	$c_j - Z_j$			0	0	-1	-5/4	-1/4	0	

The optimum solution is $x_1 = 15$, $x_2 = 10$ $d_1^{-1} = 105$.

Here in this problem 105 is under achievement of the maximum profit goal 200. Thus actual profit is 200 – 105 = 95 which is the same as the maximum profit obtained in LP model.

Therefore in a single goal programming problem the solution of the problem remains the same as obtained using LP model of the problem.

The main difference between the LP model (12.1) and GP model (12.2) is the objective function. In GP model we include deviational variable(s) in the objective function. Also the problem in LP model may be a maximization or minimization but in GP model it is always a minimization problem where we minimize the some (weighted) of deviations.

12.5 Goal Programming Algorithm

The standard simplex method can easily be used in solving goal programming problems. This is accomplish by assigning values to the priority coefficients in the objective function of G.P formulation so that the values reflect the same order of relationship as the priorities.

In the chapter on linear programming, we have already described in detail the steps and procedure of simplex method. Here we shall describe how the simplex method algorithm can be modified to solve a goal programming problem.

The following steps are used in G.P algorithm:

Step 1: Construct the initial modified simplex table which is similar to that of simplex table with only difference in $(c_j - z_j)$ row, where it is splitted into as many different goals row as the number of priorities assigned.

Step 2: Check for optimality. If there is no positive entry in $(c_j - Z_j)$ row for highest priority row p_k, then the priority p_k goal has been met and go to step 6 otherwise go to step 3.

Step 3: Determine new entering variable by identifying the largest positive entry in p_k row. This fixes the column of entering variable.

Step 4: Determine the departing variable by considering the minimum of ratio of X_B column entries with corresponding non-negative entries of the column fixed in step 3 as it is done in standard simplex method.

Step 5: Develop the new table to update the coefficients in the body of the table by using elementary row operations. The new $(c_j - Z_j)$ rows are computed in the same manner as would be used in the simplex method. The only difference is the tabular representation. As an example compute Z_j by multiplying the values in the j^{th} column with corresponding entries of X_B column. Then subtract it from c_j to get $(c_j - Z_j)$. Break $(c_j - Z_j)$ into parts, where parts are associated with priority levels.

To be specific, if there are three priorities p_1, p_2 *and* p_3 arranged in the order of their importance and if $c_j - Z_j = 0 + 3p_3 - 4p_1 + 0$ then in the j^{th} column we split it as

$(c_j - Z_j)row:$	p_1	-4	$(p_1 >> p_2 >> p_3)$
	p_2	0	
	p_3	3	

This way continue steps 3, 4 and 5 till optimality conditions are satisfied for highest priority p_k row and go to step 6.

Step 6: Evaluate the next-lower priority level row $p_{k+1}(p_{k+1} << p_k)$ by identifying the largest positive coefficient for which there are no negative coefficients at a higher priority in the same column. If a tie exists in the values of the coefficients that determine the entering variable, break it arbitrarily.

Note: The lower priority goals must not be satisfied at the expense of higher priority goals.

12.6 Multiple Goal Models

There are three types of multiple goals models:

1. Multiple goals models with equal (no) priority

2. Multiple goals models with priority

3. Multiple goals models with priority and weights

For practical purposes, the multiple goals models are most useful in day to day life.

12.7 Multiple Goal Models with equal or no priority

The multiple goal equal priority model of GP is not of much practical value as compared to GP models with different priorities attached to different goals. However this model is easy to handle. We illustrate the multiple goal equal priority models through an example given below.

Example

Example.3. A firm is manufacturing B_1 *and* B_2 types of bags. Type B_1 bags are ordinary and type B_2 are luxury bags. These bags are processed through two machines M_1 *and* M_2 to get them in final form. Type B_1 bag requires 20 hours of machine M_1 and 10 hours of machine M_2. Type B_2 bag requires 10 hours each of machine M_1 *and* M_2. Time available on machine M_1 *and* M_2 is 60 hours and 40 hours respectively. The profit earned per bag is Rs.40 and Rs.80 on bags M_1 *and* M_2 respectively. The firm wants to maximize profit as much as Rs.1000. In addition to the profit goal, the firm wants to produce at least two bags of each type and consider this second goal equally as important as the first profit goal. Formulate and solve the given problem as a goal programming problem. Interpret the solution.

Solution: Formulation:

Let $x_1 =$ number of type B_1 bags manufactured and $x_2 =$ number of type B_2 bags manufactured.

Resource constraints:

Time taken on machine M_1 to produce x_1, x_2 number of bags of type B_1 and $B_2 = 20x_1 + 10x_2$. Thus we have

$$20x_1 + 10x_2 \le 60$$

Similarly for machine M_2, we have

$$10x_1 + 10x_2 \le 40$$

Profit constraints:

$$40x_1 + 80x_2 + d_1^- - d_1^+ = 1000$$

Production constraints:

As $x_1 \le 2, x_2 \le 2,$ the production constraints can be written as

$$x_1 + d_2^- - d_2^+ = 2, \quad x_2 + d_3^- - d_3^+ = 2.$$

As the equal priority (no priority) has been given to all the goals, the objective function is minimize the sum of deviational: Min $Z = d_1^- + d_2^- + d_3^-$

Thus the complete GP model can be described as

Min $Z = d_1^- + d_2^- + d_3^-$

s.t. $20x_1 + 10x_2 \le 60$

$10x_1 + 10x_2 \le 40$

$40x_1 + 80x_2 + d_1^- - d_1^+ = 1000$ (goal-1)

$x_1 + d_2^- - d_2^+ = 2$ (goal-2)

$x_2 + d_3^- - d_3^+ = 2$ (goal-3)

Where $x_1, x_2, d_1^-, d_1^+, d_2^-, d_2^+, d_3^-, d_3^+ \ge 0$. d_1^-, d_2^-, d_3^- are the under-achievement d_1^+, d_2^+, d_3^+ are the over-achievement of goals 1, 2 and 3 respectively.

The canonical form is

Min $Z = d_1^- + d_2^- + d_3^-$ or Max $-Z = -d_1^- - d_2^- - d_3^-$

s.t. $20x_1 + 10x_2 + x_3 = 60$

$10x_1 + 10x_2 + x_4 = 40$

$40x_1 + 80x_2 + d_1^- - d_1^+ = 1000$ (goal-1)

$x_1 + 0x_2 + d_2^- - d_2^+ = 2$ (goal-2)

$0x_1 + x_2 + d_3^- - d_3^+ = 2$ (goal-3)

Where $x_1, x_2, x_3, d_1^-, d_1^+, d_2^-, d_2^+, d_3^-, d_3^+ \ge 0$. x_3, x_4 are the slack variables d_1^-, d_2^-, d_3^- are the under achievements d_1^+, d_2^+, d_3^+ are the over achievements of goals 1, 2 and 3 respectively.

To solve the problem we use the modified simplex method:

$c_j \rightarrow$				0	0	-1	0	-1	0	-1	0	0	0	Min ratio
T. No.	C_B	B.V.	X_B	x_1	x_2	d_1^-	d_1^+	d_2^-	d_2^+	d_3^-	d_3^+	x_3	x_4	

1	0	x_3	60	20	10	0	0	0	0	0	0	1	0	60/10 = 6
	0	x_4	40	10	10	0	0	0	0	0	0	0	0	40/10 =4
	-1	d_1^-	1000	40	80	1	-1	0	0	0	0	0	0	1000/80
	-1	d_2^-	2	1	0	0	0	1	-1	0	0	0	0	……..
	-1	d_3^-	2	0	[1]	0	0	0	0	1	-1	0	0	2/1=2 min
$c_j - Z_j$				41	81↑	0	-1	0	-1	0	1	0	0	
2	0	x_3	40	20	0	0	0	0	0	-10	10	1	0	40/10 =4
	0	x_4	20	10	0	0	0	0	0	-10	[10]	0	1	20/10=2 min
	-1	d_1^-	840	40	0	1	-1	0	0	-80	80	0	0	840/80
	-1	d_2^-	2	1	0	0	0	1	-1	0	0	0	0	……
	0	x_2	2	0	1	0	0	0	0	1	-1	0	0	…..
$c_j - Z_j$				41	0	0	-1	0	-1	-81	80↑	0	0	
	0	x_3	20	10	0	0	0	0	0	0	0	1	-1	

	0	d_3^+	2	1	0	0	0	0	0	-1	1	0	1/10	
3	-1	d_1^-	680	-40	0	1	1	0	0	0	0	0	-8	
	-1	d_2^-	2	1	0	0	0	1	-1	0	0	0	0	
	0	x_2	4	1	1	0	0	0	0	0	0	0	1/10	
$c_j - Z_j$				-39	0	0	-1	0	-1	-1	0	0	-8	

The solution is $x_1 = 0, x_2 = 4, x_3 = 20, x_4 = 0, d_1^- = 680, d_1^+ = 0, d_2^- = 2, d_2^+ = 0, d_3^- = 0, d_3^+ = 2$

and $Z = 682$.

Interpretation of the solution obtained:

Here $d_1^- = 680$ implies that profit goal of Rs.1000 is under achievement by Rs.680. Thus actual profit is 1000 – 680 = 320. Also $d_2^- = 2$ indicates, the production goal of type B_1 bags was missed by 2.

Further $d_3^+ = 2$ indicates that production goal of type B_2 bags was over achievement by 2.

12.8 Summary

The Goal Programming (G.P) model serves as an extension of the Linear Programming (L.P) model. A solution in single-objective goal programming closely resembles that obtained through the L.P method. Within goal programming, the assignment of priorities to different goals facilitates the prioritized accomplishment of the most significant goal. Essentially, lower-priority goals can be achieved first, while higher-priority goals are pursued at a potentially higher cost. Goal programming consistently poses a minimization challenge, focusing on minimizing deviations from established goals within a predefined set of constraints.

12.9 Terminal Questions

Q.1. Write a short note on Goal Programming problem.

Q.2. Explain the Single and multi-goal programming problem.

Q.3. A manufacturer produces two types of products A and B. The production of each of A and B requires one hr of production capacity in the workshop. The workshop of the company has maximum production capacity of 80 hr per week. On account of restriction on the sales capacity of these products, maximum number of A and B that can be sold are respectively 8 and 10 per week. The profit on the sale of A is Rs. 100 and B is Rs. 60.

The manager wants to determine the number of units of each product that should be produced per week and sets the following goal to achieve with equal priority:

Goal 1: The production capacity should not exceed 50 hr per week.

Goal 2: The sales of the two products A and B must be as much as possible.

Formulate the problem as GP problem.

Q.4. Solve the following GP:

Min $Z = p_1 d_1^- + p_2 d_2^- + p_3 d_3^-$

s.t. $200x_1 + 300x_2 + d_1^- - d_1^+ = 4800$

$x_2 + d_2^- - d_2^+ = 15$

$x_1 + d_3^- - d_3^+ = 5$

$4x_1 + 2x_2 + x_3 = 60$

$4x_1 + 4x_2 + x_4 = 70$

and $x_1, x_2, x_3, x_4, d_1^-, d_1^+, d_2^-, d_2^+, d_3^-, d_3^+ \geq 0.$

Answer

3. Min $Z = d_1^+ + d_2^- + d_3^-$

s.t. $x_1 + x_2 + d_1^- - d_1^+ = 80$

$$x_1 + d_2^- = 8$$

$$x_2 + d_3^- = 10$$

and $x_1, x_2, d_1^-, d_1^+, d_2^-, \; d_3^- \geq 0.$

4. $x_1 = 5/2,\; x_2 = 15,\; d_3^- = 5/2, x_3 = 5, d_1^+ = 200.$

UNIT-13: Integer Programming Problem-I

Structure

13.1 Introduction

In various crucial real-life scenarios, the necessity arises for decision variables to be integers. In Linear Programming Problems (LPP), decision variables typically have the flexibility to assume any non-negative values, whether integer or fractional. However, certain practical problems, including capital budgeting, capacity expansion, shipping schedules, construction schedules, and location selection, demand that decision variables take on only integral values. An LPP wherein some or all decision variables in the optimal solution are constrained to non-negative integral values is termed an Integer Linear Programming Problem (ILPP).

This chapter will delve into several methods for solving ILPPs, including: (i) Gomory's Cutting Plane Method for All Integer Linear Programming Problems and (ii) Gomory's Cutting Plane Method for Mixed Integer Linear Programming Problems. These techniques provide effective approaches to address the challenges posed by decision variables restricted to integer values in real-world problems.

13.2 Objectives

After reading this unit the learner should be able to understand about:

- the Integer Linear Programming Problem
- Types of Integer Linear Programming Problem
- All Integer Linear Programming Problem
- Mixed Integer Linear Programming Problem
- Zero-One Integer Linear Programming Problem
- Gomory's Cutting Plane Method for All Integer Linear Programming Problem
- Algorithm for Gomory's Cutting Plane Method
- Gomory's Cutting Plane Method for Mixed Integer Linear Programming Problem
- Difference between Gomory's Cutting Plane Method for All Integer Linear Programming Problem and Mixed Integer Linear Programming Problem

13.3 Formulation of Integer Linear Programming Problem (ILPP)

The mathematical formulation of Integer Linear programming problem (ILPP) is

$$\text{Max or Min } Z = C^T X$$

$$AX = b$$

$$\text{and } \quad X \geq 0,$$

Where $X = (x_1, x_2, \ldots\ldots., x_n)^T \geq 0$;

x_j's are integer, $j = 1, 2, \ldots\ldots.., n$;

$C = (c_1, c_2, \ldots\ldots.., c_n)^T$;

$b = (b_1, b_2, \ldots\ldots.., b_m)^T$.

Here are some common important techniques for solving Integer Linear Programming problems (ILPP):

Branch and Bound

This technique systematically explores the solution space by dividing it into smaller sub problems. It uses bounds to eliminate sub problems that cannot yield an optimal integer solution, thereby reducing the search space.

Gomory's Cutting Plane Method

Gomory's method involves iteratively adding linear constraints (cutting planes) to the linear programming relaxation of the Integer Linear Programming problem. These additional constraints help tighten the solution space, leading to an integer solution.

Branch and Cut

Branch and Cut is an extension of the Branch and Bound method that incorporates cutting planes. It combines the strengths of both techniques to efficiently explore the solution space while gradually tightening the bounds.

Integer Branch and Bound

This is a modification of the traditional Branch and Bound method specifically designed for Integer Linear Programming problems. It uses branching to explore integer solutions and employs bounding techniques to eliminate non-optimal branches.

Dynamic Programming

Dynamic programming approaches are applicable to certain types of Integer Linear Programming problems. They involve breaking down the problem into smaller sub problems and solving them systematically, often with recursive formulations.

Heuristic Methods

Heuristic methods, such as simulated annealing, genetic algorithms, or tabu search, are optimization techniques that may be applied to find near-optimal solutions for large-scale Integer Linear Programming problems in a reasonable amount of time.

Mixed Integer Linear Programming (MILP)

Specialized solvers designed for Mixed Integer Linear Programming problems, which include both continuous and integer decision variables, are available in optimization software packages. These solvers often use a combination of the above techniques.

Examples

Example.1. A company has to manufacture the circular tops of cylindrical cans. Two sizes: One of smaller diameter 10 cm and other of bigger diameter 20 cm are required. They are to be cut from metal sheets of dimensions 20 cm by 70 cm. The requirement of smaller size is 30,000 and of larger size 10,000. How to cut the tops from metal sheets so that the number of sheets used is minimized. Formulate the Integer Linear Programming Problem.

Solution Suppose the sheets of size 20 cm by 70 cm be cut in the following four patterns:

Pattern 1, where entire sheet is used to cut it into the 14 taps of 10 cm diameter.

Pattern 2, one 20 cm diameter top and ten 10 cm diameter tops are cut.

Pattern 3, two 20 cm diameter and six 10 cm diameter tops are cut.

Pattern 4, three 20 cm diameter and two 10 cm diameter tops are cut from the given sheet.

Let x_1, x_2, x_3 and x_4 be the number of sheets cut according to first, second, third and fourth patterns respectively.

Then the Integer Linear Programming Problem is

$$\text{Max } Z = x_1+x_2+x_3+x_4$$

$$\text{s.t. } 14x_1+10x_2+6x_3+2x_4 \geq 30{,}000$$

$$x_2+2x_3+3x_4 \geq 10{,}000$$

$$x_1, x_2, x_3, x_4 \geq 0 \text{ and integers.}$$

13.4 Types of Integer Linear Programming Problem (ILPP)

Integer Linear Programming Problem can be classified into following three types:

1. All Integer Linear Programming Problem (AILPP)

2. Mixed Integer Linear Programming Problem (MILPP)

3. Zero-One Integer Linear Programming Problem

13.5 All Integer Linear Programming Problem (AILPP)

An Linear programming problem is said to be an all Integer Linear Programming Problem if it contains all the decision variables restricted to integer values.

13.6 Mixed Integer Linear Programming Problem (MILPP)

An Linear programming problem is said to be an all Integer Linear Programming Problem if it contains some of the decision variables (not all) restricted to integer values.

13.7 Zero-One Integer Linear Programming Problem

An Linear programming problem is said to be an all Integer Linear Programming Problem if it contains all the decision variables restricted to take value either 0 or 1.

13.8 Gomory's Cutting Plane Method for All Integer Programming Problem

In 1956, R.E. Gomory's devised a method to determine the all-integer solution of an Integer Linear Programming Problem (ILPP). He accomplished this using the dual simplex method, creating a systematic approach based on generating a sequence of linear inequalities, referred to as cuts. The boundary of these cuts is termed the cutting plane. The key characteristic of Gomory's cutting plane method is that the cuts, or additional linear constraints, are generated in a way that avoids severing the portion of the original feasible solution space containing an integer solution. This method ensures the generation of additional linear constraints in a systematic manner, guaranteeing an integer solution to a given Linear Programming Problem (LPP) within a finite number of steps.

The procedure involves initially applying the simplex method to solve the Integer Linear Programming Problem. If, in the optimal simplex table, all decision variables are integers, then it serves as the solution to the ILPP. However, if any decision variable is not an integer, the method identifies the basic variable (denoted as x_r) in the optimal table with the largest fractional value among all basic variables, restricted to being integers. This step facilitates the systematic generation of cutting planes to move towards an all-integer solution. Then the row (chosen) corresponding to x_r in the optimal simplex table can be written as

$$x_{B_r} = 1.x_r + \sum_{j \neq r} a_{rj} x_j \quad or \quad x_{B_r} = x_r + \sum_{j \neq r} a_{rj} x_j \qquad \text{….(13.1)}$$

Where $x_j,\ j \neq r$ are all the non-basic variables of the chosen row. Let us now decompose the coefficients of x_r, x_j and x_{B_r} into integer and non-negative fractional parts in (13.1) to get

$$[x_{B_r}] + f_r = x_r + \sum_{j \neq r} \{[a_{rj}] + f_{rj}\} x_j \qquad \text{....(13.2)}$$

Where $[x_{B_r}]$ and $[a_{rj}]$ are the greatest integers in x_{B_r} and $a_{rj},\ j \neq r$ respectively. Rearranging equation (13.2) so that all integer coefficients appear on the left hand side, we get

$$f_r + \left\{[x_{B_r}] - x_r - \sum_{j \neq r} [a_{rj}] x_j\right\} = \sum_{j \neq r} f_{rj} x_j \qquad \text{....(13.3)}$$

Where f_r is strictly positive fraction ($0 < f_r < 1$) and $f_{rj,\, r \neq j}$ is a non-negative fraction ($0 \leq f_{rj} < 1$). In equation (13.3) right hand side is positive and the term f_r on left hand side is a strictly positive fraction. Thus for equation (13.3) to hold good, the total of terms in the bracket on left hand side must be a non-negative integer (as all the variables including slacks assume integer values). Thus the equation (13.3) can be converted into an inequality

$$f_r \leq \sum_{j \neq r} f_{rj} x_j \qquad \text{....(13.4)}$$

or $$\sum_{j \neq r} f_{rj} x_j = f_r + s_g$$

or $$-f_r = s_g - \sum_{j \neq r} f_{rj} x_j \qquad \text{....(13.5)}$$

where s_g is a non-negative slack variable. Equation (13.5) represents Gomory's cutting plane constraint. When this new constraint is added to the bottom of simplex table, it would create a new additional row in the table along with a column of the new basic variable s_g.

13.9 Algorithm for Gomory's Cutting Plane Method

The following steps can be used to solve an Integer Linear Programming Problem by Gomory's cutting plane method:

Step-1: First solve the given Integer Linear Programming Problem to get its continuous solution using simplex method. If all the decision variables in the optimal table are integers, then problem is done. If not then go to step-II.

Step-II: Identify the row of the basic variable say x_r with largest fractional part in the optimal simplex table of step-I and write it down as an equation. Decompose the coefficients of all terms into integer and non-negative fractional part using the idea of greatest integer function.

Note: The greatest integer in a number x is the largest integer and is denoted as $[x]$. As an example greatest integer in 2.1, denoted as [2.1] =2 and greatest integer in -1.1= [-1.1] =-2.

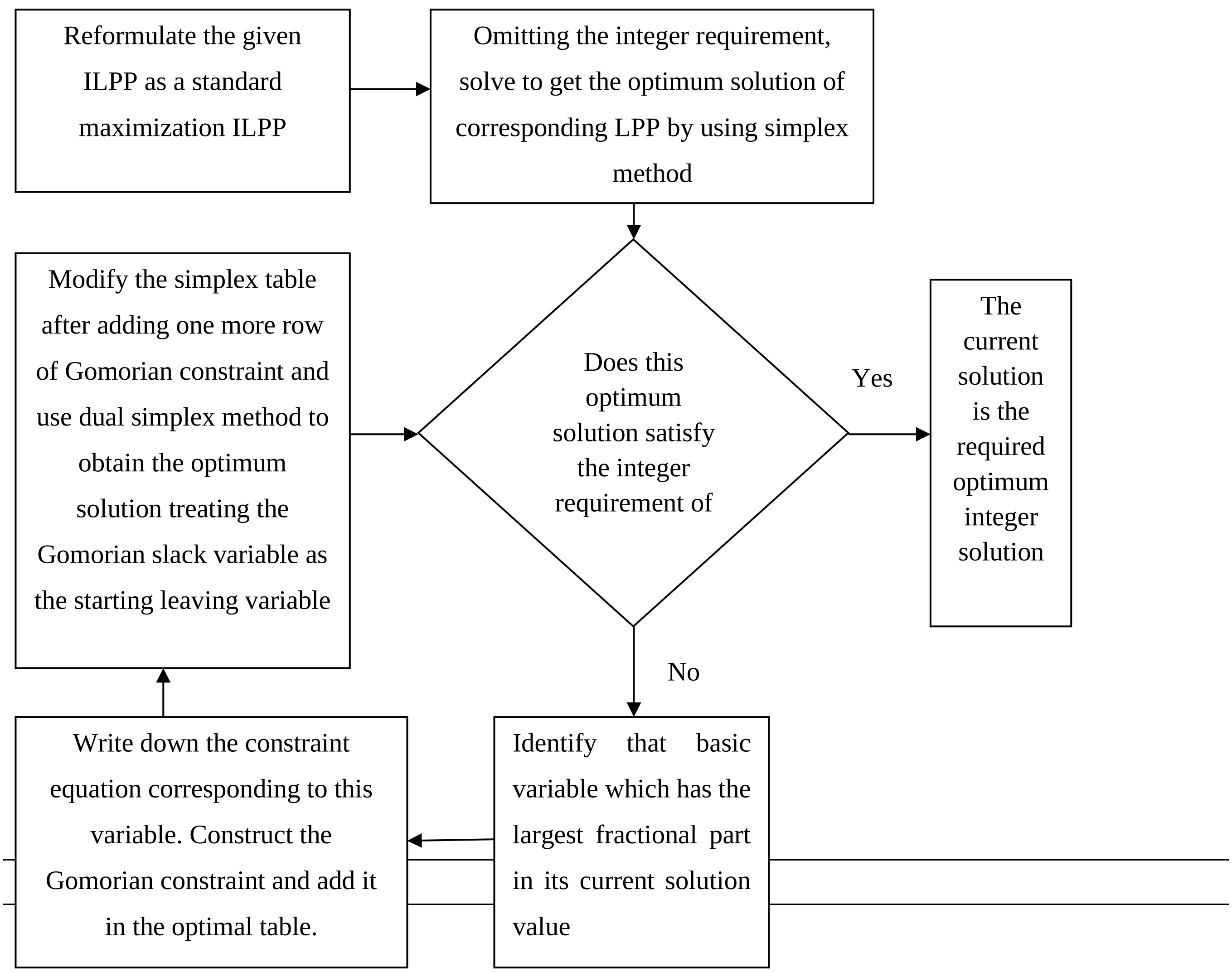
Step-III: Construct an addition constraint (Gomory's cut)

$$-f_r = s_g - \sum_{j \neq r} f_{rj} x_j$$

Where f_r ($0 < f_r < 1$) is strictly positive fractional part of the basic variable x_r (step-I), s_g is the non-negative slack variable and f_{rj}, $r \neq j$ are the non-negative fractions of the coefficients of non-basic variables in the r^{th} row equation.

Step-IV: Add the constraint of step-III as a row at the bottom of simplex table of step-I and use dual simplex method to get optimal table. If the solution is all integer then problem is done. If not, then repeat step-II through step-IV till the final optimal solution of AILPP is obtained.

All above steps of Gomory's algorithm can be summarized in the form of following flow chart:



Example.2. Find the all integer solution to the following Integer Linear Programming Problem (ILPP):

$$\text{Max } Z = 3x_1 + 2x_2$$

s.t. $x_1 + x_2 \leq 4$

$x_1 - x_2 \leq 2$

$x_1, x_2 \geq 0$ and integers.

Solution. Reformulated Integer Linear Programming Problem as a standard maximization Integer Linear Programming Problem is

$$Max\ Z = 3x_1+2x_2+0x_3+0x_4$$

s.t. $x_1+x_2+x_3 = 4$

$x_1-x_2+x_4 = 2$

$x_1, x_2, x_3, x_4 \geq 0$ and integers; x_3, x_4 are slack variables.

Now omitting the integer requirement, we solve the given LPP using simplex method:

Cost →				3	2	0	0	Minimum Ratio
Variable →				x_1	x_2	x_3	x_4	$\frac{X_{B_i}}{\alpha_i^j}$
Table No.	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	
1	0	x_3	4	1	1	1	0	$\frac{4}{1}=4$
	0	← x_4	2	[1]	-1	0	1	$\frac{2}{1}=2$ →
	c_j-Z_j			3↑	2	0	0	
2	0	← x_3	2	0	[2]	1	-1	$\frac{2}{2}=1$ →
	3	x_1	2	1	-1	0	1	
	c_j-Z_j			0	5↑	0	-3	
3	2	x_2	1	0	1	1/2	-1/2	

	3	x_1	3	1	0	1/2	1/2	
	c_j-Z_j			0	0	-5/2	-1/2	

The optimum solution is x_1=3, x_2=1 and Max Z =11. This optimum solution satisfies the integer requirement of the given ILPP and hence is its optimum solution and no further application of Gomory's cutting plane method is needed.

Example.3. Find the all integer solution to the following Integer Linear Programming Problem (ILPP):

$$\text{Max } Z = 2x_1 - x_2$$

s.t. $x_1 + x_2 \leq 2$

$x_1 - x_2 \leq 1$

$x_1, x_2 \geq 0$ and integers.

Solution Reformulated Integer Linear Programming Problem as a standard maximization Integer Linear Programming Problem is

$$Max\ Z = 2x_1 - x_2 + 0x_3 + 0x_4$$

s.t. $x_1 + x_2 + x_3 = 2$

$x_1 - x_2 + x_4 = 1$

$x_1, x_2, x_3, x_4 \geq 0$ and integers; x_3, x_4 are slack variables.

Now omitting the integer requirement in above ILPP, we solve the resulting LPP as below:

Cost →				2	-1	0	0	Minimum Ratio $\frac{X_{B_i}}{\alpha_i^j}$
Variable →				x_1	x_2	x_3	x_4	
Table No.	$\mathbf{C_B}$	Basic Variable	$\mathbf{X_B}$	α^1	α^2	α^3	α^4	

	0	x_3	2	1	1	1	0	$\frac{2}{1}=2$
1	0	$\leftarrow x_4$	1	$\boxed{1}$	-1	0	1	$\frac{1}{1}=1\rightarrow$
	c_j-Z_j			$2\uparrow$	-1	0	0	
2	0	$\leftarrow x_3$	1	0	$\boxed{2}$	1	-1	$\frac{1}{2}\rightarrow$
	2	x_1	1	1	-1	0	1	….
	c_j-Z_j			0	$1\uparrow$	0	-2	
3	-1	x_2	1/2	0	1	1/2	-1/2	
	2	x_1	3/2	1	0	1/2	1/2	
	c_j-Z_j			0	0	-1/2	-3/2	

This is optimal simplex table. The optimum solution is x_1=3/2, x_2=1/2 and Max Z =5/2, which is not all integer solution.

To get all integer solution we select the basic variable corresponding to maximum fractional value f_{B_i}.

Here $x_{B_1} = I_{B_1} + f_{B_1} = 0 + \frac{1}{2}$

and $x_{B_2} = I_{B_2} + f_{B_2} = 1 + \frac{1}{2}$.

So $f_{B_1} = f_{B_2} = \frac{1}{2}$.

Therefore $\max\left(f_{B_1}, f_{B_2}\right) = \max\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{1}{2}$, which is the same for $x_2 (= x_{B_1})$ and $x_1 (= x_{B_2})$ basic variable rows so either of the two basic row equations can be chosen to construct Gomorian cut.

Let us choose basic variable x_1 row equation (second row of the optimal table). Cut is

$$-f_{B_2} = -f_{23}x_3 - f_{24}x_4 + s_{g_2}$$

or $$-\frac{1}{2} = -\frac{1}{2}x_3 - \frac{1}{2}x_4 + s_{g_2}$$

or $$-\frac{1}{2} = 0x_1 + 0x_2 - \frac{1}{2}x_3 - \frac{1}{2}x_4 + s_{g_2}$$

Adding this Gomorian cut in above simplex table 3, we get

Cost →				2	-1	0	0	0
Variable →				x_1	x_2	x_3	x_4	s_{g_2}
Table No.	**CB**	Basic Variable	**XB**	α^1	α^2	α^3	α^4	α^{g_2}
4	-1	x_2	1/2	0	1	1/2	-1/2	0
	2	x_1	3/2	1	0	1/2	1/2	0
	0	← s_{g_2}	-1/2	0	0	[-1/2]	-1/2	1
	c_j-Z_j			0	0	-1/2↑	-2	0

The solution is infeasible. Restore feasibility using dual simplex method. When s_{g_2} leaves then x_3 enters $\left\{ as \min\left(\left|\frac{c_3 - z_3}{\alpha_3^3}\right|, \left|\frac{c_4 - z_4}{\alpha_3^4}\right| \right) = \min(1, 4) = 1 \;\; for \; x_3 \; column \right\}$ to get the next simplex table:

Cost →				2	-1	0	0	0
Variable →				x_1	x_2	x_3	x_4	s_{g_2}
Table No.	$\mathbf{C_B}$	Basic Variable	$\mathbf{X_B}$	α^1	α^2	α^3	α^4	α^{g_2}
5	-1	x_2	0	0	1	0	-1	1
	2	x_1	1	1	0	0	0	1
		x_3	1	0	0	1	1	-2
	0							
	c_j-Z_j			0	0	0	-3/2	-1

This is the all integer optimal simplex table. The all integer optimal solution of the given ILPP is $x_1 = 1$, $x_2 = 0$ and Max Z = 2.

13.10 Gomory's Cutting Plane Method for MILPP

If none of the variables in an Integer Linear Programming Problem (ILPP) are constrained to be integers, it is referred to as a Mixed Integer Linear Programming Problem (MILPP). The process of constructing a cut in a MILPP differs from that in an All Integer Linear Programming Problem (AILPP). In the case of a MILPP, the construction of a cut begins with the optimal solution of the Linear Programming Problem (LPP) obtained by applying the simplex method, disregarding the integer requirement(s).

Unlike in AILPP, where the focus is on generating cuts to preserve the integer feasibility, in MILPP, the initial emphasis is on obtaining an optimal solution within the relaxed context of non-integer variables.

The subsequent steps involve introducing additional constraints, or cuts, to progressively move towards a solution that satisfies the integer requirements while maintaining optimality. This distinction reflects the nuanced approach needed when dealing with Mixed Integer Linear Programming. Now from the optimal table thus obtained, choose the i^{th}-row corresponding to a basic variable x_{B_i} which has largest fractional value amongst those required to be integers.

The x_{B_i} -row in the optimal table has the form

$$x_i + \sum_{j=1}^{p} f_{ij}^{+} x_j + \sum_{j=p+1}^{q} f_{ij}^{-} x_j = b_i$$

or

$$x_i + \sum_{j=1}^{p} f_{ij}^{+} x_j + \sum_{j=p+1}^{q} f_{ij}^{-} x_j = [b_i] + f_{b_i} \qquad \text{….(13.6)}$$

where x_j, j=1 to p and j=p+1 to q are the non-basic variables with positive coefficients f_{ij}^{+} and negative coefficients f_{ij}^{-} respectively. $f_{b_i}, 0 < f_{b_i} < 1$ is the fractional part of right hand side and [b_i] is greater integer in b_i, which is less than or equal to b_i.

Equation (13.6) can be written as

$$\sum_{j=1}^{p} f_{ij}^{+} x_j + \sum_{j=p+1}^{q} f_{ij}^{-} x_j = f_{b_i} + \{[b_i] - x_i\} \qquad \text{….(13.7)}$$

The right hand side in equation (13.7) is a number which may be ≤ 0 or ≥ 0.

Case-I: Consider the case when $f_{b_i} + \{[b_i] - x_i\} \geq 0.$ Then $[b_i] - x_i = 0$ or 1 or 2,………………, as [b_i] and x_i are integers.

Thus the equation (13.7) becomes,

$$\sum f_{ij}^{+} x_j + \sum f_{ij}^{-} x_j \geq f_{b_i} \qquad \text{….(13.8)}$$

and clearly

$$\sum f_{ij}^{+} x_j \geq \sum f_{ij}^{+} x_j + \sum f_{ij}^{-} x_j \qquad \text{….(13.9)}$$

as x_j's ≥ 0 and $f_{ij}^{-} < 0$.

From equations (13.8) and (13.9), we have

$$\sum f_{ij}^{+} x_j \geq f_{b_i} \qquad \text{….(13.10)}$$

Case-II: In case $f_{b_i} + \{[b_i] - x_i\} < 0$ will imply $[b_i] - x_i = -1, -2, \ldots\ldots\ldots,$ as difference is an integer.

Using this fact from (13.7), we have

$$\sum f_{ij}^{+} x_j + \sum f_{ij}^{-} x_j \leq f_{b_i} - 1 \qquad \text{....(13.11)}$$

So clearly

$$\sum f_{ij}^{-} x_j \leq \sum f_{ij}^{+} x_j + \sum f_{ij}^{-} x_j \qquad \text{....(13.12)}$$

From equations (13.11) and (13.12), we have

$$\sum f_{ij}^{-} x_j \leq f_{b_i} - 1$$

or
$$\frac{f_{b_i}}{f_{b_i} - 1} \sum f_{ij}^{-} x_j > f_{b_i} \quad as \quad 0 < f_{b_i} < 1 \qquad \text{....(13.13)}$$

From equations (13.10) and (13.13), we have

$$\sum f_{ij}^{+} x_{ij} + \frac{f_{b_i}}{f_{b_i} - 1} \sum f_{ij}^{-} x_j \geq f_{b_i}$$

or
$$-f_{b_i} = -\sum f_{ij}^{+} x_j - \frac{f_{b_i}}{f_{b_i} - 1} \sum f_{ij}^{-} x_j + s_{g_i} \qquad \text{....(13.14)}$$

This is Gomory's cut for MILPP.

Note that for MILPP in Gomory's cut, f_{ij}^{+}, f_{ij}^{-} are just the positive and negative coefficients of non-basic variables without being converted into positive fractional values but f_{b_i} is non-negative fractional part of basic variable x_i.

Examples

Example.4. Solve the following MILPP:

Max $Z = 4x_1+6x_2+2x_3$

s.t. $4x_1-4x_2 \leq 5$

$-x_1+6x_2 \leq 5$

$-x_1+x_2+x_3 \leq 5$

$x_1, x_2, x_3 \geq 0$ and x_2 is integer.

Solution The given problem in standard form can be written as

$Max\ Z = 4x_1+6x_2+2x_3+0x_4+0x_5+0x_6$

s.t. $4x_1-4x_2+x_4 = 5$

$-x_1+6x_2+x_5 = 5$

$-x_1+x_2+x_3+x_6 = 5$

$x_1, x_2, x_3, x_4, x_5, x_6 \geq 0$; x_2 is an integer and x_4, x_5, x_6 are the slack variables.

Ignoring the integer requirement, the optimal solution using simplex method is obtained as given in the following simplex table 1:

Cost →				4	6	2	0	0	0
Variable →				x_1	x_2	x_3	x_4	x_5	x_6
Table No.	$\mathbf{C_B}$	Basic Variable	$\mathbf{X_B}$	α^1	α^2	α^3	α^4	α^5	α^6
1	4	x_1	5/2	1	0	0	3/10	1/5	0
	6	x_2	5/4	0	1	0	1/20	1/5	0
	2	x_3	25/4	0	0	1	1/4	0	1
	c_j-Z_j			0	0	0	-2	-2	-2

The non-integer optimal solution is x_1=5/2, x_2=5/4, x_3=25/4 and Max Z=30.

Here x_2=5/4 which is non-integer is required to be integer. Thus the x_2-row (second row in above simplex table) can be written as

$$\frac{5}{4} = 0x_1 + 1x_2 + 0x_3 + \frac{1}{20}x_4 + \frac{1}{5}x_5 + 0x_6$$

The Gomory's cut for MILPP is

$$-f_{b_i} = -\sum f_{ij}^{+} x_j - \frac{f_{b_i}}{f_{b_i}-1}\sum f_{ij}^{-} x_j + s_{g_i}$$

Here $i=2,\quad f_{b_i}=\frac{1}{4},\quad f_{24}^{+}=\frac{1}{20},\quad f_{25}^{+}=\frac{1}{5},\quad f_{26}^{-}=0.$

Thus the Gomory's cut is

$$-\frac{1}{4} = -\frac{1}{20}x_4 - \frac{1}{5}x_5 + s_{g2}$$

Adding this constraint in above table 1, the new simplex table 2 is

Cost →				4	6	2	0	0	0	0
Variable →				x_1	x_2	x_3	x_4	x_5	x_6	s_{g2}
Table No.	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6	α^{g2}
2	4	x_1	5/2	1	0	0	3/10	1/5	0	0
	6	x_2	5/4	0	1	0	1/20	1/5	0	0
	2	x_3	25/4	0	0	1	1/4	0	1	0
	0	← s_{g2}	-1/4	0	0	0	-1/20	-1/5	0	1
	c_j-Z_j			0	0	0	-2	-2↑	-2	0

The solution is infeasible. Restore feasibility using dual simplex method. When s_{g_2} leaves then x_5 enters

$$\left\{ as \min\left(\left| \frac{c_4 - z_4}{\alpha_4^4} \right|, \left| \frac{c_5 - z_5}{\alpha_5^4} \right| \right) = \min\ (40, 10) = 10\ \ for\ x_5\ column \right\}.$$

The next simplex table 3 is

Cost →				4	6	2	0	0	0	0
Variable →				x_1	x_2	x_3	x_4	x_5	x_6	s_{g_2}
Table No.	**C_B**	Basic Variable	**X_B**	α^1	α^2	α^3	α^4	α^5	α^6	α^{g2}
3	4	x_1	9/4	1	0	0	1/4	0	0	1
	6	x_2	1	0	1	0	0	0	0	1
	2	x_3	25/4	0	0	1	1/4	0	1	0
	0	x_5	5/4	0	0	0	1/4	1	0	-5
	c_j-Z_j			0	0	0	-3/2	0	-2	-10

This is the optimal table with x_2 an integer. Thus the optimum solution is x_1=9/4, x_2=1, x_3=25/4 and Max Z=55/2.

13.11 Difference between Gomory's Cutting Plane Method for AILPP and MILPP

In All Integer Linear Programming Problem (AILPP), for Gomory's cut we use fractional parts of all non-basic variables of the optimum simplex table including the basic variable. In Mixed Integer Linear Programming Problem (MILPP), we use only coefficients and not the fractional parts of all non-basic variables. However, we use fractional part of basic variable having most positive fractional part in optimal simplex table. Because of above differences, the Gomory's cut is different for AILPP and MILPP.

13.12 Summary

Integer Linear Programming (ILP) techniques are employed to address optimization problems wherein some or all of the decision variables must assume integer values. Unlike traditional linear programming, which allows variables to be continuous, ILP imposes the constraint that certain variables must be integers. This restriction is particularly relevant in real-world scenarios where solutions need to be whole numbers, such as in resource allocation, scheduling, or network design problems. The goal of ILP is to find the optimal integer values for decision variables that satisfy the problem constraints and maximize or minimize the objective function. In contrast to All Integer Linear Programming Problems (AILPP), where the primary emphasis is on generating cuts to preserve integer feasibility, the approach in Mixed Integer Linear Programming Problems (MILPP) diverges. In MILPP, the initial focus is on attaining an optimal solution within the more flexible framework of non-integer variables. The process involves utilizing methods like the simplex method to find an optimal solution without strict adherence to integer constraints. Subsequently, additional constraints, or cuts, are introduced strategically to gradually transition towards a solution that not only satisfies the integer requirements but also maintains optimality. This distinction underscores the different strategic considerations in addressing MILPP compared to AILPP.

13.13 Terminal Questions

Q.1. Explain the Integer Linear Programming Problem.

Q.2. Write a short note on Integer Linear Programming Problem.

Q.3. Find the optimum all integer solution to the following ILPP:

$$\text{Max } Z = x_1+2x_2$$

s.t. $2x_2 \leq 7$

$x_1+x_2 \leq 7$

$2x_1 \leq 11$

$x_1, x_2 \geq 0$ and integers.

Q.4. Find the optimum all integer solution to the following ILPP:

$$\text{Max } Z = 2x_1+20x_2-10x_3$$

s.t. $2x_1+20x_2+4x_3 \leq 15$

$6x_1+20x_2+4x_3 = 20$

$x_1, x_2, x_3 \geq 0$ and integers.

Q.5. Solve the following MILPP:

$$\text{Max } Z = 2x_1 + x_2$$

$$\text{s.t.} \quad 3x_1 + 2x_2 \leq 5$$

$$x_2 \leq 2$$

$$x_1, x_2 \geq 0 \text{ and } x_1 \text{ is integer.}$$

Answer

3. $x_1=4$, $x_2=3$ and Max Z =10.

4. $x_1=2$, $x_2=0$, $x_3=2$ and Max Z =-16.

5. $x_1=0$, $x_2=2$ and Max Z=2.

UNIT-14: Integer Programming Problem-II

Structure

14.1 Introduction

Branch and Bound is a general algorithmic technique for finding optimal solutions to combinatorial optimization problems. It is commonly used for solving Integer Linear Programming (ILP) problems and other discrete optimization problems. Integer Linear Programming (ILP) techniques are used to solve optimization problems where some or all of the decision variables are required to take on integer values.

Choosing the appropriate technique depends on the characteristics of the specific Integer Linear Programming Problem, including its size, structure, and the nature of the integer constraints. The most effective method may vary from one problem instance to another.

14.2 Objectives

After reading this unit the learner should be able to understand about:

- Branch and Bound Technique to find the solution of an Integer Linear Programming Problem
- Zero-One Integer Linear Programming Problem

- Balas Additive Algorithm
- Solution Procedure of Zero-One Integer Linear Programming Problem

14.3 Branch and Bound Techniques

The Dankin's branch and bound method is used for problems with a finite number of feasible solutions. As the number of variables in a problem increases, the total number of feasible solutions also grows exponentially, making it impractical to examine each one individually. In the Dankin's branch and bound method, the approach involves systematically filtering large subsets of feasible solutions to identify and converge towards the optimal solution.

This method is particularly useful for problems with a large number of variables and constraints, allowing for more efficient exploration of the feasible solution space.

The following steps are used to find the optimal solution of an ILPP using Dankin's branch and bound techniques:

Step-I: Let the given Integer Linear Programming Problem be a maximization problem. If it is a minimization problem, convert it into maximization problem by multiplying the objective function with (-1).

Step-II: Solve the given ILPP using simplex method ignoring the integer requirement. If the solution obtained is integer then it is optimal solution. If it is not an integer solution, then go to step III.

Step-III: Consider the objective function value Z in a maximization problem as the upper bound along with the solution obtained in Step-II and call it node-1.

Now if x_k^* is the non-integer value of x_k in node 1, which is required to be an integer, then introduce two branches (sub problems):

(i) $x_k \leq [x_k^*]$ and

(ii) $x_k \geq [x_k^*]+1$, where $[x_k^*]$ is the greatest integer $\leq x_k^*$.

Step-IV: Solve the above two sub problems (i) and (ii) using sensitivity analysis and dual simplex method getting two node say, node 2 and node 3.

With these nodes associate the value of Z as the upper bound for the purpose of comparison.

Step-IV: Among the terminal nodes, branch off from the node for which the value of Z (upper bound) is largest.

Finally the process will end if the terminal nodes are:

(i) with integer solution

(ii) non-feasible solution or

(iii) Fathomed nodes {Let α be the upper bound with best node (that is with largest Z) and β be the Z-value of some another node such that $\beta \leq \alpha$.

Then this node is not branched off further, as its branching will not give a better Z-value (Such nodes are called fathomed nodes}.

The best out of the nodes of type (i) gives the optimal solution of ILPP.

Examples

Example.1. Solve the following ILPP using Dankin's Branch and Bound Techniques:

$$\text{Max } Z = 5x_1 - 3x_2$$

s.t.

$$-x_1 + x_2 \leq 1$$

$$2x_1 + x_2 \leq 2$$

$$2x_1 \leq 1$$

$x_1, x_2 \geq 0$ and integers.

Solution The given problem in standard form can be written as

$$Max\ Z = 5x_1 - 3x_2 + 0x_3 + 0x_4 + 0x_5$$

s.t.

$$-x_1 + x_2 + x_3 = 1$$

$$2x_1 + x_2 + x_4 = 2$$

$$2x_1 + x_5 = 1$$

$x_1, x_2, x_3, x_4, x_5 \geq 0$ and integers; x_3, x_4, x_5 are the slack variables.

Now omitting the integer requirement in above ILPP, we solve the resulting LPP as below:

Cost →				5	-3	0	0	0	Minimum Ratio
Variable →				x_1	x_2	x_3	x_4	x_5	$\frac{X_{B_i}}{\alpha_i^j}$
Table	$\mathbf{C_B}$	Basic	$\mathbf{X_B}$	α^1	α^2	α^3	α^4	α^5	

No.		Variable							
	0	x_3	1	-1	1	1	0	0	$\frac{2}{2}=1$
	0	x_4	2	2	1	0	1	0	
1	0	$\leftarrow x_5$	1	[2]	0	0	0	1	$\frac{1}{2}\rightarrow$
	c_j-Z_j			5↑	-3	0	0	0	
	0	x_3	3/2	0	1	1	0	1/2	
	0	x_4	1	0	1	0	1	-1	
2	5	x_1	1/2	1	0	0	0	1/2	
	c_j-Z_j			0	-3	0	0	-5/2	

The optimum solution is x_1=1/2, x_2=0 and Max Z =5/2, which is not all integer solution. Consider the objective function value Z in this maximization problem as the upper bound along with the solution obtained x_1=1/2, x_2=0 and Max Z =5/2 and call it node-1.

Here 1/2 is the non-integer value of x_1 at node 1, which is required to be an integer. To have x_1 as an integer, [x_1]=0, so introduce two branches (subproblems):

(i) $x_1 \geq 1$ and

(ii) $x_1 \leq 0$, where 0 is the greatest integer $\leq x_1$.

Now solve the above two subproblems, using sensitivity analysis and dual simplex method getting two nodes say, node 2 and node 3 (Figure 14.1).

(i) For node 2, subproblem becomes:

$$Max\ Z = 5x_1-3x_2+0x_3+0x_4+0x_5$$

s.t. $-x_1+x_2+x_3 = 1$

$$2x_1+x_2+x_4 = 2$$

$$2x_1+x_5= 1$$

$-x_1 + x_6 = -1$

$x_1, x_2, x_3, x_4, x_5, x_6 \geq 0$ and integers; x_3, x_4, x_5, x_6 are the slack variables.

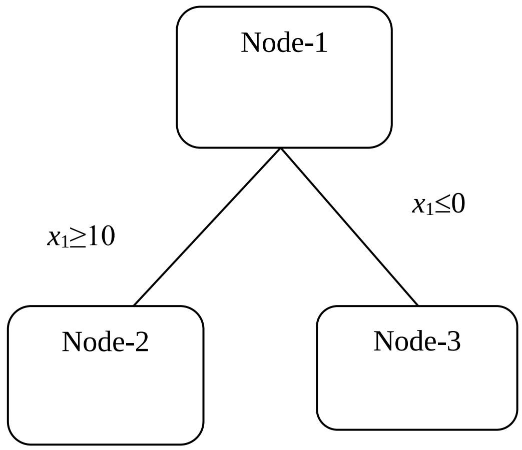


Figure 14.1 Branching at Node 1

Adding this additional constraint the new simplex table is:

Cost →				5	-3	0	0	0	0
Variable →				x_1	x_2	x_3	x_4	x_5	x_6
Table No.	**C_B**	Basic Variable	**X_B**	α^1	α^2	α^3	α^4	α^5	α^6
3	0	x_3	3/2	0	1	1	0	1/2	0
	0	x_4	1	0	1	0	1	-1	0
	5	x_1	1/2	1	0	0	0	1/2	0
	0	x_6	-1	-1	0	0	0	0	1
	$c_j - Z_j$			0	-3	0	0	-5/2	0

Above simplex table is not in the standard format (as α^1 column is not an identity column).

To reduce it in the standard format, adding x_1-row entries in the corresponding x_6-row entries we get the following simplex table:

Cost →				2	3	0	0	0	0
Variable →				x_1	x_2	x_3	x_4	x_5	x_6
Table No.	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6
4	0	x_3	3/2	0	1	1	0	1/2	0
	0	x_4	1	0	1	0	1	-1	0
	5	x_1	1/2	1	0	0	0	1/2	0
	0	x_6	-1/2	0	0	0	0	1/2	1
	c_j-Z_j			0	-3	0	0	-5/2	0

The solution is infeasible. Restore feasibility using dual simplex method, variable x_6 leaves but there is no variable to enter (as there is no negative entry in x_6-row). Thus solution is infeasible and say it node-2.

(i) For node 3, the subproblem becomes

$$Max\ Z = 5x_1 - 3x_2 + 0x_3 + 0x_4 + 0x_5$$

s.t.

$$-x_1 + x_2 + x_3 = 1$$

$$2x_1 + x_2 + x_4 = 2$$

$$2x_1 + x_5 = 1$$

$$x_1 + x_7 = 0$$

$x_1, x_2, x_3, x_4, x_5, x_7 \geq 0$ and integers; x_3, x_4, x_5, x_7 are the slack variables.

Adding this additional constraint in simplex table 2, the new simplex table obtained is:

Cost →				5	-3	0	0	0	0
Variable →				x_1	x_2	x_3	x_4	x_5	x_6
Table No.	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6
	0	x_3	3/2	0	1	1	0	1/2	0
	0	x_4	1	0	1	0	1	-1	0
5	5	x_1	1/2	1	0	0	0	1/2	0
	0	x_7	0	1	0	0	0	0	1
	c_j-Z_j			0	-3	0	0	-5/2	0

Above simplex table is not in the standard format (as α^1 column is not an identity column).. To reduce it in the standard format, subtracting x_1-row entries from the corresponding x_7-row entries we get the following simplex table:

Cost →				5	-3	0	0	0	0
Variable →				x_1	x_2	x_3	x_4	x_5	x_6
Table No.	C_B	Basic Variable	X_B	α^1	α^2	α^3	α^4	α^5	α^6

6	0	x_3	3/2	0	1	1	0	1/2	0
	0	x_4	1	0	1	0	1	-1	0
	5	x_1	1/2	1	0	0	0	1/2	0
	0	x_7	-1/2	0	0	0	0	-1/2	1
	c_j-Z_j			0	-3	0	0	-5/2	0

The solution is infeasible. Restore feasibility using dual simplex method. When x_7 leaves then x_5 enters, to get the next simplex table 7:

Cost →				5	-3	0	0	0	0
Variable →				x_1	x_2	x_3	x_4	x_5	x_6
Table No.	**C_B**	Basic Variable	**X_B**	α^1	α^2	α^3	α^4	α^5	α^6
7	0	x_3	3/2	0	1	1	0	0	1
	0	x_4	1	0	1	0	1	0	-2
	5	x_1	0	1	0	0	0	0	1
	0	x_5	1	0	0	0	0	1	-2
	c_j-Z_j			0	-3	0	0	0	-5

The optimum solution is x_1=0, x_2= 0 and Max Z =0, which is all integer solution at node-3.

Finally the process is end because node 2 has infeasible solution and node 3 has integer solution, which are fathomed nodes. Hence the optimum solution is x_1=0, x_2=0 and Max Z =0 of the given ILPP at node

3.

14.4 Zero-One Integer Linear Programming Problem

The first zero-one algorithm called the additive algorithm was proposed to solve in 1965 by E. Balas nearly seven years after the development of branch and bound techniques. In the beginning it (zero-one ILPP) appeared unrelated to branch and bound technique in the sense that it does not require solving LP problems but the main computational work is simple additions and subtractions. However short after, it becomes evident that zero-one algorithm is a special case of branch and bound algorithm.

14.5 Balas Additive Algorithm

$$\left.\begin{aligned} &MinZ = \sum_{j=1}^{n} C_j x_j \\ &\sum_{j=1}^{n} a_{ij} x_j \le b_i,\ i = 1,\ 2, ...,m \\ &x_j = 0\ or\ 1, C_j > 0,\ \forall j \end{aligned}\right\} \quad .... (14.1)$$

Or

$$\left.\begin{aligned} &\text{Min } Z = \sum_{j=1}^{n} C_j x_j \\ &s_i = b_i - \sum_{j=1}^{n} a_{ij} x_j,\ i = 1,\ 2, ...,m \\ &x_j = 0\ or\ 1, C_j > 0,\ \forall j \\ &\text{and } s_i \text{ is the slack variable added to } i^{\text{th}} \text{ constraint },\ i = 1, 2, ...,m \end{aligned}\right\} \quad .... (14.2)$$

Thus for Balas additive algorithm following conditions must be satisfied:

1. Problem must be of minimization type.

2. All objective function coefficients C_j*'s* must be non-negative.

3. All constraints must be of "≤" type.

Any zero-one LP (where variables assume value o or 1) can be put in the form of additive algorithm.

The following terms are used in Baas additive algorithm to solve a zero-one ILPP:

Branching Nodes

The partial solution (Solution of sub problem) from where branching is done is called a branching node.

Branching Variable

The variable selected to raise it from level 0 to level 1 is called branching variable.

Slack Feasibility

The sum of the non-positive slack values, given that a zero variable x_j is elevated to level 1 is called slack infeasibility of slack variable s_i, denoted as I_j and is given as

$$I_j = \sum_i \min\,[s_i - a_{ij}, 0], \qquad \ldots\ldots (14.3)$$

Where s_i= current value of slack variable i and a_{ij} = constraint coefficient of the j^{th} variable x_j in constraint i.

Criterion of Selection of Branching Variable

The variable x_j for which the amount added to bring a non-positive I_j at zero level is minimum, is selected for branching. If the minimum amount is same for more than one variable then the variable whose objective function coefficient is smaller should be preferred for branching. In case objective function coefficients are also equal than break the tie arbitrarily.

Note: Branching is done in both branch and bound and zero-one ILPP but the method and the selection criteria of branching variable are different for these.

Fathomed Nodes

A node (subproblem solution) is said to be fathomed if one of following happens:

(i) Branching from this cannot give a feasible solution.

(ii) Branching from this cannot yield a better upper bound.

(iii) Branching from this cannot yield a feasible integer solution.

Free Variables

Variable which is free to assume value 0 or 1 is called a free variable.

14.6 Solution Procedure of Zero-One Integer Linear Programming Problem

The following steps are used in solving zero-one ILPP:

Step-I: Bring the problem in the initial form required by additive algorithm i.e., problem must be a minimization problem with all objective function coefficients non-negative and all of "≤" type.

Step-II: Express the slack variable s_i (s_i is the slack variable added to i^{th} constraint) in terms of remaining variables by transposing every term except s_i on the right hand side of equality.

Step-III: Find an initial all zero binary solution (slack solution) and call it node 1 attaching with it the Z-value (objective function value) as upper bound. If all the slacks were non-negative, we would conclude that the all zero binary solution is optimum. If not go to step IV.

Step-IV: If some of the slack variable(s) are infeasible (negative), then using slack infeasibility test identify the branching variable. Let it be x_j. The two branches with x_j as the branching variables are x_j =1 and x_j=0 creating nodes 2 and 3 respectively.

Proceeding this way we arrive at a situation where all terminal nodes are fathomed. Out of these we pick up the best one giving the optimal solution of given problem

Examples

Example.2. Convert the following 0-1 problem to satisfy the starting requirements of the additive algorithm:

$$\text{Max } Z = 2x_1 - 7x_2$$

$$\text{s.t.} \quad x_1 + x_2 = 5$$

$$4x_1 + 6x_2 \geq 4$$

$$x_1 \geq 4$$

$$x_2 \leq 5$$

$$x_i\text{'s} = 0 \text{ or } 1,\ i = 1, 2.$$

Solution First convert the given problem to minimization with all '≤' constraints as follows:

(i) Multiplying the objective function row (Z-row) by (-1) to get minimize $Z' = -2x_1 + 7x_2$.

(ii) Convert the first constraint equation into two constraints of '≤' type to obtain

$$x_1 + x_2 \leq 5$$

and $-x_1 - x_2 \leq -5$.

(iii) Multiply the second and third constraints by (-1) to obtain constraints $-4x_1 - 6x_2 \leq -4$ and $-x_1 \leq -4$ respectively.

Now using above computations the problem is written after adding slacks as

$$Min \quad Z' = -Z = -2x_1 + 7x_2$$

s.t. $x_1+x_2+s_1 = 5$

$-x_1-x_2+s_2 = -5$

$-4x_1-6x_2+s_3 = -4$

$-x_1+s_4 = -4$

$x_2+s_5 = 5$

s_i's ≥ 0, $i = 1, 2, ..., 5$ are slack variables and $x_1, x_2 = 0$ or 1.

To ensure that the coefficients in objective function are non-negative, substitute $x_1=1-y_1$ (as x_1 is with negative coefficient) and $x_2=y_2$ (as x_2 is with positive coefficient). Change the left hand side of the constraints accordingly, to get the given 0-1 problem fulfilling the requirements of the additive algorithm. The required form for additive algorithm is

$$Min \quad Z' = -2(1-y_1)+7y_2$$

s.t. $(1-y_1)+y_2+s_1 = 5$

$-(1-y_1) -y_2+s_2 = -5$

$-4(1-y_1)-6y_2+s_3 = -4$

$-(1-y_1)+s_4 = -4$

$y_2+s_5 = 5$

s_i's ≥ 0, $i = 1, 2, ..., 5$ are slack variables and $y_1, y_2 = 0$ or 1.

or

$$Min \quad Z' = 2y_1+7y_2-2$$

s.t. $-y_1+y_2+s_1 = 4$

$y_1 -y_2+s_2 = -4$

$4y_1-6y_2+s_3 = 0$

$y_1+s_4 = -3$

$y_2+s_5 = 5$

s_i's ≥ 0, $i = 1, 2, ..., 5$ are slack variables and $y_1, y_2 = 0$ or 1.

Example.3. Solve the following 0-1 problem:

Min $Z = 2x_1+3x_2$

s.t. $x_1+x_2 \leq 2$

$-x_1+x_2 \leq 1$

$x_1, x_2 = 0$ or 1.

Solution The problem can be put in initial form required by the additive algorithm using following operations:

(i) Add the two slack variables s_1 and s_2 to convert the two constraints into equations, we get . $x_1+x_2+s_1=2$ and $-x_1+x_2+s_2 = 1$.

$$Min\ Z = 2x_1+3x_2$$

s.t. $s_1 = 2-x_1-x_2$

$s_2 = 1+x_1-x_2$

As we seek the minimum of the objective function, a logical starting solution is when all binary variables are at zero level. In this case slacks will act as basic variables and the initial all zero binary solution is $s_1=2$, $s_2=1$, Z=0. Here all the slacks are non-negative i.e., the all zero binary solution is the optimum solution. Hence the optimum solution is $x_1= x_2=0$ and Z=0.

14.7 Summary

Branch and Bound is particularly effective for solving discrete optimization problems where the solution space is large and needs to be systematically explored. It efficiently narrows down the search space by bounding and pruning, making it suitable for problems like ILP where integer solutions are sought. Branch and bound technique used only AILPP. To solve a zero-one ILPP using Balas additive algorithm, the problem must be a minimization problem with all objective function coefficients non-negative and all of "≤" type.

14.8 Terminal Questions

Q.1. Write a short note on Branch and Bound techniques.

Q.2. What do you mean by Zero-one Integer Linear programming problem

Q.3. Solve the following ILPP using Dankin's Branch and Bound Techniques:

$$\text{Max } Z = 2x_1+3x_2$$

s.t. $-x_1+x_2 \leq 1$

$2x_1+x_2 \leq 2$

$x_1, x_2 \geq 0$ and integers.

Q.4. Solve the following 0-1 problem:

$$\text{Max } W = 3x_1+2x_2-5x_3-2x_4+3x_5$$

s.t. $x_1+x_2+x_3+2x_4+x_5 \leq 4$

$7x_1+3x_3-4x_4+3x_5 \leq 8$

$11x_1-6x_2+3x_4-3x_5 \geq 3$

x_i's = 0 or 1, i = 1, 2, ..., 5.

Answer

3. is x_1=0, x_2= 1 and Max Z =3.

4. x_1=1, x_2=1, x_3=x_4= x_5=0 and W=5.