CORE - 2
OPERATIONS RESEARCH

SYLLABUS

UNIT – I

Importance-The History of OR-Definition-Features-Scope of Operations Research –Linear Programming: Introduction-Advantages of using LP-Application areas of LP- Formation of mathematical modelling, Graphical method, the Simplex Method; Justification, interpretation of Significance of All Elements In the Simplex Tableau, Artificial variable techniques: Big M method.

UNIT II

Transportation, Assignment Models: Definition and application of the transportation model, methods for finding initial solution-tests for optimality-variations in transportation problem, the Assignment Model, Travelling Salesman Problem.

UNIT – III

Dynamic Programming – Applications of D.P. (Capital Budgeting, Production Planning, Solving Linear Programming Problem) – Integer Programming – Branch and Bound Method.

UNIT – IV

Game Theory: Introduction – Two Person Zero-Sum Games, Pure Strategies, Games with Saddle Point, Mixed strategies, Rules of Dominance, Solution Methods of Games without Saddle point – Algebraic, matrix and arithmetic methods. Simulation – Simulation Inventory and Waiting Lines.

UNIT – V

P.E.R.T. & C.P.M. and Replacement Model: Drawing networks – identifying critical path – probability of completing the project within given time- project crashing – optimum cost and optimum duration. Replacement models comprising single replacement and group replacement.

UNIT – I
SIGNIFICANCE OF OPERATIONS RESEARCH & LINEAR PROGRAMMING

Learning Objectives

Importance-The History of OR-Definition-Features-Scope of Operations Research –Linear Programming: Introduction-Advantages of using LP-Application areas of LP- Formation of mathematical modelling, Graphical method, the Simplex Method; Justification, interpretation of Significance of All Elements In the Simplex Tableau, Artificial variable techniques: Big M method.

HISTORY OF OPERATIONS RESEARCH

Historically, the term Operations Research originated during Second World War when U.S.A. and Great Britain's Armed Forces sought the assistance of Scientists to solve complex and very difficult strategical and tactical problems of warfare, like making mines harmless or increasing the efficiency of antisubmarine aerial warfare, etc.
Operations research employs mathematical logic to complex problems requiring managerial decisions.
Operations research aids, in solving diverse business problems and in planning and investigation of major operational decisions.

The term operations Research was first coined in 1940 by McClosky and Trefthen in a small town, Bowdsey, of the United Kingdom. This new science came into existence in military context. During world war II, military management called on scientists from various disciplines and organized them into teams to assist in solving strategic and tactical problems (ie) to discuss, evolve and suggest ways and means to improve the execution of various military projects. By their joint efforts, experience and deliberations, they suggested certain approaches that showed remarkable progress. This new approach to systematic and scientific study of the operations of the system was called the Operations Research or Operational Research (abbreviated as O.R). During the year 1950, O.R achieved recognition as a subject worthy of academic study in the Universities. Since then, the subject has been gaining more and more importance for students of Economics, Management, Public Administration, Behavioral Sciences, Social work, Mathematics, Commerce and Engineerin
Operations Research Society of America was formed in 1950 and in 1957 the International Federation of O.R Societies was established. In several Countries, International Scientific Journals in O.R began to appear in different languages. The primary journals are Operations Research, Transportation Science, Management Sciences, Operational Research Quarterly, Journal of the Canadian Operational Research Society, Mathematics of Operational Research, International journal of Game Theory etc.

OPERATIONAL RESEARCH IN INDIA

In India, Operational Research came into existence in 1949 with the opening of an Operational Research Unit at the Regional Research Laboratory at Hyderabad. In 1953, an Operational Research Unit was established in the Indian Statistical Institute, Calcutta for the application of Operational Research methods in national planning & survey . Operational Research Society of India was formed in 1957. It became a member of the International Federation of Operational Research Societies in 1959. The first Conference of Operational Research Society of India was held in Delhi in 1959. Operational Research Society of India started a journal

"Opsearch" in 1963. Other journals which deal with Operational Research are : Journal of the National Productivity Council, Materials Management journal of India and the Defence Science journal.

Definition

Because of the wide scope of applications of Operational Research, giving a precise definition is difficult. However, a few definitions of Operational Research are as under :

"Operational Research is the application of scientific methods, techniques and tools to problems involving the Operations of a system so as to provide those in control of the system with optimum solutions to the problem".

- C.W.Churchman, R.L.Ackoff & E.L.Arnoff

"Operational Research is the art of giving bad answers to problems which otherwise have worse answers".

- T.L.Saaty

MODELS IN OPERATIONAL RESEARCH

A model in Operational Research is a simplified representation of an operation or a process in which only the basic aspects or the most important features of a typical problem under investigation are considered.

Types of Models

There are many ways to classify models and therefore, the decision-maker must identify which type of model best suits the decision problem.

Physical Models

These models provide a physical appearance of the real object under study either reduced in size or scaled up physical models are useful only in design problems because they are easy to observe, build and describe.

1. **Iconic models:**Iconic model retain some of the physical properties and characteristics of the system they represent.
2. **Analogue models :** The models represent a system by the set of properties different from that of the original system and does not resemble physically.

Symbolic Models

These models use letters, numbers and other symbols to represent the properties of the system.

1. **Verbal Models :** These models describes a situation in written or spoken language. eg:- Written Sentences, books etc.,
2. **Mathematical Models:** These models involve the use of mathematical Symbols, letters, numbers and mathematical operators (+, -, ÷, ×) to represent relationship among various variables of the systems to describe its properties or behaviour.

Descriptive Models

These models simply describe some aspects of a situation, based on observation, survey, questionnaire results or other available data of a situation and do not predict or recommend.

Eg:- Plant layout diagram

Predictive Models

These models are used to predict the outcomes due to a given set of alternatives for problem. These models do not have an objective function as a part of the model to evaluate decision alternatives.

Optimization Models

These models provide the ‘best’ or ‘optimal’ solution to problems subject to certain limitations of the use of resources.

Static Models

Static models present a system at some specified time and do not account for changes over time.

Dynamic Models

In a dynamic model, time is considered as one of the variables and admit the impact of changes generated by time in the selection of the optimal courses of action.

Deterministic Models

If all the parameters, constants and functional relationships are assumed to be known with certainty when the decision is made, then the model is said to be deterministic. For a specific set of input values, there is a uniquely determined output which represents the solution of the model under conditions of certainty.

Eg:- Linear Programming Model.

Probabilistic (Stochastic Models)

Models in which atleast one parameter or decision variable is a random variable are called probabilistic (or Stochastic) models. Since atleast one decision variable is random, therefore, an independent variable which is the function of dependent variable(s) will also be random. This means consequences or payoff due to certain changes in the independent variable cannot be predicted with certainty. However, it is possible to predict a pattern of values of both the variable by their probability distribution.

Eg:- Insurance against risk of fire, accidents, sickness etc.

Analytical Models

These models have a specific mathematical structure and thus can be solved by known analytical or mathematical techniques. Any optimization model (which requires maximization or minimization of an objective function) is an analytical model.

Simulation Models

These models also have a mathematical structure but are not solved by applying mathematical structure but are not solved by applying mathematical techniques to get a solution. Instead, a simulation model is essentially a computer assisted experimentation on a mathematical structure of a real-life problem in order to describe and evaluate its behavior under certain assumptions over a period of time.

CHARACTERISTICS OF OPERATIONS RESEARCH

There are three primary characteristics of all operations research efforts:

1. Optimization- The purpose of operations research is to achieve the best performance under the given circumstances. Optimization also involves comparing and narrowing down potential options.
2. Simulation- This involves building models or replications to try out and test solutions before applying them.
3. Probability and statistics- This includes using mathematical algorithms and data to uncover helpful insights and risks, make reliable predictions and test possible solutions.

IMPORTANCE OF OPERATIONS RESEARCH

The field of operations research provides a more powerful approach to decision making than ordinary software and data analytics tools. Employing operations research professionals can help companies achieve more complete datasets, consider all available options, predict all possible outcomes, and estimate risk. Additionally, operations research can be tailored to specific business processes or use cases to determine which techniques are most appropriate to solve the problem.

SCOPE OF OPERATION RESEARCH

Operations research can be applied to a variety of use cases, including:

- Scheduling and time management.
- Urban and agricultural planning.
- Enterprise resource planning (ERP) and supply chain management (SCM).
- Inventory management.
- Network optimization and engineering.
- Packet routing optimization.
- Risk management.

Methodology of Operations Research

It is essential to follow some steps that everybody agrees as being helpful in planning, organizing, directing and controlling Operations Research activities within an organization. The steps are listed below :

1. **Formulation of the problem**: It involves analysis of the physical system, setting-up of objectives, determination of restriction constraints against which decision should be adopted, alternative courses of action and measurement of effectiveness.

2. **Construction of a Mathematical model:** After formulation of the problem, the next step is to express all the relevant variables (activities) of the problem into a mathematical model. A generalized mathematical model might take the form :

$$E = f(x_i, y_j)$$

Where f represents a system of mathematical relationships between the measures of effectiveness of the objective and the variables, both controllable and uncontrollable (y_j).

3. **Deriving the solution from the model:** Once the mathematical model is formulated, the next step is to determine the values of decision variables that optimize the given objective function. This deals with the mathematical calculations for obtaining the solution to the model.

4. **Validity of the model:** The model should be validated to measure its accuracy. A model is valid or accurate if (a) it contains al the objectives, constraints, and decision variables relevant to the problem,(b) the objectives, constraints, and decision variables included in the model are all relevant to, or actually part of the problem, and (c) the functional relationships are valid.

5. **Establishing control over the solution :** After testing the model and its solution, the next step of the study is to establish control over the solution, by proper feedback of the information on variables which deviated significantly, the solution goes out of control. In such situation the model may accordingly be modified.

6. **Implementation of the final results :** Finally, the tested results of the model are implemented to work. This would basically involve a careful explanation of the solution to be adopted and its relationship with the operating realities. This stage of Operations Research investigation is executed primarily through the cooperation of both the operations Research experts and those who are responsible for managing and operating the system.

Applications of Operations Research

Operations Research is mainly concerned with the techniques of applying scientific knowledge, besides the development of science. It provides an understanding which gives the expert/manager new insights and capabilities to determine better solutions in his decision-making problems, with great speed, competence, and confidence. In recent years, Operations Research has successfully entered many different areas of research in Defence, Government, Service Organizations, and Industry. We briefly describe some applications of Operations Research in the functional areas of management:

Finance, Budgeting, and Investment

1. Cash flow analysis, long range capital requirements, dividend policies, investment portfolios.
2. Credit policies, credit risks and delinquent account procedures.
3. Claim and complain procedure.

Marketing

1. Product selection, timing, competitive actions.
2. Advertising media with respect to cost and time.
3. Number of salesmen, frequency of calling of account etc.
4. Effectiveness of market research.

Physical Distribution

1. Location and size of warehouses, distribution centres, retail outlets etc.
2. Distribution policy.

Purchasing, Procurement and Exploration

1. Rules for buying.
2. Determining the quantity and timing of purchase.
3. Bidding policies and vendor analysis.
4. Equipment replacement policies

Personnel

1. Forecasting the manpower requirement, Recruitment policies and assignment of jobs.
2. Selection of suitable personnel with due consideration for age and skills, etc.
3. Determination of optimum number of persons for each service centre.

Production

1. Scheduling and sequencing the production run by proper allocation of machines.
2. Calculating the optimum product mix.
3. Selection, location and design of the sites for the production plant.

Research and Development

1. Reliability and evaluation of alternative designs.
2. Control of developed projects.
3. Co-ordination of multiple research projects.
4. Determination of time and cost requirements.

Besides the above mentioned applications of Operations Research in the context of modern management, its use has now extended to a wide range of problems, such as the problems of communication and information, socio-economic fields and national planning.

Uses and Limitations of Operations Research

Uses

1. **Optimum use of production factors.** Linear programming techniques indicate how a manager can most effectively employ his production factors by more efficiently selecting and distributing these elements.
2. **Improved quality of decision**. The computation table gives a clear picture of the happenings within the basic restrictions and the possibilities of compound behavior of

the elements involved in the problem. The effect on the profitability due to changes in the production pattern will be clearly indicated in the table, e.g., simplex table.

3. ***Preparation of future managers.*** These methods substitute a means for improving the knowledge and skill of young managers.
4. **Modification of mathematical solution.** Operations Research presents a possible practical solution when one exists, but it is always a responsibility of the manager to accept or modify the solution before its use. The effect of these modifications may be evaluated from the computational steps and tables.
5. ***Alternative solutions.*** Operations Research techniques will suggest all the alternative solutions available for the same profit so that the management may decide on the basis of its strategies.

Limitations of Operations Research

Operations Research has certain limitations. However, these limitations are mostly related to the time and money factors involved in its applications rather than its practical utility. These limitations are as follows :

a) Magnitude of computation. Operations Research tries to find out the optimal solution taking all the factors into account. In the modern society, these factors are numerous and expressing them in quantity and establishing relationship among these, requires huge calculations. All these calculations cannot be handled manually and require electronic computers which bear very heavy cost. Thus, the use of Operations Research is limited only to very large organizations.

b) Absence of quantification. Operations Research provides solution only when all the elements related to a problem can be quantified. The tangible factors such as price, product, etc., can be expressed in terms of quantity, but intangible factors such as human relations etc, cannot be quantified. Thus, these intangible elements of the problem are excluded from the study, though these might be equally or more important than quantifiable intangible factors as far as possible.

c) Distance between managers and Operations Research. Operations Research, being specialists' job, requires a mathematician or a statistician, who might not be aware of the business problems. Similarly, a manager may fail to understand the complex working of Operations Research. Thus, there is a gap between one who provides the solution and one who uses the solution.

LINEAR PROGRAMMING INTRODUCTION

In a decision-making embroilment, model formulation is important because it represents the essence of business decision problem. The term formulation is used to mean the process of converting the verbal description and numerical data into mathematical expressions which represents the relevant relationship among decision factors, objectives and restrictions on the use of resources. Linear Programming (LP) is a particular type of technique used for economic allocation of 'scarce' or 'limited' resources, such as labour, material, machine, time, warehouse space, capital, energy, etc. to several competing activities, such as products, services, jobs, new equipment, projects, etc. on the basis of a given criterion of optimally. The phrase scarce resources mean resources that are not in unlimited in availability during the planning period. The criterion of optimality generally is either performance, return on investment, profit, cost, utility, time, distance, etc.

George B Dantzing while working with US Air Force during World War II, developed this technique, primarily for solving military logistics problems. But now, it is being used extensively in all functional areas of management, hospitals, airlines, agriculture, military

operations, oil refining, education, energy planning, pollution control, transportation planning and scheduling, research and development, etc. Even though these applications are diverse, all I.P models consist of certain common properties and assumptions. Before applying linear programming to a real-life decision problem, the decision-maker must be aware of all these properties and assumptions.
The word linear refers to linear relationship among variables in a model. Thus, a given change in one variable will always cause a resulting proportional change in another variable. For example, doubling the investment on a certain project will exactly double the rate of the return.
The word programming refers to modelling and solving a problem mathematically that involves the economic allocation of limited resources by choosing a particular course of action or strategy among various alternative strategies to achieve the desired objective.

WHAT IS LINEAR PROGRAMMING?

Now, what is linear programming? Linear programming is a simple technique where we **depict** complex relationships through linear functions and then find the optimum points. The important word in the previous sentence is depicted. The real relationships might be much more complex – but we can simplify them to linear relationships.
Applications of linear programming are everywhere around you. You use linear programming at personal and professional fronts. You are using linear programming when you are driving from home to work and want to take the shortest route. Or when you have a project delivery you make strategies to make your team work efficiently for on-time delivery.

General Linear Programming Problem

The linear programming involving more than two variables may be expressed as follows :
Maximize (or) Minimize $Z = c_1x_1 + c_2x_2 + c_3x_3 + ... c_nx_n$

subject to the constraints

$a_{11}x_1 + a_{12}x_2 + ... + a_{1n}\, x_n$ £ or = or ³ b_1
$a_{21}x_1 + a_{22}x_2 + ... + a_{2n}\, x_n$ £ or = or ³ b_2
$a_{31}x_1 + a_{32}x_2 + ... + a_{3n}\, x_n$ £ or = or ³ b_3
....
$a_{m1}x_1 + a_{m2}\, x_2 + ... + a_{mn}\, x_n$ £ or = or ³ b_m

and the non-negativity restrictions

$x_1, x_2, x_3, ... x_n$ ³ 0.

Note : Some of the constraints may be equalities, some others may be inequalities of (£) type or (³) type or all of them are of same type.
Solution: A set of values x_1, x_2 ...x_n which satisfies the constraints of the LPP is called its solution.
Feasible solution: Any solution to a LPP which satisfies the non-negativity restrictions of the LPP is called its feasible solution.
Optimum Solution or Optimal Solution: Any feasible solution which optimizes (maximizes or minimizes) the objective function of the LPP is called its optimum solution or optimal solution.
Slack Variables: If the constraints of a general LPP be

$$\sum_{j=1}^{n} a_{ij}\ x_j \le b_i \qquad (i = 1, 2, 3, ...k)$$

...........(1)

then the non-negative variables s_i which are introduced to convert the inequalities (1) to the equalities are called slack variables. The value of these variables can be interpreted as the amount of unused resource.

$$\sum_{j=1}^{n} a_{ij}\ x_j + s_i = b_i \qquad (i = 1, 2, 3, \ldots k)$$

Surplus Variables: If the constraints of a general LPP be

$$\sum_{j=1}^{n} a_{ij}\ x_j \geq b_i \qquad (i = k+1, k+2, \ldots)$$

............(2)

then the non-negative variables s_i which are introduced to convert the inequalities (1) to the equalities

$$\sum_{j=1}^{n} a_{ij}\ x_j - s_i = b_i \qquad (i = k+1, k+2, \ldots)$$

are called surplus variables. The value of these variables can be interpreted as the amount over and above the required level.

CANONICAL AND STANDARD FORMS OF LPP :

After the formulation of LPP, the next step is to obtain its solution. But before any method is used to find its solution, the problem must be presented in a suitable from. Two forms are dealt with here, the canonical form and the standard form.

The canonical form : The general linear programming problem can always be expressed in the following form :

Maximize $Z = c_1 x_1 + c_2 x_2 + c_3 x_3 + \ldots c_n x_n$

subject to the constraints

$a_{11}x_1 + a_{12}x_2 + \ldots. + a_{1n} x_n \leq b_1$

$a_{21}x_1 + a_{22}x_2 + \ldots. + a_{2n} x_n \leq b_2$

....

$a_{m1}x_1 + a_{m2} x_2 + \ldots + a_{mn} x_n \leq b_m$

and the non-negativity restrictions

$x_1, x_2, x_3, \ldots x_n \geq 0.$

This form of LPP is called **the canonical form of the** LPP.

In matrix notation the canonical form of LPP can be expressed as :

Maximize Z = CX (objective function)

Subject to AX $\leq$ b (constraints)

and X $\geq$ 0 (non-negativity restrictions)

where C = $(c_1\ c_2 \ldots c_n)$,

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots a_{1n} \\ a_{12} & a_{22} & \ldots a_{2n} \\ \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots\ldots\ldots a_{mn} \end{bmatrix}, X = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_m \end{pmatrix}$$

3.2.2 Characteristics of the Canonical form :

(i) The objective function is of maximization type.

Min f(x) = - Max {-f(x) } (or)

Min Z = - Max (-Z)

(ii) All constraints are of (≤) type, except for the non-negative restrictions.

An inequality of "≥" type can be changed to an inequality of the type "≤" type by multiplying both sides of the inequality y -1.

For example, the linear constraint

$a_{1l}x_1 + a_{12}x_2 + ... + a_{1n}x_n \geq b_i$

is equivalent to

$-a_{il}x_1 - a_{i2}x_2 - ... - a_{in}x_n \leq - b_i$

An equation may be replaced by two weak inequalities in opposite directions.
For example

$a_{il}x_1 + a_{i2}x_2 + ... + a_{in}x_n = b_i$

is equivalent to

$a_{il}x_1 + a_{i2}x_2 + ... + a_{in}x_n \geq b_i$

and $a_{il}x_1 + a_{i2}x_2 + ... + a_{in}x_n \leq b_i$

(iii) All variables are non-negative.

A variable which is unrestricted in sign is equivalent to the difference between two non-negative variables. Thus if x_j is unrestricted in sign, it can be replaced by $(x_j^l - x_j^{ll})$, where x_j^l and x_j^{ll} are both non-negative,

i.e., $x_j = x_j^l - x_j^{ll}$, where $x_j^l \geq 0$ and $x_j^{ll} \geq 0$

The Standard From :

The general linear programming problem in the form

Maximize or Minimize

$Z = c_1x_1 + c_2x_2 + \ldots\ldots + c_nx_n$

Subject to the constraints

$a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1n}x_n = b_1$

$a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2n}x_n = b_2$

$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

$a_{ml}x_1 + a_{m2}x_2 + \ldots\ldots + a_{mn}x_n = b_m$

and $x_1, x_2, \ldots\ldots x_n \geq 0$ is known as standard form

In matrix notation the standard form of LPP can be expressed as :

Maximize or Minimize Z = CX (objective function)

Subject to constraints AX = b and $X \geq 0$

Where, $c = (c_1, c_2, \ldots, c_n)$,

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots a_{1n} \\ a_{12} & a_{22} & \ldots a_{2n} \\ \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} \ldots\ldots & \ldots a_{mn} \end{bmatrix}, X = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_m \end{pmatrix}$$

Characteristics of the standard form :

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

The inequalities can be changed into equation by introducing a non-negative variable on the left hand side of such constraint. It is to be added (slack variable) if the constraint is of " ≤ " type and subtracted (surplus variable) if the constraint is of " ≥ " type.

Basic Solution.: Given a system of m simultaneous linear equations with n variables (m < n).

$Ax=b, x^T \varepsilon R^n$

where A is an m x n matrix of rank m. Let B be any m x m sub matrix, formed by m linearly independent columns of A. Then a solution obtained by setting n-m variables not associated with the columns of B, equal to zero, and solving the resulting system, is called a basic solution to the given system of equations.

The m variables, which may be all different from zero, are called basic variables. The m x m non-singular sub matrix B called a basis matrix with the columns of B as basis vectors. The (n-m) variables which are put to zero are called as non-basic variables.

Example:

Obtain all the basic solutions to the following system of linear equation:

$x_1 + 2x_2 + x_3 = 4$

$2x_1 + x_2 + 5x_3 = 5$

Solution:

The given system of equations can be written in the matrix form

Ax=b

where,

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 5 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

since rank of A is 2, the maximum number of linearly independent columns of A is 2. Thus we can take any of the following, 2x 2 sub-matrices as basis matrix B:

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 2 & 5 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix}$$

The variables not associated with the columns of B are x_3, x_2 and x_1 respectively, in the three different cases.

Let us first take $B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

A basic solution to the given system is now obtained by setting $x_3 = 0$, and solving the system

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

Thus a basic (non-basic) solution to the given system is

(Basic) $x_1 = 2$, $x_2 = 1$; (Non-basic) $x_3 = 0$,

(Basic) $x_1 = 5$, $x_3 = -1$; (Non-basic) $x_2 = 0$,

(Basic) $x_2 = 5/3$, $x_3 = 2/3$; (Non-basic) $x_1 = 0$.

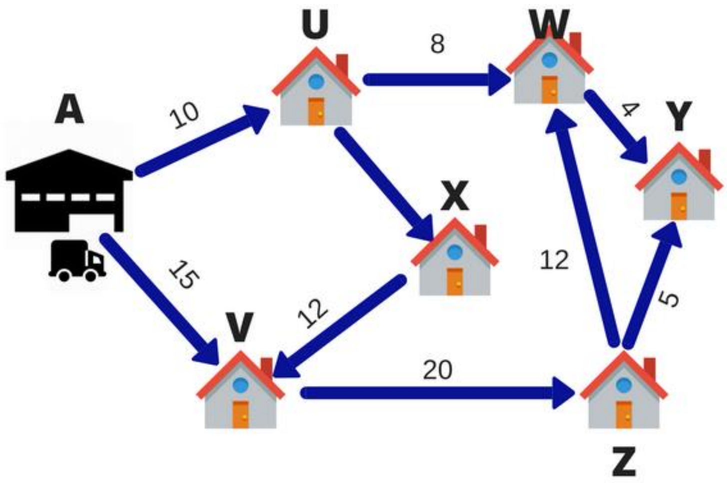
We observe that all the above three basic solutions are non-degenerate solution.

Degenerate Basic Solution: A basic solution is said to be a degenerate basic solution if one or more of the basic variables are zero.

Basic Feasible Solution: A feasible solution to a LPP., which is also a basic solution to the problem is called a basic feasible solution to the LPP.

Example of a linear programming problem

Let's say a FedEx delivery man has 6 packages to deliver in a day. The warehouse is located at point A. The 6 delivery destinations are given by U, V, W, X, Y, and Z. The numbers on the lines indicate the distance between the cities. To save on fuel and time the delivery person wants to take the shortest route.



So, the delivery person will calculate different routes for going to all the 6 destinations and then come up with the shortest route. This technique of choosing the shortest route is called linear programming.

In this case, the objective of the delivery person is to deliver the parcel on time at all 6 destinations. The process of choosing the best route is called Operation Research. Operation research is an approach to decision-making, which involves a set of methods to operate a system. In the above example, my system was the Delivery model.

Linear programming is used for obtaining the most optimal solution for a problem with given constraints. In linear programming, we formulate our real-life problem into a mathematical model. It involves an objective function, linear inequalities with subject to constraints.

Is the linear representation of the 6 points above representative of the real-world? Yes and No. It is an oversimplification as the real route would not be a straight line. It would likely have multiple turns, U-turns, signals and traffic jams. But with a simple assumption, we have reduced the complexity of the problem drastically and are creating a solution that should work in most scenarios.

Formulating a problem – Let's manufacture some chocolates

Example: Consider a chocolate manufacturing company that produces only two types of chocolate – A and B. Both the chocolates require Milk and Choco only. To manufacture each unit of A and B, the following quantities are required:

- Each unit of A requires 1 unit of Milk and 3 units of Choco
- Each unit of B requires 1 unit of Milk and 2 units of Choco

The company kitchen has a total of 5 units of Milk and 12 units of Choco. On each sale, the company makes a profit of

- Rs 6 per unit A sold
- Rs 5 per unit B sold.

Now, the company wishes to maximize its profit. How many units of A and B should it produce respectively?

Solution: The first thing I'm gonna do is represent the problem in a tabular form for better understanding.

	Milk	Choco	Profit per unit
A	1	3	Rs 6
B	1	2	Rs 5
Total	5	12	

Let the total number of units produced by A be = X
Let the total number of units produced by B be = Y
Now, the total profit is represented by Z
The total profit the company makes is given by the total number of units of A and B produced multiplied by its per-unit profit of Rs 6 and Rs 5 respectively.

Profit: Max $Z = 6X+5Y$

which means we have to maximize Z.
The company will try to produce as many units of A and B to maximize the profit. But the resources Milk and Choco are available in a limited amount.
As per the above table, each unit of A and B requires 1 unit of Milk. The total amount of Milk available is 5 units. To represent this mathematically,

$$X+Y \leq 5$$

Also, each unit of A and B requires 3 units & 2 units of Choco respectively. The total amount of Choco available is 12 units. To represent this mathematically,

$$3X+2Y \leq 12$$

Also, the values for units of A can only be integers.
So we have two more constraints, $X \geq 0$ **&** $Y \geq 0$
For the company to make maximum profit, the above inequalities have to be satisfied.

This is called formulating a real-world problem into a mathematical model.
Common terminologies used in Linear Programming
Let us define some terminologies used in Linear Programming using the above example.

- **Decision Variables:** The decision variables are the variables that will decide my output. They represent my ultimate solution. To solve any problem, we first need to identify the decision variables. For the above example, the total number of units for A and B denoted by X & Y respectively are my decision variables.
- **Objective Function:** It is defined as the objective of making decisions. In the above example, the company wishes to increase the total profit represented by Z. So, profit is my objective function.
- **Constraints:** The constraints are the restrictions or limitations on the decision variables. They usually limit the value of the decision variables. In the above example, the limit on the availability of resources Milk and Choco are my constraints.
- **Non-negativity restriction:** For all linear programs, the decision variables should always take non-negative values. This means the values for decision variables should be greater than or equal to 0.

The process to formulate a Linear Programming problem
Let us look at the steps of defining a Linear Programming problem generically:

1. Identify the decision variables
2. Write the objective function
3. Mention the constraints
4. Explicitly state the non-negativity restriction

For a problem to be a linear programming problem, the decision variables, objective function and constraints all have to be linear functions.
If all the three conditions are satisfied, it is called a **Linear Programming Problem**.

Mathematical Formulation of The Problem
In fact, the most difficult problem in the application of management science is the formulation of a model. Therefore, it is important to consider model formulation before launching into the details of linear programming solution. Model formulation is the process of transforming a real word decision problem into an operations research model. In the sections that follow, we give

several Lilliputian examples so that you can acquire some experience of formulating a model. All the examples that we provide in the following sections are of static models, because they deal with decisions that occur only within a single time period

ALGORITHM:

The procedure for mathematical formulation of linear programming problem consists of the following major steps:

Step1: Write down the decision variables of the problem.

Step2: Formulate the objective function to be optimized (maximized or minimized) as a linear function of the decision variables.

Step3: Formulate the other conditions of the problem such as resource limitations, market constraints ,inter-relation between variables etc. as linear equations or in equations in terms of the decision variables.

Step4: Add the 'Non-negativity' constraint from the consideration that negative values of the decision variables do not have any valid physical interpretation

The objective function, the set of constraints and the non-negative constraint together form a Linear Programming Problem.

SAMPLE PROBLEMS

Ex1:

(Production allocation problem). A manufacturer produces two types of models M_1 and M_2.Each M_1 model requires 4 hours of grinding and 2 hours of polishing; whereas each M_2 model requires 2 hours of grinding and 5 hours of polishing. The manufacturer has 2 grinders and 3 polishers. Each grinder works for 40 hours a week and each polisher works for 60 hours a week. Profit on an M_1 model is Rs 3.00 and on an M_2 model is Rs. 4.00. Whatever is produced in a week is sold in the market. How should the manufacturer allocate his production a week is sold in the market. How should the manufacturer allocate his production capacity to the two types of models so that he may make the maximum profit in a week ?

Solution:

Here is a real situation from an industry. The manufacturer can, if he so chooses, produce only model M_2 , but then his grinders would be idle and his profits may be maximum. This is only a guess. To be definite, we must tell how many M_1 models and how many M_2 models should be produced per week, in order that his profits may be maximum.

Mathematical Formulation

Decision variables: Let

X_1 = number of units of M_1 model, and

X_2 = number of units of M_2 model.

Objective function: The objective of the manufacturer is to determine the number of M_1 and M_2 models so as to maximize the total profit.

$Z= 3x_1 + 4x_2$

Constraints: For grindinding since each M_1 model requires 4 hours and each M_2 model requires 2 hours the total number of grinding hours needed per week is given by $4x_1 + 2x_2$. Similarly for polishing, the total number of polishing hours needed per week is $2x_1+5x_2$ Further, since the manufacturer does not have more than 2 x 40(=80) hours of grinding and 3x60(=180) hours of polishing, the time constraints are

$4x_1 + 2x_2 \leq 80$ and $2x_1+ 5x_2 \leq 180$

Non- negativity constraints: Since the production of negative number of models is meaningless, we must have $x_1 \geq 0$ and $x_2 \geq 0$

Hence, the manufacturer's allocation problem can be put in the following mathematical form:

Find two real numbers, x_1 and x_2 such that

1. $4x_1 + 2x_2 \leq 80$
2. $2x_1 + 5x_2 \leq 180$
3. $x_1 \geq 0$, $x_2 \geq 0$

and for which the expression (objective function) $z = 3x_1 + 4x_2$.

Ex .2: Universal Corporation manufactures two products- P_1 and P_2. The profit per unit of the two products is Rs. 50 and Rs. 60 respectively. Both the products require processing in three machines. The following table indicates the available machine hours per week and the time required on each machine for one unit of P_1 and P_2. Formulate this product mix problem in the linear programming form.

Machine	**Product**		**Available Time (in machine hours per week)**
	P_1	**P_2**	
1	2	1	300
2	3	4	509
3	4	7	812
Profit	Rs. 50	Rs. 60	

Solution:

Let x_1 and x_2 be the amounts manufactured of products P_1 and P_2 respectively. The objective here is to **maximize** the profit, which is given by the linear function

Maximize $z = 50x_1 + 60x_2$

Since one unit of product P_1 requires two hours of processing in machine 1, while the corresponding requirement of P_2 is one hour, the first constraint can be expressed as

$2x_1 + x_2 \leq 300$

Similarly, constraints corresponding to machine 2 and machine 3 are

$3x_1 + 4x_2 \leq 509$

$4x_1 + 7x_2 \leq 812$

In addition, there **cannot** be any **negative production** that may be stated algebraically as

$x_1 \geq 0$, $x_2 \geq 0$

(A variable that is also allowed to assume negative values is said to be unrestricted in sign.)

The problem can now be stated in the standard linear programming form as

Maximize $z = 50x_1 + 60x_2$

subject to

$2x_1 + x_2 \leq 300$

$3x_1 + 4x_2 \leq 509$

$4x_1 + 7x_2 \leq 812$

$x_1 \geq 0$, $x_2 \geq 0$

This procedure is commonly referred to as the formulation of the problem.

Ex.3: The Best Stuffing Company manufactures two types of packing tins- round & flat. Major production facilities involved are cutting and joining. The cutting department can process 200 round tins or 400 flat tins per hour. The joining department can process 400 round tins or 200 flat tins per hour. If the contribution towards profit for a round tin is the same as that of a flat tin, what is the optimal production level?

Solution:

Let

x_1 = number of round tins per hour

x_2 = number of flat tins per hour

Since the contribution towards profit is identical for both the products, the objective function can be expressed as **$x_1 + x_2$.** Hence, the problem can be formulated as

Maximize $Z = x_1 + x_2$

subject to
$(1/200)x_1 + (1/400)x_2 \leq 1$
$(1/400)x_1 + (1/200)x_2 \leq 1$
$x_1 \geq 0,\ x_2 \geq 0$
i.e., $2x_1 + x_2 \leq 400$
$x_1 + 2x_2 \leq 400$
$x_1 \geq 0,\ x_2 \geq 0$

Solve Linear Programs by Graphical Method
A linear program can be solved by multiple methods. In this section, we are going to look at the Graphical method for solving a linear program. This method is used to solve a two-variable linear program. If you have only two decision variables, you should use the graphical method to find the optimal solution.
A graphical method involves formulating a set of linear inequalities subject to the constraints. Then the inequalities are plotted on an X-Y plane. Once we have plotted all the inequalities on a graph the intersecting region gives us a feasible region. The feasible region explains what all values our model can take. And it also gives us the optimal solution.
Let's understand this with the help of an example.

Example: A farmer has recently acquired a 110 hectares piece of land. He has decided to grow Wheat and barley on that land. Due to the quality of the sun and the region's excellent climate, the entire production of Wheat and Barley can be sold. He wants to know how to plant each variety in the 110 hectares, given the costs, net profits and labor requirements according to the data shown below:

Variety	Cost (Price/Hec)	Net Profit (Price/Hec)	Man-days/Hec
Wheat	100	50	10
Barley	200	120	30

The farmer has a budget of US$10,000 and availability of 1,200 man-days during the planning horizon. Find the optimal solution and the optimal value.

Solution: To solve this problem, first we gonna formulate our linear program.
Formulation of Linear Problem
Step 1: Identify the decision variables
The total area for growing Wheat = X (in hectares)
The total area for growing Barley = Y (in hectares)
X and Y are my decision variables.

Step 2: Write the objective function
Since the production from the entire land can be sold in the market. The farmer would want to maximize the profit for his total produce. We are given net profit for both Wheat and Barley. The farmer earns a net profit of US$50 for each hectare of Wheat and US$120 for each Barley.
Our objective function (given by Z) is, **Max Z = 50X + 120Y**

Step 3: Writing the constraints
1. It is given that the farmer has a total budget of US$10,000. The cost of producing Wheat and Barley per hectare is also given to us. We have an upper cap on the total cost spent by the farmer. So our equation becomes:

$$\mathbf{100X + 200Y \leq 10,000}$$

2. The next constraint is the upper cap on the availability of the total number of man-days for the planning horizon. The total number of man-days available is 1200. As per the table, we are given the man-days per hectare for Wheat and Barley.

$$\mathbf{10X + 30Y \leq 1200}$$

3. The third constraint is the total area present for plantation. The total available area is 110 hectares. So the equation becomes,

$$\mathbf{X + Y \leq 110}$$

Step 4: The non-negativity restriction

The values of X and Y will be greater than or equal to 0. This goes without saying.

$$\mathbf{X \geq 0, Y \geq 0}$$

We have formulated our linear program. It's time to solve it.

Solving an LP through Graphical method

Since we know that X, Y ≥ 0. We will consider only the first quadrant.

To plot for the graph for the above equations, first I will simplify all the equations.

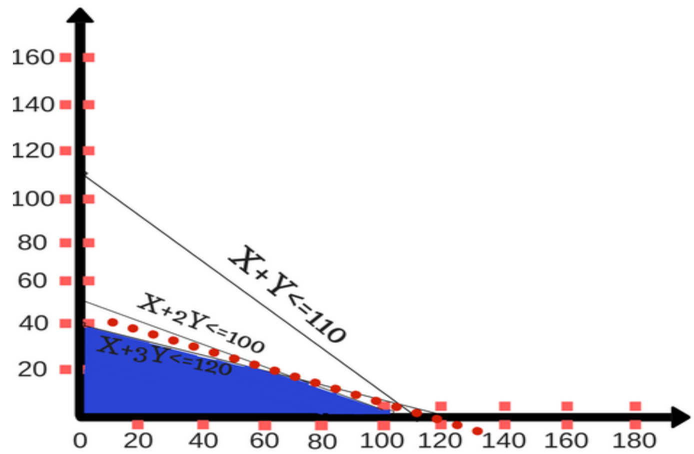
100X + 200Y ≤ 10,000 can be simplified to X + 2Y ≤ 100 by dividing by 100.

10X + 30Y ≤ 1200 can be simplified to X + 3Y ≤ 120 by dividing by 10.

The third equation is in its simplified form, X + Y ≤ 110.

Plot the first 2 lines on a graph in the first quadrant (like shown below)

The optimal feasible solution is achieved at the point of intersection where the budget & man-days constraints are active. This means the point at which the equations X + 2Y ≤ 100 and X + 3Y ≤ 120 intersect gives us the optimal solution.



The values for X and Y which gives the optimal solution is at (60,20).

To maximize profit the farmer should produce Wheat and Barley in 60 hectares and 20 hectares of land respectively.

The maximum profit the company will gain is,

Max Z = 50 * (60) + 120 * (20)

= US$5400

Simplex Method

Simplex Method is one of the most powerful & popular methods for linear programming. The simplex method is an iterative procedure for getting the most feasible solution. In this method, we keep transforming the value of basic variables to get maximum value for the objective function.

A linear programming function is in its **standard form** if it seeks to maximize the objective

$$z = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$$

function. subject to constraints,

$$a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n \leq b_1$$

$$a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n \leq b_2$$

. . . . . .

. . . . . .

$$a_{m1}\, x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \le b_m$$

where, $x_i \ge 0$ and $b_i \ge 0$. After adding slack variables, the corresponding system of constraint equation is,

$$a_{11}\, x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + s_1 = b_1$$

$$a_{21}\, x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + s_2 = b_2$$

. . . .

$$a_{m1}\, x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + s_m = b_m$$

where, $s_i \ge 0$

The variables, s_1 s_2 ………………. s_m are called slack variables. They are non-negative numbers that are added to remove the inequalities from an equation.

The above explanation gives the theoretical explanation of the simplex method. Now, I am gonna explain how to use the simplex method in real life using Excel.

Example: The advertising alternatives for a company include television, newspaper, and radio advertisements. The cost for each medium with its audience coverage is given below.

[su_table]

	Television	Newspaper	Radio
Cost per advertisement ($)	2000	600	300
Audience per advertisement	100,000	40,000	18,000

[/su_table]

The local newspaper limits the number of advertisements from a single company to ten. Moreover, to balance the advertising among the three types of media, no more than half of the total number of advertisements should occur on the radio. And at least 10% should occur on television. The weekly advertising budget is $18,200. How many advertisements should be run in each of the three types of media to maximize the total audience?

Solution: First I am going to formulate my problem for a clear understanding.

Step 1: Identify Decision Variables

Let X_1 , X_2 , X_3 represent the total number of ads for television, newspaper, and radio respectively.

Step 2: Objective Function

The objective of the company is to maximize the audience. The objective function is given by:

$$Z = 100{,}000X_1 + 40{,}000X_2 + 18{,}000X_3$$

Step 3: Write down the constraints

Now, I will mention each constraint one by one.

It is clearly given that we have a budget constraint. The total budget which can be allocated is \$18,200. And the individual costs per television, newspaper and radio advertisement is \$2000, \$600 and \$300 respectively. This can be represented by the equation,

$$2000X_1 + 600X_2 + 300X_3 \leq 18{,}200$$

For a newspaper advertisement, there is an upper cap on the number of advertisements to 10.

My first constraints are, $X_2 \leq 10$

The next constraint is the number of advertisements on television. The company wants at least 10% of the total advertisements to be on television. So, it can be represented as:

$$X_1 \geq 0.10(X_1 + X_2 + X_3)$$

The last constraint is the number of advertisements on the radio cannot be more than half of the total number of advertisements. It can be represented as

$$X_3 \leq 0.5(X_1 + X_2 + X_3)$$

Now, I have formulated my linear programming problem. We are using the simplex method to solve this. I will take you through the simplex method one by one.

To reiterate all the constraints are as follows. I have simplified the last two equations to bring them in standard form.

$$2000X_1 + 600X_2 + 300X_3 \leq 18{,}200$$

$$X_2 \leq 10 \quad -9X_1 + X_2 + X_3 \leq 0$$

$$-X_1 - X_2 + X_3 \leq 0$$

We have a total of 4 equations. To balance out each equation, I am introducing 4 slack variables, S_1 S_2 , S_3 and S_4 .

So our equations are as follows:

$$2000X_1 + 600X_2 + 300X_3 + S_1 = 18{,}200$$

$$X_2 + S_2 = 10$$

$$-9X_1 + X_2 + X_3 + S_3 = 0$$

$$-X_1 - X_2 + X_3 + S_4 = 0$$

I hope now you are available to make sense of the entire advertising problem. All the above equations are only for your better understanding. Now if you solve these equations, you will get the values for X1= 4, X2= 10 and X3= 14.

On solving the objective function, you will get the maximum weekly audience as 1,052,000..

NORTHWEST CORNER METHOD AND LEAST COST METHOD

The northwest corner method is a special type method used for transportation problems in linear programming. It is used to calculate the feasible solution for transporting commodities from one place to another. Whenever you are given a real-world problem, which involves supply and demand from one source of a different sources. The data model includes the following:

- The level of supply and demand at each source is given
- The unit transportation of a commodity from each source to each destination

The model assumes that there is only one commodity. The demand for which can come from different sources. The objective is to fulfill the total demand with minimum transportation cost. The model is based on the hypothesis that the total demand is equal to the total supply, i.e the model is balanced. Let's understand this with the help of an example.

Example: Consider there are 3 silos which are required to satisfy the demand from 4 mills. (A silo is a storage area of the farm used to store grain and Mill is a grinding factory for grains).

From	Mill 1	Mill 2	Mill 3	Mill 4	Supply (in units)
Silo 1	$10	$2	$20	$11	15
Silo 2	$12	$7	$9	$20	25
Silo 3	$4	$14	$16	$18	10
Demand (in units)	5	15	15	15	

Solution: Let's understand what the above table explains.

The cost of transportation from Silo *i* to Mill *j* is given by the cost in each cell corresponding to the supply from each silo 1 and the demand at each Mill. For example, The cost of transporting from Silo 1 to Mill 1 is $10, from Silo 3 to Mill 5 is $18. It is also given the total demand & supply for mill and silos. The objective is to find the minimal transportation cost such that the demand for all the mills is satisfied.

As the name suggests Northwest corner method is a method of allocating the units starting from the top-left cell. The demand for Mill 1 is 5 and Silo 1 has a total supply of 15. So, 5 units can be allocated to Mill1 at a cost of $10 per unit. The demand for Mill1 is met. then we move to the top-left cell of Mill 2. The demand for Mill 2 is 15 units, which it can get 10 units from Silo 1 at a cost of $2 per unit and 5 units from Silo 2 at a cost of $7 per unit. Then we move onto Mill 3, the northwest cell is S2M3. The demand for Mill 3 is 15 units, which it can get from Silo 2 at a cost of $9 per unit. Moving on to the last Mill, Mill 4 has a demand of 15 units. It will get 5 units from a Silo 2 at a cost of $20 per unit and 10 units from Silo 3 at a cost of $18 per unit.

The total cost of transportation is = 5*10+(2*10+7*5)+9*15+(20*5+18*10) = $520

From	Mill 1	Mill 2	Mill 3	Mill 4	Supply
Silo 1	**5** 10	**10** 2	20	11	15
Silo 2	12	**5** 7	**15** 9	**5** 20	25
Silo 3	4	14	16	**10** 18	10
Demand	5	15	15	15	

LEAST COST METHOD

Least Cost method is another method to calculate the most feasible solution for a linear programming problem. This method derives more accurate results than Northwest corner method. It is used for transportation and manufacturing problems. To keep it simple I am explaining the above transportation problem.

From	Mill 1	Mill 2	Mill 3	Mill 4	Supply
Silo 1	10	**15** 2	20	11	15
Silo 2	12	7	**15** 9	**10** 20	25
Silo 3	**5** 4	14	16	**5** 18	10
Demand	5	15	15	15	

According to the least cost method, you start from the cell containing the least unit cost for transportation. So, for the above problem, I supply 5 units from Silo 3 at a per-unit cost of $4. The demand for Mill1 is met. For Mill 2, we supply 15 units from Silo 1 at a per unit cost of $2. Then For Mill 3 we supply 15 units from Silo 2 at a per-unit cost of $9. Then for Mill 4 we supply 10 units from Silo 2 at a per unit cost of $20 and 5 units from Silo 3 an $18 per unit. The total transportation costs are $475.

	A	B	C	D	E	F	G	H	I
2									
3		From	Mill 1	Mill 2	Mill 3	Mill 4	**Supply**	Total	
4		Silo 1					**15**	0	
5		Silo 2					**25**	0	
6		Silo 3					**10**	0	
7		**Demand**	**5**	**15**	**15**	**15**			
8		Total	0	0	0	0	50		
9		Total Cost	0						
10									
11		From	Mill 1	Mill 2	Mill 3	Mill 4	**Supply**	Total	
12		Silo 1	10	2	20	11	**15**	43	
13		Silo 2	12	7	9	20	**25**	48	
14		Silo 3	4	14	16	18	**10**	52	
15		**Demand**	**5**	**15**	**15**	**15**			
16		Total	26	23	45	49	50		

	A	B	C	D	E	F	G	H
1								
2								
3		From	Mill 1	Mill 2	Mill 3	Mill 4	**Supply**	Total
4		Silo 1	10	2	20	11	**15**	43
5		Silo 2	12	7	9	20	**25**	48
6		Silo 3	4	14	16	18	**10**	52
7		**Demand**	**5**	**15**	**15**	**15**		
8		Total	26	23	45	49	50	
9								

Well, the above method explains we can optimize our costs further with the best method. Let's check this using Excel Solver. Solver is an in-built add-on in Microsoft Excel. It's an add-in plug available in Excel. Go to file->options->add-ins->select solver->click on manage->select solver->click Ok. Your solver is now added in excel. You can check it under the Data tab.

The first thing I am gonna do is enter my data in excel. After entering the data in excel, I have calculated the total of C3:F3. Similarly for others. This is done to take the total demand from Silo 1 and others.

After this, I am gonna break my model into two. The first table gives me the units supplied and the second table gives me the unit cost.

	A	B	C	D	E	F	G	H	I
2									
3		From	Mill 1	Mill 2	Mill 3	Mill 4	**Supply**	Total	
4		Silo 1						0	
5		Silo 2						0	
6		Silo 3						0	
7		**Demand**							
8		Total	0	0	0	0	0		
9									
10									
11		From	Mill 1	Mill 2	Mill 3	Mill 4	**Supply**	Total	
12		Silo 1	10	2	20	11	**15**	43	
13		Silo 2	12	7	9	20	**25**	48	
14		Silo 3	4	14	16	18	**10**	52	
15		**Demand**	**5**	**15**	**15**	**15**			
16		Total	26	23	45	49	50		
17									

Now, I am calculating my total cost which will be given by Sumproduct of unit cost and units supplied.

Now I am gonna use Solver to compute my model. Similar to the above method. Add the objective function, variable cells, constraints.

Now your model is ready to be solved. Click on solve and you will get your optimal cost. The minimum transportation cost is $435.

From	Mill 1	Mill 2	Mill 3	Mill 4	Supply	Total
Silo 1					15	0
Silo 2					25	0
Silo 3					10	
Demand	5	15	15	15		
Total	0	0	0	0	50	
Total Cost	0					

Variable Cells - C3:F6

Objective function - C9

Constraints: H4:H6 <= G4:G6
C8:F8 >= C7:F7
C3:F6 >= 0
C3: F6 -> int

Add Constraint

Cell Reference: Constraint:

OK Add Cancel

From	Mill 1	Mill 2	Mill 3	Mill 4	Supply	Total
Silo 1	10	2	20	11	15	43
Silo 2	12	7	9	20	25	48
Silo 3	4	14	16	18	10	52
Demand	5	15	15	15		
Total	26	23	45	49	50	

From	Mill 1	Mill 2	Mill 3	Mill 4	Supply	Total
Silo 1	0	5	0	10	15	15
Silo 2	0	10	15	0	25	25
Silo 3	5	0	0	5	10	10
Demand	5	15	15	15		
Total	5	15	15	15	50	
Total Cost	435					

From	Mill 1	Mill 2	Mill 3	Mill 4	Supply	Total
Silo 1	10	2	20	11	15	43
Silo 2	12	7	9	20	25	48
Silo 3	4	14	16	18	10	52
Demand	5	15	15	15		
Total	26	23	45	49	50	

ADVANTAGES OF LINEAR PROGRAMMING

Following are certain advantages of linear programming:

1. Linear programming helps in attaining the optimum use of productive resources. It also indicates how a decision-maker can employ his productive factors effectively by selecting and distributing (allocating) these resources.
2. Linear programming techniques improve the quality of decisions. The decision-making approach of the user of this technique becomes more objective and less subjective.
3. Linear programming techniques provide possible and practical solutions since there might be other constraints operating outside the problem which must be taken into account. Just because we can produce so many units docs not mean that they can be sold. Thus, necessary modification of its mathematical solution is required for the sake of convenience to the decision-maker.
4. Highlighting of bottlenecks in the production processes is the most significant advantage of this technique. For example, when a bottleneck occurs, some machines cannot meet demand while other remains idle for some of the time.
5. Linear programming also helps in re-evaluation of a basic plan for changing conditions. If conditions change when the plan is partly carried out, they can be determined so as to adjust the remainder of the plan for best results.

LIMITATIONS OF LINEAR PROGRAMMING

1. There should be an objective which should be clearly identifiable and measurable in quantitative terms. It could be, for example, maximisation of sales, of profit, minimisation of cost, and so on, which is not possible in real life.
2. The activities to be included should be distinctly identifiable and measurable in quantitative terms, for instance, the products included in a production planning problem and all the activities can't be measured in quantitative terms for example if labour is sick, which will decrease his performance which can't be measured.
3. The resources of the system which arc to be allocated for the attainment of the goal should also be identifiable and measurable quantitatively. They must be in limited supply. The technique would involve allocation of these resources in a manner that would trade off the returns on the investment of the resources for the attainment of the objective.
4. The relationships representing the objective as also the resource limitation considerations, represented by the objective function and the constraint equations or inequalities, respectively must be linear in nature, which is not possible.
5. There should be a series of feasible alternative courses of action available to the decision makers, which are determined by the resource constraints.

When these stated conditions are satisfied in a given situation, the problem can be expressed in algebraic form, called the Linear Programming Problem (LPP) and then solved for optimal decision.

- While solving an LP model, there is no guarantee that we will get integer valued solutions.

For example, in finding out how many men and machines would be required lo perform a particular job, a non-integer valued solution will be meaningless. Rounding off the solution to the nearest integer will not yield an optimal solution. In such cases, integer programming is used to ensure integer value to the decision variables.

- Linear programming model does not take into consideration the effect of time and uncertainty. Thus, the LP model should be defined in such a way that any change due to internal as well as external factors can be incorporated.
- Sometimes large-scale problems can be solved with linear programming techniques even when assistance of computer is available. For it, the main problem can be fragmented into several small problems and solving each one separately.
- Parameters appearing in the model are assumed to be constant but in real-life situations, they are frequently neither known nor constant.
- Parameters like human behaviour, weather conditions, stress of employees, demotivated employee can't be taken into account which can adversely effect any organisation
- Only one single objective is dealt with while in real life situations, problems come with multi-objectives.

Applications of Linear Programming

Linear programming and Optimization are used in various industries. The manufacturing and service industry uses linear programming on a regular basis. In this section, we are going to look at the various applications of Linear programming.

1. Manufacturing industries use linear programming for **analyzing their supply chain operations**. Their motive is to maximize efficiency with minimum operation cost. As per the recommendations from the linear programming model, the manufacturer can reconfigure their storage layout, adjust their workforce and reduce the bottlenecks. Here

is a small Warehouse case study of Cequent a US-based company, watch this video for a more clear understanding.

2. Linear programming is also used in organized retail for **shelf space optimization**. Since the number of products in the market has increased in leaps and bounds, it is important to understand what does the customer want. Optimization is aggressively used in stores like Walmart, Hypercity, Reliance, Big Bazaar, etc. The products in the store are placed strategically keeping in mind the customer shopping pattern. The objective is to make it easy for a customer to locate & select the right products. This is subject to constraints like limited shelf space, a variety of products, etc.
3. Optimization is also used for **optimizing Delivery Routes**. This is an extension of the popular traveling salesman problem. The service industry uses optimization for finding the best route for multiple salesmen traveling to multiple cities. With the help of clustering and greedy algorithm, the delivery routes are decided by companies like FedEx, Amazon, etc. The objective is to minimize the operation cost and time.
4. Optimizations are also used in **Machine Learning**. Supervised Learning works on the fundamental of linear programming. A system is trained to fit on a mathematical model of a function from the labeled input data that can predict values from an unknown test data.

APPLICATION AREAS OF LINEAR PROGRAMMING

Linear programming is the most widely used technique of decision-making in business and Industry and in various other fields. In this section, we will discuss a few of the broad application areas of linear programming.

Agricultural Applications

These applications fall into categories of farm economics and farm management. The former deals with agricultural economy of a nation or region, while the latter is concerned with the problems of the individual farm.

The study of farm economics deals with inter-regional competition and optimum allocation of crop production. Efficient production patterns can be specified by a linear programming model under regional land resources and national demand constraints.

Linear programming can be applied in agricultural planning, e.g. allocation of limited resources such as acreage, labour, water supply and working capital, etc. in a way so as to maximise net revenue.

Military Applications

Military applications include the problem of selecting an air weapon system against enemy so as to keep them pinned down and at the same time minimising the amount of aviation gasoline used. A variation of the transportation problem that maximises the total tonnage of bombs dropped on a set of targets and the problem of community defence against disaster, the solution of which yields the number of defence units that should be used in a given attack in order to provide the required level of protection at the lowest possible cost.

Production Management

1. Product mix: A company can produce several different products, each of which requires the use of limited production resources. In such cases, it is essential to determine the quantity of each product to be produced knowing its marginal contribution and amount of available resource used by it. The objective is to maximise the total contribution, subject to all constraints.
2. Production planning: This deals with the determination of minimum cost production plan over planning period of an item with a fluctuating demand, considering the initial number of units in inventory, production capacity, constraints on production, manpower and all relevant cost factors. The objective is to minimise total operation costs.

3. Assembly-line balancing: This problem is likely to arise when an item can be made by assembling different components. The process of assembling requires some specified sequence(s). The objective is to minimise the total elapse time.
4. Blending problems: These problems arise when a product can be made from a variety of available raw materials, each of which has a particular composition and price. The objective here is to determine the minimum cost blend, subject to availability of the raw materials, and minimum and maximum constraints on certain product constituents.
5. Trim loss When an item is made to a standard size (e.g. glass, paper sheet), the problem that arises is to determine which combination of requirements should be produced from standard materials in order to minimise the trim loss.

Financial Management

1. Portfolio selection: This deals with the selection of specific investment activity among several other activities. The objective is to find the allocation which maximises the total expected return or minimises risk under certain limitations.
2. Profit planning: This deal with the maximisation of the profit margin from investment in plant facilities and equipment, cash in hand and inventory.

Marketing Management

1. Media selection: Linear programming technique helps in determining the advertising media mix so as to maximise the effective exposure, subject to limitation of budget, specified exposure rates to different market segments, specified minimum and maximum number of advertisements in various media. (if) Travelling salesman problem The problem of salesman is to find the shortest route from a given city, visiting each of the specified cities and then returning to the original point of departure, provided no city shall be visited twice during the tour. Such type of problems can be solved with the help of the modified assignment technique.
2. Physical distribution: Linear programming determines the most economic and efficient manner of locating manufacturing plants and distribution centres for physical distribution.

Personnel Management

- Staffing problem: Linear programming is used to allocate optimum manpower to a particular job so as to minimise the total overtime cost or total manpower.
- Determination of equitable salaries: Linear programming technique has been used in determining equitable salaries and sales incentives.
- Job evaluation and selection: Selection of suitable person for a specified job and evaluation of job in organisations has been done with the help of linear programming technique.

Other applications of linear programming lie in the area of administration, education, fleet utilisation, awarding contracts, hospital administration and capital budgeting.

STRUCTURE OF LINEAR PROGRAMMING

General Structure of LP Model

The general structure of LP model consists of three components.

1. Decision variables (activities): We need to evaluate various alternatives (courses of action) for arriving at the optimal value of objective function. Obviously, if there are no alternatives to select from, we would not need LP. The evaluation of various alternatives is guided by the nature of objective function and availability of resources. For this, we pursue certain activities usually denoted by x1, x2…xn. The value of these activities represents the extent to which each of these is performed. For example, in a product-mix manufacturing, the management may use LP to decide how many units of each

of the product to manufacture by using its limited resources such as personnel, machinery, money, material, etc.

These activities are also known as decision variables because they arc under the decision maker's control. These decision variables, usually interrelated in terms of consumption of limited resources, require simultaneous solutions. All decision variables are continuous, controllable and non-negative. That is, x1>0, x2>0, ….xn>0.

1. The objective function: The objective function of each L.P problem is a mathematical representation of the objective in terms of a measurable quantity such as profit, cost, revenue, distance, etc. In its general form, it is represented as:

Optimise (Maximise or Minimise) Z = c1x1 + c2X2. … cnxn

Where Z is the measure-of-performance variable, which is a function of x1, x2 …, xn. Quantities c1, c2…cn are parameters that represent the contribution of a unit of the respective variable x1, x2…, xn to the measure-of-performance Z. The optimal value of the given objective function is obtained by the graphical method or simplex method.

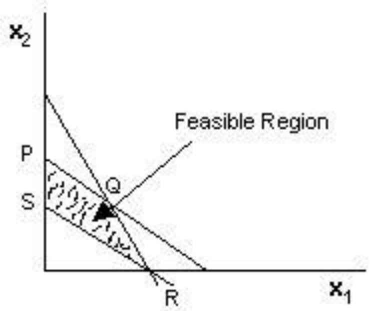
1. The constraints: There are always certain limitations (or constraints) on the use of resources, e.g. labour, machine, raw material, space, money, etc. that limit the degree to which objective can be achieved. Such constraints must be expressed as linear equalities or inequalities in terms of decision variables. The solution of an L.P model must satisfy these constraints. The linear programming method is a technique for choosing the best alternative from a set of feasible alternatives, in situations in which the objective function as well as the constraints can be expressed as linear mathematical functions.

GRAPHICAL METHOD

Linear programming problems with two decision variables can be easily solved by graphical method. A problem of three dimensions can also be solved by this method, but their graphical solution becomes complicated.

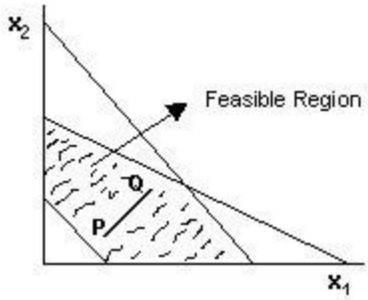
Feasible Region

It is the collection of all feasible solutions. In the following figure, the shaded area represents the feasible region.



2.3 Convex Set

A region or a set R is convex, if for any two points on the set R, the segment connecting those points lies entirely in R. In other words, it is a collection of points such that for any two points on the set, the line joining the points belongs to the set. In the following figure, the line joining P and Q belongs entirely in R.



Thus, the collection of feasible solutions in a linear programming problem form a convex set

2.4 Graphical Method - Algorithm

1. Formulate the mathematical model of the given linear programming problem.
2. Treat inequalities as equalities and then draw the lines corresponding to each equation and non-negativity restrictions.
3. Locate the end points (corner points) on the feasible region.
4. Determine the value of the objective function corresponding to the end points determined in step 3.

5. Find out the optimal value of the objective function.

2.5 Redundant Constraint

It is a constraint that does not affect the feasible region.

Consider the linear programming problem:

Maximize 1170 x_1 + 1110x_2

subject to

$9x_1 + 5x_2 \geq 500$

$7x_1 + 9x_2 \geq 300$

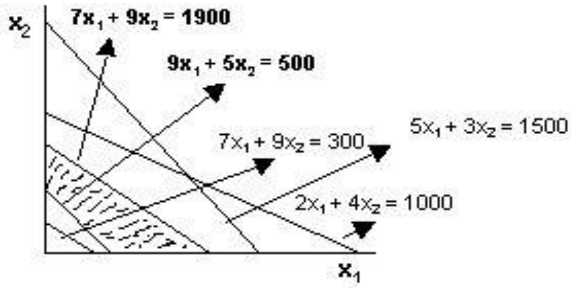
$5x_1 + 3x_2 \leq 1500$

$7x_1 + 9x_2 \leq 1900$

$2x_1 + 4x_2 \leq 1000$

$x_1, x_2 \geq 0$

The feasible region is indicated in the following figure:



The critical region has been formed by the two constraints.

$9x_1 + 5x_2 \geq 500$

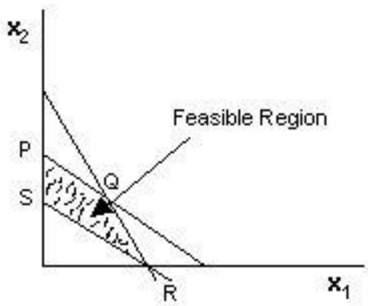
$7x_1 + 9x_2 \leq 1900$

$x_1, x_2 \geq 0$

The remaining three constraints are not affecting the feasible region in any manner. Such constraints are called **redundant constraints**.

2.6 Extreme Point

Extreme points are referred to as vertices or corner points. In the following figure, P, Q, R and S are extreme points.



Example 1.

Maximize z = 18x_1 + 16x_2

subject to

$15x_1 + 25x_2 \leq 375$

$24x_1 + 11x_2 \leq 264$

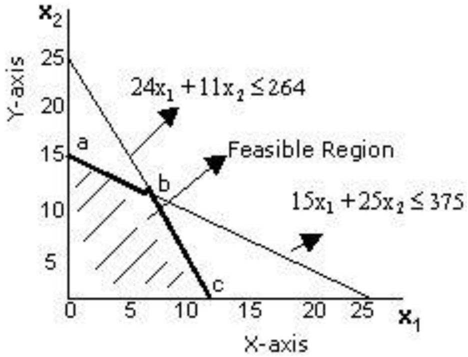
$x_1, x_2 \geq 0$

Solution:

If only x_1 and no x_2 is produced, the maximum value of x_1 is 375/15 = 25. If only x_2 and no x_1 is produced, the maximum value of x_2 is 375/25 = 15. A line drawn between these two points (25, 0) & (0, 15), represents the constraint factor $15x_1 + 25x_2 \leq 375$. Any point which lies on or below this line will satisfy this inequality and the solution will be somewhere in the region bounded by it.

Similarly, the line for the second constraint $24x_1 + 11x_2 \leq 264$ can be drawn. The polygon *oabc* represents the region of values for x_1 & x_2 that satisfy all the constraints. This polygon is called the solution set.

The solution to this simple problem is exhibited graphically below.



The end points (corner points) of the shaded area are (0,0), (11,0), (5.7, 11.58) and (0,15). The values of the objective function at these points are 0, 198, 288 (approx.) and 240. Out of these four values, 288 is maximum.

The optimal solution is at the extreme point b, where $x_1 = 5.7$ & $x_2 = 11.58$, and z = 288.

Example 2

Maximize z = $6x_1 - 2x_2$

subject to

$2x_1 - x_2 \leq 2$

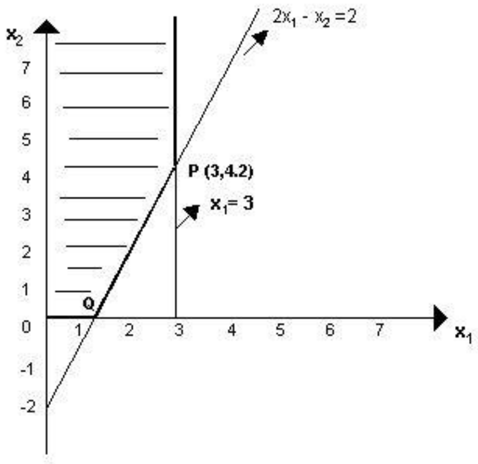
$x_1 \leq 3$

$x_1, x_2 \geq 0$

Solution.

First, we draw the line $2x_1 - x_2 \leq 2$, which passes through the points (1, 0) & (0, -2). Any point which lies on or below this line will satisfy this inequality and the solution will be somewhere in the region bounded by it.

Similarly, the line for the second constraint $x_1 \leq 3$ is drawn. Thus, the optimal solution lies at one of the corner points of the dark shaded portion bounded by these straight lines.



Optimal solution is $x_1 = 3$, $x_2 = 4.2$, and the maximum value of z is 9.6.

Multiple Optimal Solutions

Graphical Method II - Special Cases

The linear programming problems discussed in the previous section possessed unique solutions. This was because the optimal value occurred at one of the extreme points (corner points). But situations may arise, when the optimal solution obtained is not unique. This case may arise when the line representing the objective function is parallel to one of the lines bounding the feasible region. The presence of multiple solutions is illustrated through the following example.

Example:

Maximize z = $x_1 + 2x_2$

subject to

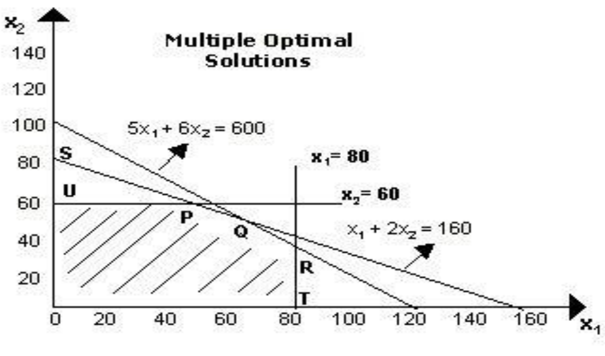
$x_1 \leq 80$

$x_2 \leq 60$

$5x_1 + 6x_2 \leq 600$

$x_1 + 2x_2 \leq 160$

$x_1, x_2 \geq 0$.



In the above figure, there is no unique outer most corner cut by the objective function line. All points from P to Q lying on line PQ represent optimal solutions and all these will give the same optimal value (maximum profit) of Rs. 160. This is indicated by the fact that both the points P with co-ordinates (40, 60) and Q with co-ordinates (60, 50) are on the line $x_1 + 2x_2 = 160$. Thus, every point on the line PQ maximizes the value of the objective function and the problem has multiple solutions.

3.2. Infeasible Problem

In some cases, there is no feasible solution area, i.e., there are no points that satisfy all constraints of the problem. An infeasible LP problem with two decision variables can be identified through its graph. For example, let us consider the following linear programming problem.

Minimize z = 200x_1 + 300x_2

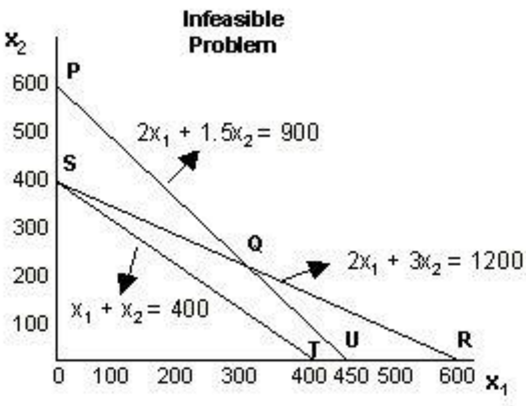
subject to

$2x_1 + 3x_2 \geq 1200$

$x_1 + x_2 \leq 400$

$2x_1 + 1.5x_2 \geq 900$

$x_1, x_2 \geq 0$



The region located on the right of *PQR* includes all solutions, which satisfy the first and the third constraints. The region located on the left of *ST* includes all solutions, which satisfy the second constraint. Thus, the problem is infeasible because there is no set of points that satisfy all the three constraints.

3.3. Unbounded Solutions

It is a solution whose objective function is infinite. If the feasible region is unbounded then one or more decision variables will increase indefinitely without violating feasibility, and the value of the objective function can be made arbitrarily large. Consider the following model:

Minimize z = 40x_1 + 60x_2

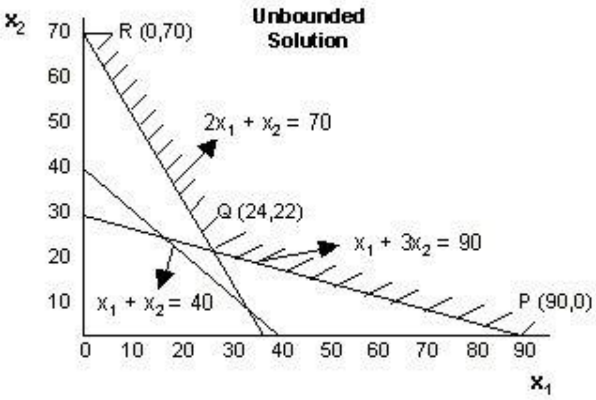
subject to

$2x_1 + x_2 \geq 70$

$x_1 + x_2 \geq 40$

$x_1 + 3x_2 \geq 90$

$x_1, x_2 \geq 0$



The point (x_1, x_2) must be somewhere in the solution space as shown in the figure by shaded portion.

The three extreme points (corner points) in the finite plane are:

P = (90, 0); Q = (24, 22) and R = (0, 70) The values of the objective function at these extreme points are: Z(P) = 3600, Z(Q) = 2280 and Z(R) = 4200.

In this case, no maximum of the objective function exists because the region has no boundary for increasing values of x_1 and x_2. Thus, it is not possible to maximize the objective function in this case and the solution is unbounded.

Note:

Although it is possible to construct linear programming problems with unbounded solutions numerically, but no linear programming problem formulated from a real life situation can have unbounded solution.

Limitations of Linear Programming

- **Linearity of relations:** A primary requirement of linear programming is that the objective function and every constraint must be linear. However, in real life situations, several business and industrial problems are nonlinear in nature.
- **Single objective:** Linear programming takes into account a single objective only, i.e., profit maximization or cost minimization. However, in today's dynamic business environment, there is no single universal objective for all organizations.
- **Certainty:** Linear Programming assumes that the values of co-efficient of decision variables are known with certainty.

- Due to this restrictive assumption, linear programming cannot be applied to a wide variety of problems where values of the coefficients are probabilistic.
- **Constant parameters:** Parameters appearing in LP are assumed to be constant, but in practical situations it is not so.
- **Divisibility:** In linear programming, the decision variables are allowed to take non-negative integer as well as fractional values. However, we quite often face situations where the planning models contain integer valued variables. For instance, trucks in a fleet, generators in a powerhouse, pieces of equipment, investment alternatives and there are a myriad of other examples. Rounding off the solution to the nearest integer will not yield an optimal solution. In such cases, linear programming techniques cannot be used.

Summary:

This chapter initiated your study of linear models. Linear programming is a fascinating topic in operations research with wide applications in various problems of management, economics, finance, marketing, transportation and decision making pertaining to the operations of virtually any private or public organization. Unquestionably, linear programming techniques are among the most commercially successful applications of operations research.
In this chapter, you learned how to formulate a linear programming problem, and then we discussed the graphical method of solving an LPP with two decision variables.

THE SIMPLEX METHOD

Simplex method was developed by George B. Datzing in 1947 is an iterative and an efficient method for solving linear programming problems. It is an algebraic procedure that starts at a feasible extreme point of the simplex(or convex), normally the origin, and systematically moves from one feasible extreme point to another until an optimum(or optimal) extreme point is located. At each iteration, the procedure tests the one extreme (corner) point for optimality, and if not optimum, chooses another extreme point, of the convex set that is formed by the constraints and non-negativity conditions of the linear programming problem. Since the number of extreme points (i.e., corners or vertices) of the convex set of all feasible solutions is finite, the method leads to the optimum extreme point (i.e., optimum or optimal solution) in a finite number of steps or indicates that there exists an unbounded solution.

Basic Definitions:

Slack Variable*:*

It is a variable that is added to the left-hand side of a less than or equal to type constraint to convert the constraint into an equality. In economic terms, slack variables represent left-over or unused capacity.

Specifically:

$$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + ...+ a_{in}x_n \leq b_i$$

can be written as

$$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + ...+ a_{in}x_n + s_i = b_i$$

Where i = 1, 2, ..., m

Surplus variable*:*

It is a variable subtracted from the left-hand side of a greater than or equal to type constraint to convert the constraint into equality. It is also known as negative slack variable. In economic terms, surplus variables represent over fulfillment of the requirement.

Specifically:

$$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + ...+ a_{in}x_n \geq b_i$$

can be written as

$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + ...+ a_{in}x_n - s_i = b_i$

Where i = 1, 2, ..., m

Artificial variable:

It is a non negative variable introduced to facilitate the computation of an initial basic feasible solution. In other words, a variable added to the left-hand side of a greater than or equal to type constraint to convert the constraint into equality is called an artificial variable.

1.3 Canonical Form:

The general formulation of L.P.P can always be put in the following form:

Maximize $z = c_1x_1 + c_2x_2 + \ldots\ldots c_nx_n$ subject to the constraints

$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + ...+ a_{in}x_n \leq b_i$, $i = 1,2,3,....m$

$x_1, x_2, x_3, ... x_n \geq 0$

by making use of some elementary transformations. This form of L.P.P. is called the **canonical form** of L.P.P.

1.4 Characteristic of Canonical form:

(i) The objective function is of the maximization type

Minimization of f(x) = -Maximization (-f(x))

If we have the objective function as min z then change that as min z = -Max (-z).

(ii) All the constraints are of the "≤ "type, except for the non-negative restriction

If we have an inequality of "≥" type, multiply both sides by -1. if the constraints is an equation say

$a_{i1}x_1 + a_{i2}x_2 + \ldots\ldots\ldots a_{in}x_n = b_i$

write it as

$a_{i1}x_1 + a_{i2}x_2 + \ldots\ldots\ldots a_{in}x_n \leq b_i$ and

$a_{i1}x_1 + a_{i2}x_2 + \ldots\ldots\ldots a_{in}x_n \geq b_i$ (or $-a_{i1}x_1 - a_{i2}x_2 - \ldots\ldots\ldots -a_{in}x_n \leq -b_i$)

(iii) All the variables are non-negative

If a variable is unrestricted in sign ((i.e.) positive, negative or zero) then replace that variable by difference between two non-negative variables. If x_j is unrestricted in sign replace x_j by x_j = where both and are both non-negative.

Matrix Notation of Canonical form:

In matrix notation the canonical form of L.P.P can be expressed as:

Maximize or Minimize z = cx (Objective Function)

Subject to the constraints: (Constraints)

$AX \leq b, X \geq 0$ (Non-negative restrictions)

$$\text{where } X = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix}, c = (c_1, c_2, \ldots. c_n), b = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ b_n \end{pmatrix} \text{ and } A = \begin{pmatrix} a_{11} & a_{12} \ldots\ldots a_{1n} \\ a_{21} & a_{22} \ldots\ldots a_{2n} \\ \ldots\ldots & \ldots\ldots \\ a_{m1} & a_{m2} \ldots\ldots a_{mn} \end{pmatrix}$$

1.5 Standard Form:

The general formulation of L.P.P can always be put in the following form:

Maximize $z = c_1x_1 + c_2x_2 + \ldots\ldots c_nx_n$ subject to the constraints

$a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + ...+ a_{in}x_n = b_i$, $i = 1,2,3,....m$

$x_1, x_2, x_3, ... x_n \geq 0$

is known as in **Standard form.**

1.6 Characteristic of Standard form:

(i) The objective function is of the maximization type.

(ii) All the constraints are expressed in the form of equations except for the non-negative constraints.

If the constraints are of "≤" type add the slack variable to the left hand side and if the constraints are of "≥" type subtract surplus variable to the left hand side.

(iii) The right hand side of each constraint equation is of non-negative.

1.7 Matrix Notation of Standard form:

In matrix notation the Standard form of L.P.P can be expressed as:

Maximize or Minimize z = cx (Objective Function)

Subject to the constraints: (Constraints)

AX = b, X ≥ 0 (Non-negative restrictions)

$$\text{where } X = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix}, \; c = (c_1, c_2, \ldots c_n), \; b = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ b_n \end{pmatrix} \text{ and } A = \begin{pmatrix} a_{11} & a_{12} \ldots\ldots a_{1n} \\ a_{21} & a_{22} \ldots\ldots a_{2n} \\ \ldots\ldots & \ldots\ldots \\ a_{m1} & a_{m2} \ldots\ldots a_{mn} \end{pmatrix}$$

Basic Solution, Basic and Non- basic variable*:*

Given a system of m linear equations with n variables (m < n). The solution obtained by setting (n – m) variables equal to zero and solving for the remaining m variables is called a **basic solution.**

The m variables are called **basic variables** and they form the basic solution. The n-m variables which are put to zero are called as **non-basic variables.**

Degenerate basic solution:

A basic solution is said to be a **degenerate basic solution** if one or more of the basic variables are zero.

Basic Feasible solution:

A feasible solution which is also basic is called a **Basic Feasible solution.**

ALGORITHM OF SIMPLEX METHOD

Assuming the existence of an initial basic feasible solution, an optimal solution to any L.P.P by simplex method is found as follows:

Step 1: Check whether the objective function is to be maximized or minimized. If it is to be minimized, then convert it into a problem of maximization, by

Minimize Z = -Maximize (-Z)

Step 2: Check whether all b_i's are positive. If any of the b_i's is negative, multiply both sides of that constraint by -1 so as to make its right hand side positive.

Step 3: By introducing slack / surplus variables, convert the inequality constraints into equations and express the given L.P.P into its standard form.

Step 4: Find an initial basic feasible solution and express the above information conveniently in the following simplex table.

		C_j	C_1	C_2	C_3	…………………	0	0	0	……………
C_B	Y_B	X_B	X_1	X_2	X_3	…………………	S_1	S_2	S_3	……………
C_{B1}	S_1	b_1	a_{11}	a_{12}	a_{13}	…………………	1	0	0	……………
C_{B2}	S_2	b_2	a_{21}	a_{22}	a_{23}	…………………	0	1	0	……………
C_{B3}	S_3	b_3	a_{31}	a_{32}	a_{33}	…………………	0	0	1	……………

Body Matrix **Unit Matrix**

Zj-Cj Z_0 Z_1-C_1 Z_2-C_2 Z-C_3 ………………… ……………… ………

C_j = Row denotes the coefficients of the variables in the objective function.

C_B = Column denotes the coefficients of the basic variables in the objective function.

Y_B = Column denotes the basic variables.

X_B = Column denotes the values of the basic variables.

The coefficients of the non-basic variables in the constraint equations constitute the body matrix while the coefficients of the basic variables constitute the unit matrix. The row

Z_j –C_j denotes the net evaluations or index for each column.

Step 5: Compute the net evaluations Z_j –C_j (j= 1,2,3…..n) by using the relation

Z_j –C_j = $C_B a_j$- C_j

Examine the sign of Z_j –C_j

a) If all Z_j –$C_j \geq 0$ then the current basic solution X_B is optimal.

b) If atleast Z_j –$C_j < 0$ then the current basic solution X_B is not optimal, go to next step.

Step 6: (To find the entering variable)

The entering variable is the non-basic variable corresponding to the most negative value Z_j –C_j. Let it be x_r for some j = r. The entering variable column is known at the bottom. If more than one variable has the same most negative Z_j –C_j, any of these variables may be selected arbitrarily as the entering variable.

Step 7: (To find the leaving variable)

Compute the ratio;

$$\theta = \text{Min} \left\{ \frac{X_{Bi}}{a_{ir}}, a_{ir} > 0 \right\}$$

(i.e) the ratio between the solution column and the entering variable column by considering only the positive denominators.

a) If all $a_{ir} \leq 0$, then there is an unbounded solution to the given L.P.P.

b) If all $a_{ir} > 0$, then the leaving variable is the basic variable corresponding to the minimum ratio θ. If,

$$\theta = \frac{X_{Bk}}{a_{kr}}$$

then the basic variable x_k leaves the basis. The leaving variable row is called the **key row** or **pivot row** or **pivot equation** and the element at the intersection of the pivot column and pivot row is called the **pivot element** or **key element** or **leading element**.

Step 8:

Drop the leaving variable and introduce the entering variable along with its associated value under C_B column. Convert the pivot element to unity by dividing the pivot equation by the pivot element and all other elements in its column to zero by making use of

- New pivot equation = old pivot equation / pivot element
- New equation (all other rows including $Z_j - C_j$ row)

New equation = (Corresponding element) X New pivot equation

Step 9: Go to step 5 and repeat the procedure until either an optimum solution is obtained or there is an indication of an unbounded solution.

Example:

Maximize $z = 3x_1 + 2x_2$

Subject to

$-x_1 + 2x_2 \leq 4$

$3x_1 + 2x_2 \leq 14$

$x_1 - x_2 \leq 3$

$x_1, x_2 \geq 0$

Solution:

First, convert every inequality constraints in the LPP into an equality constraint, so that the problem can be written in a standard from. This can be accomplished by adding a slack variable to each constraint. Slack variables are always added to the less than type constraints.

Converting inequalities to equalities

$-x_1 + 2x_2 + x_3 = 4$

$3x_1 + 2x_2 + x_4 = 14$

$x_1 - x_2 + x_5 = 3$

$x_1, x_2, x_3, x_4, x_5 \geq 0$

where x_3, x_4 and x_5 are slack variables.

Since slack variables represent unused resources, their contribution in the objective function is zero. Including these slack variables in the objective function, we get

$$\textbf{Maximize } z = 3x_1 + 2x_2 + 0x_3 + 0x_4 + 0x_5$$

Initial basic feasible solution

Now we assume that nothing can be produced. Therefore, the values of the decision variables are zero. $x_1 = 0$, $x_2 = 0$, $z = 0$

When we are not producing anything, obviously we are left with unused capacity $x_3 = 4$, $x_4 = 14$, $x_5 =$

We note that the current solution has three variables (slack variables x_3, x_4 and x_5) with non-zero solution values and two variables (decision variables x_1 and x_2) with zero values. **Variables with non-zero values are called basic variables. Variables with zero values are called non-basic variables.**

Iteration 1:

	c_j	3	2	0	0	0	
c_B	**Basic variables B**	x_1	x_2	x_3	x_4	x_5	**Solution values b ($=X_B$)**
0	x_3	-1	2	1	0	0	4
0	x_4	3	2	0	1	0	14
0	x_5	1	-1	0	0	1	3
z_j-c_j		-3	-2	0	0	0	

$a_{11} = -1$, $a_{12} = 2$, $a_{13} = 1$, $a_{14} = 0$, $a_{15} = 0$, $b_1 = 4$
$a_{21} = 3$, $a_{22} = 2$, $a_{23} = 0$, $a_{24} = 1$, $a_{25} = 0$, $b_2 = 14$
$a_{31} = 1$, $a_{32} = -1$, $a_{33} = 0$, $a_{34} = 0$, $a_{35} = 1$, $b_3 = 3$
Calculating values for the index row ($z_j – c_j$)
$z_1 – c_1 = (0 \times (-1) + 0 \times 3 + 0 \times 1) - 3 = -3$
$z_2 – c_2 = (0 \times 2 + 0 \times 2 + 0 \times (-1)) - 2 = -2$
$z_3 – c_3 = (0 \times 1 + 0 \times 0 + 0 \times 0) - 0 = 0$
$z_4 – c_4 = (0 \times 0 + 0 \times 1 + 0 \times 0) - 0 = 0$
$z_5 – c_5 = (0 \times 0 + 0 \times 0 + 0 \times 1) – 0 = 0$
Choose the smallest negative value from $z_j – c_j$ (i.e., – 3). So column under x_1 is the key column.
Now find out the minimum positive value
Minimum (14/3, 3/1) = 3
So row x_5 is the key row.
Here, the pivot (key) element = 1 (the value at the point of intersection).
Therefore, x_5 departs and x_1 enters.
We obtain the elements of the next table using the following rules:

1. If the values of $z_j – c_j$ are positive, the inclusion of any basic variable will not increase the value of the objective function. Hence, the present solution maximizes the objective function. If there are more than one negative values, we choose the variable as a basic variable corresponding to which the value of $z_j – c_j$ is least (most negative) as this will maximize the profit.
2. The numbers in the replacing row may be obtained by dividing the key row elements by the pivot element and the numbers in the other two rows may be calculated by using the formula:

$$\text{New number} = \text{old number} - \frac{(\text{corresponding no. of key row}) \times (\text{corresponding no. of key column})}{\text{pivot element}}$$

Calculating values for Iteration 2
x_3 row
$a_{11} = -1 – 1 \times ((-1)/1) = 0$
$a_{12} = 2 – (-1) \times ((-1)/1) = 1$
$a_{13} = 1 – 0 \times ((-1)/1) = 1$
$a_{14} = 0 – 0 \times ((-1)/1) = 0$
$a_{15} = 0 – 1 \times ((-1)/1) = 1$
$b_1 = 4 – 3 \times ((-1)/1) = 7$
x_4 row
$a_{21} = 3 – 1 \times (3/1) = 0$
$a_{22} = 2 – (-1) \times (3/1) = 5$
$a_{23} = 0 – 0 \times (3/1) = 0$
$a_{24} = 1 – 0 \times (3/1) = 1$
$a_{25} = 0 – 1 \times (3/1) = -3$
$b_2 = 14 – 3 \times (3/1) = 5$

x_1 row
a_{31} = 1/1 = 1
a_{32} = -1/1 = -1
a_{33} = 0/1 = 0
a_{34} = 0/1 = 0
a_{35} = 1/1 = 1
b_3 = 3/1 = 3
Iteration 2:

	c_j	3	2	0	0	0	
c_B	**Basic variables B**	x_1	x_2	x_3	x_4	x_5	**Solution values b (= X_B)**
0	x_3	0	1	1	0	1	7
0	x_4	0	5	0	1	-3	5
3	x_1	1	-1	0	0	1	3
z_j-c_j		0	-5	0	0	3	

Calculating values for the index row ($z_j - c_j$)
$z_1 - c_1 = (0 \times 0 + 0 \times 0 + 3 \times 1) - 3 = 0$
$z_2 - c_2 = (0 \times 1 + 0 \times 5 + 3 \times (-1)) - 2 = -5$
$z_3 - c_3 = (0 \times 1 + 0 \times 0 + 3 \times 0) - 0 = 0$
$z_4 - c_4 = (0 \times 0 + 0 \times 1 + 3 \times 0) - 0 = 0$
$z_5 - c_5 = (0 \times 1 + 0 \times (-3) + 3 \times 1) - 0 = 3$
Key column = x_2 column
Minimum (7/1, 5/5) = 1
Key row = x_4 row
Pivot element = 5
x_4 departs and x_2 enters.
Calculating values for table 3
x_3 row
$a_{11} = 0 - 0 \times (1/5) = 0$
$a_{12} = 1 - 5 \times (1/5) = 0$
$a_{13} = 1 - 0 \times (1/5) = 1$
$a_{14} = 0 - 1 \times (1/5) = -1/5$
$a_{15} = 1 - (-3) \times (1/5) = 8/5$
$b_1 = 7 - 5 \times (1/5) = 6$
x_2 row
$a_{21} = 0/5 = 0$
$a_{22} = 5/5 = 1$
$a_{23} = 0/5 = 0$
$a_{24} = 1/5$
$a_{25} = -3/5$
$b_2 = 5/5 = 1$
x_1 row
$a_{31} = 1 - 0 \times (-1/5) = 1$
$a_{32} = -1 - 5 \times (-1/5) = 0$
$a_{33} = 0 - 0 \times (-1/5) = 0$
$a_{34} = 0 - 1 \times (-1/5) = 1/5$
$a_{35} = 1 - (-3) \times (-1/5) = 2/5$
$b_3 = 3 - 5 \times (-1/5) = 4$

Note: Don't convert the fractions into decimals, because many fractions cancel out during the process while the conversion into decimals will cause unnecessary complications.
Final iteration:

	c_j	3	2	0	0	0	
c_B	**Basic variables B**	x_1	x_2	x_3	x_4	x_5	**Solution values b (= X_B)**
0	x_3	0	0	1	-1/5	8/5	6
2	x_2	0	1	0	1/5	-3/5	1
3	x_1	1	0	0	1/5	2/5	4
z_j-c_j		0	0	0	1	0	

Result:
Since all the values of zj – c_j are positive, this is the optimal solution**.**
x_1 = 4, x_2 = 1
Max z = 3 X $_4$ + 2 X $_1$ = 14.
The largest profit of Rs.14 is obtained, when 1 unit of x_2 and 4 units of x_1 are produced. The above solution also indicates that 6 units are still unutilized, as shown by the slack variable x_3 in the X_B column.

Minimization Case:
In the previous section, the simplex method was applied to linear programming problems where the objective was to maximize the profit with less than or equal to type constraints. In many cases, however, constraints may of type ≥ or = and the objective may be minimization (e.g., cost, time, etc.). Thus, in such cases, simplex method must be modified to obtain an optimal policy.
Consider the general linear programming problem
Minimize z = $c_1x_1 + c_2x_2 + c_3x_3 + ...+ c_nx_n$
subject to
$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ...+ a_{1n}x_n \geq b_1$
$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + ...+ a_{2n}x_n \geq b_2$
....
$a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + ...+ a_{mn}x_n \geq b_m$
$x_1, x_2,....., x_n \geq 0$

Changing the sense of the optimization
Any linear minimization problem can be viewed as an equivalent linear maximization problem, and vice versa.
Min. z = $c_1x_1 + c_2x_2 + c_3x_3 + ...+ c_nx_n$
It can be written as
Max. z = - ($c_1x_1 + c_2x_2 + c_3x_3 + ...+ c_nx_n$)

Converting inequalities to equalities
Introducing surplus variables (negative slack variables) to convert inequalities to equalities
$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ...+ a_{1n}x_n - s_1 = b_1$
$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + ...+ a_{2n}x_n - s_2 = b_2$
…..
$a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + ...+ a_{mn}x_n - s_m = b_m$
$x_1, x_2,....., x_n \geq 0$
$s_1, s_2,....., s_m \geq 0$
An initial basic feasible solution is obtained by setting $x_1 = x_2 =........ = x_n = 0$

$-s_1 = b_1$ or $s_1 = -b_1$
$-s_2 = b_2$ or $s_2 = -b_2$
....
$-s_m = b_m$ or $s_m = -b_m$
which is not feasible because it violates the non-negativity stipulation, (i.e., $s_1 \geq 0$).
Therefore, we need artificial variables.
After introducing artificial variables, the set of constraints can be written as
$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ...+ a_{1n}x_n - s_1 + A_1 = b_1$
$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + ...+ a_{2n}x_n - s_2 + A_2 = b_2$
....
$a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + ...+ a_{mn}x_n - s_m + A_m = b_m$
$x_1, x_2,....., x_n \geq 0$
$s_1, s_2,....., s_m \geq 0$
$A_1, A_2,....., A_m \geq 0$
Now, an initial basic feasible solution can be obtained by setting all the decision and surplus variables to zero. Thus, an initial basic feasible solution to LPP is

$A_1 = b_1$, $A_2 = b_2$, ..., $A_m = b_m$

Now to obtain an optimal solution, we must drive out the artificial variables. The two methods to solve linear programming problems in such cases are:

- Two Phase method
- Big-M- method

LIMITATIONS OF SIMPLEX METHOD:

- Inability to deal with multiple objectives.
- Inability to handle problems with integer variables.
- Solution methods to LP problems with integer or Boolean variables are still far less efficient than those which include continuous variables only.

PROBLEMS

Example:

Maximize $z = 3x_1 + 2x_2$

Subject to
$-x_1 + 2x_2 \leq 4$
$3x_1 + 2x2 \leq 14$
$x_1 – x_2 \leq 3$
$x_1, x_2 \geq 0$

Solution:

First, convert every inequality constraints in the LPP into an equality constraint, so that the problem can be written in a standard from. This can be accomplished by adding a slack variable to each constraint. Slack variables are always added to the less than type constraints.

Converting inequalities to equalities

$-x_1 + 2x_2 + x_3 = 4$
$3x_1 + 2x_2 + x_4 = 14$
$x_1 – x_2 + x_5 = 3$
$x_1, x_2, x_3, x_4, x_5 \geq 0$
where x_3, x_4 and x_5 are slack variables.
Since slack variables represent unused resources, their contribution in the objective function is zero. Including these slack variables in the objective function, we get

Maximize $z = 3x_1 + 2x_2 + 0x_3 + 0x_4 + 0x_5$

Initial basic feasible solution

Now we assume that nothing can be produced. Therefore, the values of the decision variables are zero. $x_1 = 0$, $x_2 = 0$, $z = 0$

When we are not producing anything, obviously we are left with unused capacity

$$x_3 = 4, x_4 = 14, x_5 = 3$$

We note that the current solution has three variables (slack variables x_3, x_4 and x_5) with non-zero solution values and two variables (decision variables x_1 and x_2) with zero values.
Variables with non-zero values are called basic variables. Variables with zero values are called non-basic variables.

Iteration 1:

	c_j	3	2	0	0	0	
c_B	**Basic variables B**	x_1	x_2	x_3	x_4	x_5	**Solution values b (=X_B)**
0	x_3	-1	2	1	0	0	4
0	x_4	3	2	0	1	0	14
0	x_5	1	-1	0	0	1	3
z_j-c_j		-3	-2	0	0	0	

$a_{11} = -1$, $a_{12} = 2$, $a_{13} = 1$, $a_{14} = 0$, $a_{15} = 0$, $b_1 = 4$
$a_{21} = 3$, $a_{22} = 2$, $a_{23} = 0$, $a_{24} = 1$, $a_{25} = 0$, $b_2 = 14$
$a_{31} = 1$, $a_{32} = -1$, $a_{33} = 0$, $a_{34} = 0$, $a_{35} = 1$, $b_3 = 3$

Calculating values for the index row ($z_j – c_j$)

$z_1 – c_1 = (0 \times (-1) + 0 \times 3 + 0 \times 1) - 3 = -3$
$z_2 – c_2 = (0 \times 2 + 0 \times 2 + 0 \times (-1)) - 2 = -2$
$z_3 – c_3 = (0 \times 1 + 0 \times 0 + 0 \times 0) - 0 = 0$
$z_4 – c_4 = (0 \times 0 + 0 \times 1 + 0 \times 0) - 0 = 0$
$z_5 – c_5 = (0 \times 0 + 0 \times 0 + 0 \times 1) – 0 = 0$

Choose the smallest negative value from $z_j – c_j$ (i.e., – 3). So column under x_1 is the key column.
Now find out the minimum positive value
Minimum (14/3, 3/1) = 3,
So row x_5 is the key row.
Here, the pivot (key) element = 1 (the value at the point of intersection).Therefore, x_5 departs and x_1 enters.
We obtain the elements of the next table using the following rules:

1. If the values of $z_j – c_j$ are positive, the inclusion of any basic variable will not increase the value of the objective function. Hence, the present solution maximizes the objective function. If there are more than one negative values, we choose the variable as a basic variable corresponding to which the value of $z_j – c_j$ is least (most negative) as this will maximize the profit.
2. The numbers in the replacing row may be obtained by dividing the key row elements by the pivot element and the numbers in the other two rows may be calculated by using the formula:

$$\text{New number} = \text{old number} - \frac{(\text{correspon ding no. of key row}) \times (\text{correspon ding no. of key column})}{\text{pivot element}}$$

Calculating values for Iteration 2

x_3 row

$a_{11} = -1 – 1 \times ((-1)/1) = 0$
$a_{12} = 2 – (-1) \times ((-1)/1) = 1$

$a_{13} = 1 – 0 \text{ x } ((-1)/1) = 1$
$a_{14} = 0 – 0 \text{ x } ((-1)/1) = 0$
$a_{15} = 0 – 1 \text{ x } ((-1)/1) = 1$
$b_1 = 4 – 3 \text{ x } ((-1)/1) = 7$
x_4 row
$a_{21} = 3 – 1 \text{ x } (3/1) = 0$
$a_{22} = 2 – (-1) \text{ x } (3/1) = 5$
$a_{23} = 0 – 0 \text{ x } (3/1) = 0$
$a_{24} = 1 – 0 \text{ x } (3/1) = 1$
$a_{25} = 0 – 1 \text{ x } (3/1) = -3$
$b_2 = 14 – 3 \text{ x } (3/1) = 5$
x_1 row
$a_{31} = 1/1 = 1$
$a_{32} = -1/1 = -1$
$a_{33} = 0/1 = 0$
$a_{34} = 0/1 = 0$
$a_{35} = 1/1 = 1$
$b_3 = 3/1 = 3$
Iteration 2:

	c_j	**3**	**2**	**0**	**0**	**0**	
c_B	**Basic variables B**	x_1	x_2	x_3	x_4	x_5	**Solution values b (= X_B)**
0	x_3	0	1	1	0	1	7
0	x_4	0	5	0	1	-3	5
3	x_1	1	-1	0	0	1	3
z_j-c_j		0	-5	0	0	3	

Calculating values for the index row ($z_j – c_j$)
$z_1 – c_1 = (0 \text{ x } 0 + 0 \text{ x } 0 + 3 \text{ x } 1) - 3 = 0$
$z_2 – c_2 = (0 \text{ x } 1 + 0 \text{ x } 5 + 3 \text{ x } (-1)) – 2 = -5$
$z_3 – c_3 = (0 \text{ x } 1 + 0 \text{ x } 0 + 3 \text{ x } 0) - 0 = 0$
$z_4 – c_4 = (0 \text{ x } 0 + 0 \text{ x } 1 + 3 \text{ x } 0) - 0 = 0$
$z_5 – c_5 = (0 \text{ x } 1 + 0 \text{ x } (-3) + 3 \text{ x } 1) – 0 = 3$
Key column = x_2 column
Minimum (7/1, 5/5) = 1
Key row = x_4 row
Pivot element = 5
x_4 departs and x_2 enters.
Calculating values for table 3
x_3 row
$a_{11} = 0 – 0 \text{ x } (1/5) = 0$
$a_{12} = 1 – 5 \text{ x } (1/5) = 0$
$a_{13} = 1 – 0 \text{ x } (1/5) = 1$
$a_{14} = 0 – 1 \text{ x } (1/5) = -1/5$
$a_{15} = 1 – (-3) \text{ x } (1/5) = 8/5$
$b_1 = 7 – 5 \text{ x } (1/5) = 6$
x_2 row
$a_{21} = 0/5 = 0$
$a_{22} = 5/5 = 1$

$a_{23} = 0/5 = 0$
$a_{24} = 1/5$
$a_{25} = -3/5$
$b_2 = 5/5 = 1$
x_1 row
$a_{31} = 1 – 0 \text{ x } (-1/5) = 1$
$a_{32} = -1 – 5 \text{ x } (-1/5) = 0$
$a_{33} = 0 – 0 \text{ x } (-1/5) = 0$
$a_{34} = 0 – 1 \text{ x } (-1/5) = 1/5$
$a_{35} = 1 – (-3) \text{ x } (-1/5) = 2/5$
$b_3 = 3 – 5 \text{ x } (-1/5) = 4$
Note: Don't convert the fractions into decimals, because many fractions cancel out during the process while the conversion into decimals will cause unnecessary complications.
Final iteration:

	c_j	**3**	**2**	**0**	**0**	**0**	
c_B	**Basic variables B**	**x_1**	**x_2**	**x_3**	**x_4**	**x_5**	**Solution values b (= X_B)**
0	x_3	0	0	1	-1/5	8/5	6
2	x_2	0	1	0	1/5	-3/5	1
3	x_1	1	0	0	1/5	2/5	4
z_j-c_j		0	0	0	1	0	

Result:
Since all the values of zj – c_j are positive, this is the optimal solution**.**
$x_1 = 4$, $x_2 = 1$

Max z = 3 X $_4$ + 2 X $_1$ = 14.

The largest profit of Rs.14 is obtained, when 1 unit of x_2 and 4 units of x_1 are produced. The above solution also indicates that 6 units are still unutilized, as shown by the slack variable x_3 in the X_B column.

DEGENERACY
The problem of obtaining a degenerate basic feasible solution in a Linear programming problem is known as Degeneracy. Degeneracy in an L.P.P may arise

- At the initial stage
- At any Subsequent iteration stage.

In case (i) at least one basic variable is zero in the initial basic feasible solution, whereas in case (ii) at any iteration of the simplex method more than one vector is eligible to leave the basis, and hence the next simplex iteration produces a degenerate solution in which at least one basic variable is zero. This means that the subsequent iterations may not produce improvements in the value of the objective function. As a result it is possible to repeat the same sequence of simplex iterations endlessly without improving the solution. This concept is known as Cycling.
As you know, the simplex algorithm starts at a corner point and moves to an adjacent corner point by increasing the value of a non-basic variable x's with a negative value in the z-row (objective function).

Typically, the entering variable x’s does increase in value, and the objective value z improves. But it is possible that that x’s does not increase at all. It will happen when one of the RHS coefficients is 0.

In this case, the objective value and solution does not change, but there is an exiting variable. This situation is called degeneracy.

The Finiteness of the Simplex Algorithm when there is no degeneracy

Recall that the simplex algorithm tries to increase a non-basic variable x’s. If there is no degeneracy, then x’s will be positive after the pivot, and the objective value will improve. Recall also that each solution produced by the simplex algorithm’s a basic feasible solution with m basic variables, where m is the number of constraints. There are a finite number of ways of choosing the basic variables. (An upper bound is n! / (n-m)! m! , which is the number of ways of selecting m basic variables out of n.) So, the simplex algorithm moves from bfs to bfs. And it never repeats a bfs because the objective is constantly improving. This shows that the simplex method is finite, so long as there is no degeneracy.

2.3 Procedure to avoid cycling

A computation procedure to avoid cycling at any stage consists of the following steps

Step 1:

Let y_r enter the basis and suppose $\min\limits_i \left\{ \dfrac{x_{Bi}}{y_{ir}},\ y_{ir} > 0 \right\}$ is not unique.

Step 2:

Re arrange the column vectors of A, so that the starting initial basis B is chosen by selecting the first m column vectors of A. Then

$$y_j = B^{-1} a_j = e_j \ (j = 1,2\ldots,m)$$

Step 3:

Compute the non-negative ratios $\min\limits_i \left\{ \dfrac{y_{i1}}{y_{ir}},\ y_{ir} > 0 \right\}$ for those values of I

for which $\min\limits_i \left\{ \dfrac{x_{Bi}}{y_{ir}},\ y_{ir} > 0 \right\}$ have a tie.

If this minimum is unique for some I = k, then the vector y_k leaves the basis. Otherwise go to next step.

Step 4:

Compute the non-negative ratios $\min\limits_i \left. \dfrac{y_{i2}}{y_{ir}},\ y_{ir} > 0 \right\}$ for those values of i

for which $\min\limits_i \left\{ \dfrac{y_{i1}}{y_{ir}},\ y_{ir} > 0 \right\}$ have a tie.

If this minimum is unique for some i = k, then the vector y_k leaves the basis. Otherwise go to next step.

Step 5:

Compute the non-negative ratios $\min_i \left\{ \frac{y_{i3}}{y_{ir}},\ y_{ir} > 0 \right\}$ for those values of i

for which $\min_i \left\{ \frac{y_{i2}}{y_{ir}},\ y_{ir} > 0 \right\}$ have a tie.

If this minimum is unique for some I = k, then the vector y_k leaves the basis. Otherwise continue the procedure until a unique minimum nonnegative replacement ratio is obtained.

Note:

The above procedure is applicable to any iteration. However we have to maintain the same ordering of the column vectors in every iteration. Any artificial vector, if included, should also not be removed at any stage.

2.4 Degeneracy and the Simplex Algorithm

The simplex method without degeneracy	**The simplex method without degeneracy**
The solution changes after each pivot. The objective value strictly improves after a pivot.	The solution may stay the same after a pivot. The objective value may stay the same.
The Simplex method is guaranteed to be finite.	The Simplex method may cycle and be finite. But it becomes finite if we use the perturbation approach or several other approaches.
Two different tableaus in canonical form give two different solutions.	It is possible that there are two different sets of basic variables that give the same solution.

Degeneracy is important because we want the simplex method to be finite, and the generic simplex method is not finite if bases are permitted to be degenerate.

In principle, cycling can occur if there is degeneracy. In practice, cycling does not arise, but no one really knows why not. Perhaps it does occur, but people assume that the simplex algorithm is just taking too long for some other reason, and they never discover the cycling.

Researchers have developed several different approaches to ensure the finiteness of the simplex method, even if the bases can be degenerate. One such method is called the perturbation approach. The perturbation approach (in the form described here) is not practical, but it serves its purpose. It does give a way of doing simplex pivoting that is guaranteed to be finite.

Duality in Linear Programming

Associated with every L.P.P is always a corresponding L.P.P called the Dual problem of the given L.P.P. The original (given) L.P.P is called the Primal problem. However if we state the dual problem as the primal one, then the other can be considered to be the dual of this primal. The two problems can thus be said to constitute a pair of dual problems. Moreover as will be seen very soon, the two problems can be derived from each other and there is a unique dual (primal) problem associated with the primal(dual) problem. It will turn out that while solving a L.P.P. by simplex method, we shall simultaneously be solving its associated dual problem as well.

The following table gives the amounts of two vitamins v_1 and v_2 per unit , present in two different foods f_1 and f_2 respectively.

Vitamin	**Food**		**Daily requirements**
	f_1	f_2	
v_1	2	4	40

v_2	3	2	50
cost	3	2.5	

The last column of this table represents the number of units of the minimum daily requirements for the two vitamins; whereas the last row represents the cost per for the two foods. The problem is to determine the minimum quantities of the two foods f_1 and f_2 so that the minimum daily requirements of the two vitamins are met and that at the same time, the cost of purchasing these quantities of f_1 and f_2 is a minimum. To formulate the problem mathematically, let x_j be the number of units of food f_j,(j = 1,2) to be purchased, then the above problem is to determine two real numbers x_1 and x_2 so as to

Minimize $z = 3x_1 + 2.5\ x_2$

Subject to the constraints :

$$2x_1 + 4x_2 \geq 40$$
$$3x_1 + 2x_2 \geq 50$$
$$x_1, x_2 \geq 0$$

Here, of course, in the construction of constraints, we have assumed that any intake of more than the minimum requirements is not harmful and that negative purchase levels are prohibited. We shall consider this L.P.P. as the primal problem.

Now, let us consider a different problem, which is associated with above problem. Suppose there is a whole sale dealer who sells the two vitamins v_1 and v_2 along with some other commodities. The local shopkeepers purchase the vitamin from him and form the two foods f_1 and f_2 (the details are same as in the given table). The dealer knows very well that the foods f_1 and f_2 have their market values only because of their vitamin contents. The problem of the dealer is to fix the maximum per unit selling prices, for the two vitamins v_1 and v_2, in such a way that the resulting prices of foods f_1 and f_2 do not exceed their existing market prices.

To formulate the problem mathematically, let the dealer decide to fix up the two prices at w_1 and w_2 per unit, respectively. Then, the dealer’s problem can be started mathematically has to determine to real numbers w_1 and w_2 so as to

Maximize $z = 40w_1 + 50w_2$

Subject to the constraints:

$$2w_1 + 3w_2 \leq 3$$
$$4w_1 + 2w_2 \leq 2.5$$
$$w_1, w_2 \geq 0$$

Since negative price levels are prohibited.

Let us place the matrix formulation of this L.P.P. side by with that of the primal one :

Diet Problem

$$\text{Min } z = C^T X$$
$$A^T X \geq B^T$$
$$X \geq 0$$

Dealer's problem

$$\text{Max } z^* = BW$$
$$AW \leq C$$
$$W \geq 0$$

$$\text{where } c = (3 \quad 2.5),\ b = (40 \quad 50),\ A = \begin{pmatrix} 2 & 3 \\ 4 & 2 \end{pmatrix},\ X = (x_1 \quad x_2) \text{ and } W = (w_1 \quad w_2)$$

These two problems remarkably symmetric

- The cost vector associated with the objective function of one is just the right hand side vector in the other's set of constraints.
- The constraint coefficient matrix associated with one problem is simply the transpose of the constraint matrix associated with the other.

However the two problems differ in one respect, that one of the problem is a maximization problem while the other is a minimization problem.

3.2 Duality theorems:

Theorem 1:

The dual of the dual is the primal.

Theorem 2: (Weak Duality Theorem)

Let x_0 be a feasible solution to the primal problem

$$\text{Max } f(x) = cx \quad \text{subject to } Ax \leq b, x \geq 0$$
$$\text{where } x^T \text{ and } c \in R^n,\ b^T \in R^m \text{ and A is an } m \times n \text{ real matrix}$$

If w_0 be a feasible solution to the dual of the primal, namely

$$\text{Minimize} \quad g(w) = b^T w \quad \text{subject to } A^T w \geq c^T, w \geq 0$$
$$\text{where } w^T \in R^m \text{ then}$$
$$cx_0 \leq b^T w$$

Theorem 3: (Fundamental theorem of Duality)

If the primal or the dual has the finite optimum solution, then the other problem also possesses a finite optimum solution and the optimum values of the objective functions of the two problems are equal.

Theorem 4: (Existence theorem)

If either the primal or the dual problem has an unbounded objective function value, then the other problem has no feasible solution.

3.3 Relationship between primal and dual:

The symmetrical relationship between the primal and dual problem is summarized below (assume primal to be a minimization problem)

Primal	Dual
Minimization	Maximization
Number of variables	Number of constraints
Number of constraints	Number of variables

Less than or equal to type constraint	Non negative variable
Equal to type constraint	Unrestricted variable
Unrestricted variable	Equal to type constraint
R.H.S constant for the i^{th} constraint	Objective function coefficient for i^{th} variable
Objective function coefficient for j^{th} variable	R.H.S constant for the j^{th} constarint
Coefficient (a_{ij}) for i^{th} variable in j^{th} constraint	Coefficient (a_{ij}) for j^{th} variable in i^{th} constraint

Rules for Constructing the Dual from the Primal (or primal from the dual) are:

1. If the objective of one problem is to be maximized, the objective of the other is to be minimized.
2. The Maximization problem should have all constraints and the minimization problem has all constraints
3. All primal and dual variables must be non-negative (0)
4. The elements of right hand side of the constraints in one problem are the respective coefficients of the objective function in the other problem.
5. The matrix of constraints coefficients for one problem is the transpose of the matrix of constraint coefficient for the other problem.

Simplex method -Problems

A farmer has 1,000 acres of land on which he can grow corn, wheat or soyabeans. Each acre of corn costs Rs. 100 for preparation, requires 7 man-days of work and yields a profit of Rs. 30. An acre of wheat costs Rs. 120 to prepare, requires 10 man-days of work and yields a profit of Rs. 40. An acre of soyabeans costs Rs. 70 to prepare, requires 8 man-days of work and yields a profit of Rs. 20. The farmer has Rs. 1,00,000 for preparation and 8,000 man-days of work. Set-up the linear programming equation for the problem.

The given LPP is Max Z = 30x +40y +20z

Subject to

$$10x+12y+7z \le 10000$$
$$7x+10y+8z \le 8000$$
$$x+y+z \le 1000$$
$$x,y,z \ge 0$$

Step 1: Check up whether the objective function is of maximization type. Otherwise apply the equation Min(Z)=-Max(-Z) to convert the objective function to maximization type.

Step 2: Convert all the inequalities(≤ type) into equalities by introducing slack Variables. Then the problem becomes

$$\text{Max } Z = 30x+40y+20z+0.s_1+0.s_2+0.s_3$$

Subject to

$$10x+12y+7z+s_1 = 10000$$
$$7x+10y+8z+s_2 = 8000$$
$$x+y+z+s_3 = 1000$$
$$x, y, z, s_1, s_2, s_3 \ge 0$$

Step 3: Identify a suitable initial feasible solution by putting x=y=z=0.Then s_1=1000; s_2=8000 and s_3=1000 and these are the basic variables and the original variables x,y,z are non-basic variables. p

Step 4: Form the initial simplex table

			30	40	20	0	0	0
C_B	Y_B	X_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
0	Y_4	10000	10	12	7	1	0	0
0	Y_5	8000	7	10	8	0	1	0
0	Y_6	1000	1	1	1	0	0	1
		0	-30	-40	-20	0	0	0

Step 5: Compute the value of the objection by using C_BX_B. Also compute all the 'Net Evaluations' using

$$z_j - c_j = C_BY_j - c_j$$

and write them in the last row.

Step 6: If all the net evaluations are ≥ 0, the the optimal solution is reached. In the above example, the net evaluations corresponding to x1,x2 and x3 are negative. So the current solution is not optimal.

Step 7: Identify the most negative net evaluation and the variable corresponding to it will enter the basis. In the above problem, the most negative net evaluation is **-40** and it corresponds to the variable x_2 (or Y_2). So Y_2 enters into the basis.

			30	40	20	0	0	0	
C_B	Y_B	X_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	
0	Y_4	10000	10	12	7	1	0	0	10000/12=833…
0	Y_5	8000	7	**10**	8	0	1	0	**8000/10=800**
0	Y_6	1000	1	1	1	0	0	1	1000/1=1000
		0	-30	**-40**	-20	0	0	0	→

↑

←↓

Entering variable leaving variable

Step 8: To find the 'leaving variable', select the variable for which the ratio is a minimum (where j refers to the entering variable). So in the above example we have to find the minimum of the ratios (10000/12,8000/10,1000/1). The minimum ratio is 800 and it corresponds to the variable Y_5. So Y_5 leaves the basis. The row corresponding to Y_5 is called 'pivot row' and the column corresponding to Y_2 is called 'pivot column'. The element at the intersection of 'pivot row' and 'pivot column' is known as 'pivot element' and in the above example the pivot element is 10.

Step9: Form the next simplex table. Dividing all the pivot row elements by the pivot element.

			30	40	20	0	0	0	
C_B	Y_B	X_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	
0	Y_4	10000	10	12	7	1	0	0	10000/12=833…
40	Y_2	8000	7	**10**	8	0	1	0	**8000/10=800**
0	Y_6	1000	1	1	1	0	0	1	1000/1=1000
		0	-30	**-40**	-20	0	0	0	

Step 10: Convert all the other elements in the pivot column to 0 as follows:

Pivot Row: Row 2

800	7/10	**1**	8/10	0	1/10	0

Row 1:

10000	10	12	7	1	0	0
9600	84/10	12	96/10	0	12/10	0

Pivot Row X 12

(-)

400	16/10	0	-26/10	1	-12/10	0

Row 3:

1000	1	1	1	0	0	1
800	7/10	**1**	8/10	0	1/10	0

Pivot Row X 1

(-)

200	3/10	0	2/10	0	-10/10	1

Step 10: Form the next Simplex table:Repeat steps 7 to 9.

			30	40	20	0	0	0	
C_B	Y_B	X_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	
0	Y_4	400	**16/10**	0	-26/10	1	-12/10	0	**250**
40	Y_2	800	7/10	1	8/10	0	1/10	0	8000/7
0	Y_6	200	3/10	0	2/10	0	-1/10	1	2000/3
		32000	-2	0	12	0	4	0	

↑

Entering variable

Y_1 enters into the basis and Y_4 leaves the basis.

			30	40	20	0	0	0
C_B	Y_B	X_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
30	Y_1	250	**1**	0	-26/16	1	-12/16	0
40	Y_2	800	7/10	1	8/10	0	1/10	0
0	Y_6	200	3/10	0	2/10	0	-1/10	1
		32000						

Pivot Row:

250	**1**	0	-26/16	0	-12/16	0

Row: 2

800	7/10	1	8/10	0	1/10	0
175	7/10	0	-182/160	0	-84/160	0

Pivot Row X 7/10

(-)

625	0	1	31/16	0	10/16	0

Row 3:

200	3/10	0	2/10	0	-1/10	1
75	3/10	0	-78/160	0	-36/160	0

Pivot row X 3/10

(-)

125	0	0	11/16	0	20/160	1

Next Simplex Table:

			30	40	20	0	0	0	
C_B	Y_B	X_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	
30	Y_1	250	**1**	0	-26/16	1	-12/16	0	
40	Y_2	625	0	1	31/16	0	10/16	0	
0	Y_6	125	0	0	11/16	0	20/160	1	
		32500	**0**	**0**	**120/16**	**0**	**40/16**	**0**	

All the net evaluations are ≥ 0. So optimal solution is reached. So

Optimal Area under Crop 1 = 250 acres;
Optimal Area Under crop 2 = 625 acres
Optimal Area Under crop 3 = 0 acres
Maximum Profit: 32,500

Resource utilization

Resource	Crop 1	Crop 2	Crop 3	Total Resource Used	Total Resource available	Slack
Optimal areas	250	625	0	--	--	--
Land	1	1	1	875	1000	125
Capital	10	12	7	10000	10000	0
Labour	7	10	8	8000	8000	0

ARTIFICIAL VARIABLE TECHNIQUES

LPP in which constraints may also have ≥ and = signs after ensuring that at all b 0 i ≥ are considered in this section. In such cases basis of matrix cannot be obtained as an identity matrix in the starting simplex table, therefore we introduce a new type of variable called the artificial variable. These variables are fictitious and cannot have any physical meaning. The artificial variable technique is a device to get the starting basic feasible solution, so that simplex procedure may be adopted as usual until the optimal solution is obtained. To solve such LPP there are two methods.

1. The Big *M* Method or Method of Penalties.
2. The Two-phase Simplex Method.

BIG M METHOD.

The Big-M-Method is an alternative method of solving a linear programming problem involving artificial variables. To solve a L.P.P by simplex method, we have to start with the initial basic feasible solution and construct the initial simplex table. In the previous problems we see that the slack variables readily provided the initial basic feasible solution. However, in some problems, the slack variables cannot provide the initial basic feasible solution. In these problems atleast one of the constraint is of = or ≥ type. "Big-M-Method is used to solve such L.P.P.

ALGORITHM

The Big M-method

The big *M*-method or the method of penalties consists of the following basic steps :

Step 1:

Express the linear programming problem in the standard form by introducing slack and/or surplus variables, if any.

Step 2:

Introduce non-negative variables to the left hand side of all the constraints of (> or =) type. These variables are called artificial variables. The purpose of introducing artificial variables is just to obtain an initial basic feasible solution. However, addition of these artificial variables causes violation of the corresponding constraints. Therefore we would like to get rid of these variables and would not allow them to appear in the optimum simplex table. To achieve this, we assign a very large penalty ' — *M* ' to these artificial variables in the objective function, for maximization objective function.

Step 3:

Solve the modified linear programming problem by simplex method.

At any iteration of the usual simplex method there can arise any one of the following three cases :

(a) There is no vector corresponding to some artificial variable, in the basis y_b.

In such a, case, we proceed to *step 4.*

(b) There is at least one vector corresponding to some artificial variable, in the basis y_B, at the zero level. That is, the corresponding entry in X_B is zero. Also, the co-efficient of M in each net evaluation $Z_j - C_j (j = 1, 2, ... n)$ is non- negative.

In such a case, the current basic feasible solution is a degenerate one. This is a case when an optimum solution to the given L.P.P. includes an artificial basic variable and an optimum basic feasible solution still exists.

(c) At least one artificial vector is in the basis y_B, but not at the zero level. That is, the corresponding entry in X_B is non-zero. Also coefficient of M in each net evaluation $Z_j - C_j$ is non-negative,

In this case, the given L.P.P. does not possess any feasible solution.

Step 4:

Application of simplex method is continued until either an optimum basic feasible solution is obtained or there is an indication of the existence of an unbounded solution to the given L.P.P.

Note. While applying simplex method, whenever a vector corresponding to some artificial variable happens to leave the basis, we drop that vector and omit all the entries corresponding to its column from the simplex table.

Example:

Maximize $z = x_1 + 5x_2$

Subject to

$3x_1 + 4x_2 \leq 6$

$x_1 + 3x_2 \leq 2$

$x_1, x_2 \geq 0$

Solution:

Converting inequalities to equalities

By introducing surplus variables, slack variables and artificial variables, the standard form of LPP becomes

Maximize $x_1 + 5x_2 + 0x_3 + 0x_4 - MA_1$

Subject to

$3x_1 + 4x_2 + x_3 = 6$

$x_1 + 3x_2 - x_4 + A_1 = 2$

$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, A_1 \geq 0$

where,

x_3 is a slack variable
x_4 is a surplus variable
A_1 is an artificial variable.
Initial basic feasible solution

$$x_1 = x_2 = x_4 = 0, A_1 = 2, x_3 = 6$$

Iteration 1:

	c_j	**1**	**5**	**0**	**0**	**–M**	
c_B	**Basic variables B**	x_1	x_2	x_3	x_4	A_1	**Solution values b (= X_B)**
0	x_3	3	4	1	0	0	6
–M	A_1	1	3	0	–1	1	2
$z_j–c_j$		–M–1	–3M–5	0	M	0	

Calculating values for index row ($z_j – c_j$)
$z_1 – c_1 = 0 \times 3 + (–M) \times 1 – 1 = –M – 1$
$z_2 – c_2 = 0 \times 4 + (–M) \times 3 – 5 = –3M – 5$
$z_3 – c_3 = 0 \times 1 + (–M) \times 0 – 0 = 0$
$z_4 – c_4 = 0 \times 0 + (–M) \times (–1) – 0 = M$
$z_5 – c_5 = 0 \times 0 + (–M) \times 1 – (–M) = 0$
As M is a large positive number, the coefficient of M in the $z_j–c_j$ row would decide the incoming basic variable.
As -3M < -M, x_2 becomes a basic variable in the next iteration.
Key column = x_2 column.
Minimum (6/4, 2/3) = 2/3
Key row = A1 row
Pivot element = 3.
A1 departs and x_2 enters.
Note: The iteration just completed, artificial variable A_1 was eliminated from the basis. The new solution is shown in the following table.
Iteration 2:

	c_j	**1**	**5**	**0**	**0**	
c_B	**Basic variables B**	x_1	x_2	x_3	x_4	**Solution values b (= X_B)**
0	x_3	5/3	0	1	4/3	10/3
5	x_2	1/3	1	0	–1/3	2/3
$z_j–c_j$		2/3	0	0	–5/3	

Iteration 3:

	c_j	**1**	**5**	**0**	**0**	
c_B	**Basic variables B**	x_1	x_2	x_3	x_4	**Solution values b (= X_B)**
0	x_4	5/4	0	3/4	1	5/2
5	x_2	3/4	1	1/4	0	3/2
$z_j–c_j$		11/4	0	5/4	0	

Result:

The optimal solution is $x_1 = 0$, $x_2 = 3/2$

Max z = 0 + 5 x 3/2 =15/2

SAMPLEX PROBLEMS

1. Use penalty (or Big 'M') method to

Minimize $z = 4x_i + 3x_2$

subject to the constraints :

$2x_1 + x_2 \geq 10$, $-3x_1, + 2x_2 \leq 6$

$x_1 + x_2 \geq 6$, $x_1 \geq 0$ and $x_2 \geq 0$.

Solution. Introducing surplus (negative slack) variables $x_3 \geq 0$, $x_5 \geq 0$ and slack variable $x_4 \geq 0$ in the constraint inequations, the problem becomes

Maximize $z^* = -4x_1 - 3x_2 + 0.x_3 + 0.x_4 + 0.x_5$

subject to the constraints :

$2x_1 + x_2 - x_3 = 10$, $-3x_1 + 2x_2 + x_4 = 6$

$x_l + x_2 - x_5 = 6$, $x_j \geq 0$ Q(j = 1,2, 3, 4, 5)

Clearly, we do not have a ready basic feasible solution. The surplus variables carry negative coefficients (-1). We introduce two new variables $A_1 \geq 0$ and $A_2 \geq 0$ in the first and third equations respectively. These extraneous variables, commonly termed as artificial variables, play the same role as that of slack variables in providing a starting basic feasible solution.

We assign a very high penalty cost (say $-M$, $M \geq 0$) to these variables in the objective function so that they may be driven to zero while reaching optimality.

Now the following initial basic feasible solution is available :

$A_l = 10$, $x_4 = 6$ and $A_2 = 6$

with $\mathbf{B} = (\mathbf{a_6}, \mathbf{a_4}, \mathbf{a_7})$ as the basis matrix. The cost matrix corresponding to basic feasible solution is $\mathbf{c_B} = (-M, 0, -M)$

Now, corresponding to the basic variables A_1, x_4 and A_2. the matrix $\mathbf{Y = B^{-1}A}$ and the net evaluations $z_j - c_j$ (j = 1, 2, ... 7) are computed. The initial basic feasible solution is displayed in the following simplex table :

Initial Iteration.			Introduce y_1 and drop y_6. —4	—3	0	0	0	-M	-M
c_B	y_B	x_B	y_1	y_2	y_3	y_4	y_5	y_6	y_7
$-M$	y_6	10	2*	1	-1	0	0	1	0
0	y_4	6	-3	2	0	1	0	0	0
$-M$	y_7	6	1	1	0	0	- 1	0	1
	z*	-16M	-3W+4	-2M+3	M	0	M	0	0

We observe that the most negative zj - cj is 4 - 3M (= $z_l - c_1$). The corresponding column vector y_1, therefore, enters the basis. Further, since min. = 5; the element y_{11} (=2) becomes the leading element for the first iteration

First Iteration: Introduce y_2 and drop y_7.

c_B	y_B	x_B	y_1	y_2	y_3	y_4	y_5	y_7
-4	y_1	5	1	1/2	-1/2	0	0	0
0	y_4	21	0	7/2	-3/2	1	0	0
-M	y_7	0	0	1/2*	1/2	0	-1	0
	z*	-M-20	0	$\frac{-M}{2}+1$	$\frac{-M}{2}+2$	0	M	0

In the above table, we omitted all entries of column vector $\mathbf{y_6}$, because the artificial variables A_l has left the basis and we would not like it to re-enter in any subsequent iterations.
Now since the most negative (z_j—c_j) is z_2-c_2; the non-basic vector $\mathbf{y_2}$ enters the basis. Further, since min is 2 which occurs for the element y_{32} (= 1/2), the corresponding basis vector $\mathbf{y_7}$ leaves the basis and the element y_{32} becomes the leading element for the next iteration.

Final Iteration: Optimum Solution,

c_B	y_B	x_B	y_1	y_2	y_3	y_4	y_5
-4	y_1	4	1	0	-1	0	1
0	y_4	14	0	0	-5	1	7
-3	y_2	2	0	1	1	0	-2
	z*	-22	0	0	1	0	2

It is clear from the table that all z_j - c_j are positive. Therefore an optimum basic feasible solution has been attained which is given by

x_1= 4, x_2 = 2, maximum z = 22.

2. Maximize z = $3x_1$, + $2x_2$ subject to the constraints :

$2x_1 + x_2 \leq 2$, $3x_1 + 4x_2 \geq 12$, $x_1, x_2 \geq 0$.

Solution:

Introducing slack variable $x_3 \geq 0$, surplus variable $x_5 \geq 0$ and an artificial variable $A_1 \geq 0$, the reformulated L.P.P. can be written as :

Maximize z = $3x_1 + 2x_2 + 0.x_3 + 0.x_4 – MA_1$

subject to the constraints :

$2x_1 + x_2 + x_3 = 2$,

$3x_1 + 4x_2 - x_4 + A_1 = 12$

$x_1, x_2, x_3, x_4 \geq 0$ and $A_1 \geq 0$.

An obvious starting basic feasible solution is :

x3 = 2 and A_1 = 12.

The iterative simplex tables are :

Initial Iteration: Introduce y_2 and drop y_3.

			3	2	0	0	-M
c_B	y_B	x_B	y_1	y_2	y_3	y_4	y_5
0	y_3	2	2	1*	1	0	0
-M	y_5	12	3	4	0	-1	1
	z	-12M-2	-3M-3	-4M	0	M	0

Final Iteration. No solution.

c_B	y_B	x_B	y_1	y_2	y_3	y_4	y_5
2	y_2	2	2	1	1	0	0
-M	y_5	4	-5	0	-4	-1	1
	z	4M+4	5M+1	0	4M + 2	M	0

Here the coefficient of *M* in each $z_j - c_j$ is non-negative and an artificial vector appears in the basis, not at the zero level. Thus the given L.P.P. does not possess any feasible solution.

PRACTICE

EXAMPLE:1

A company makes two products (X and Y) using two machines (A and B). Each unit of X that is produced requires 50 minutes processing time on machine A and 30 minutes processing time on machine B. Each unit of Y that is produced requires 24 minutes processing time on machine A and 33 minutes processing time on machine B.

At the start of the current week there are 30 units of X and 90 units of Y in stock. Available processing time on machine A is forecast to be 40 hours and on machine B is forecast to be 35 hours.

The demand for X in the current week is forecast to be 75 units and for Y is forecast to be 95 units. Company policy is to maximise the combined sum of the units of X and the units of Y in stock at the end of the week.

- Formulate the problem of deciding how much of each product to make in the current week as a linear program.
- Solve this linear program graphically.

Solution

Let

- x be the number of units of X produced in the current week
- y be the number of units of Y produced in the current week

then the constraints are:

50x + 24y <= 40(60) machine A time

30x + 33y <= 35(60) machine B time

x >= 75 - 30

i.e. x >= 45 so production of X >= demand (75) - initial stock (30), which ensures we meet demand

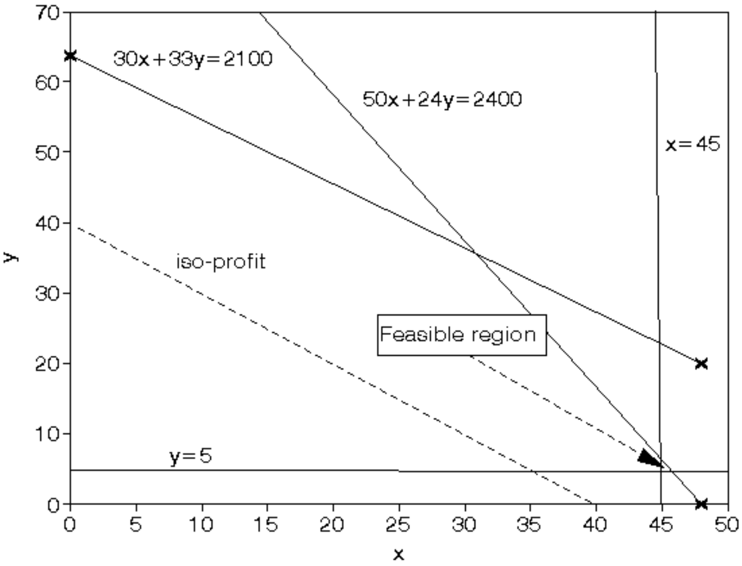
y >= 95 - 90

i.e. y >= 5 so production of Y >= demand (95) - initial stock (90), which ensures we meet demand

The objective is: maximise (x+30-75) + (y+90-95) = (x+y-50) i.e. to maximise the number of units left in stock at the end of the week

It is plain from the diagram below that the maximum occurs at the intersection of x=45 and 50x + 24y = 2400

Solving simultaneously, rather than by reading values off the graph, we have that x=45 and y=6.25 with the value of the objective function being 1.25



EXAMPLE:2

The demand for two products in each of the last four weeks is shown below.

	Week			
	1	2	3	4
Demand - product 1	23	27	34	40
Demand - product 2	11	13	15	14

Apply exponential smoothing with a smoothing constant of 0.7 to generate a forecast for the demand for these products in week 5.

These products are produced using two machines, X and Y. Each unit of product 1 that is produced requires 15 minutes processing on machine X and 25 minutes processing on machine Y. Each unit of product 2 that is produced requires 7 minutes processing on machine X and 45 minutes processing on machine Y. The available time on machine X in week 5 is forecast to be 20 hours and on machine Y in week 5 is forecast to be 15 hours. Each unit of product 1 sold in week 5 gives a contribution to profit of £10 and each unit of product 2 sold in week 5 gives a contribution to profit of £4.

It may not be possible to produce enough to meet your forecast demand for these products in week 5 and each unit of unsatisfied demand for product 1 costs £3, each unit of unsatisfied demand for product 2 costs £1.

- Formulate the problem of deciding how much of each product to make in week 5 as a linear program.
- Solve this linear program graphically.

Solution

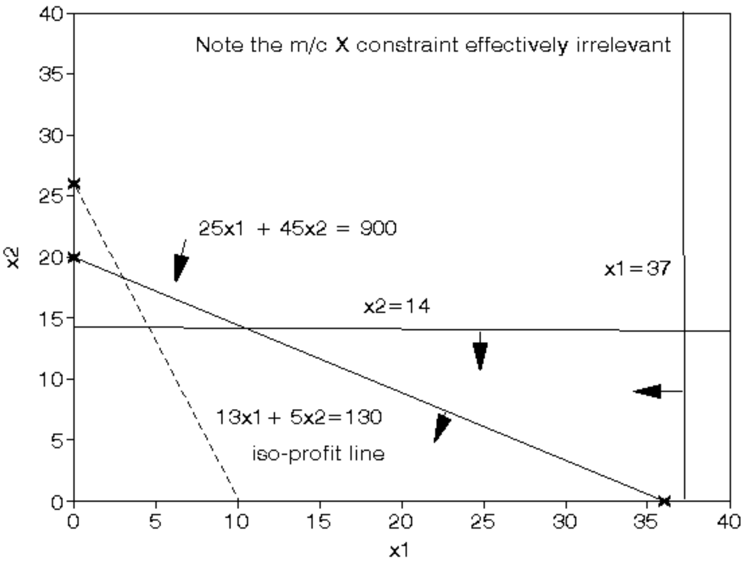
Note that the first part of the question is a forecasting question so it is solved below.

For product 1 applying exponential smoothing with a smoothing constant of 0.7 we get:

$M_1 = Y_1 = 23$

$M_2 = 0.7Y_2 + 0.3M_1 = 0.7(27) + 0.3(23) = 25.80$

$M_3 = 0.7Y_3 + 0.3M_2 = 0.7(34) + 0.3(25.80) = 31.54$
$M_4 = 0.7Y_4 + 0.3M_3 = 0.7(40) + 0.3(31.54) = 37.46$
The forecast for week five is just the average for week 4 = M_4 = 37.46 = 31 (as we cannot have fractional demand).



For product 2 applying exponential smoothing with a smoothing constant of 0.7 we get:
$M_1 = Y_1 = 11$
$M_2 = 0.7Y_2 + 0.3M_1 = 0.7(13) + 0.3(11) = 12.40$
$M_3 = 0.7Y_3 + 0.3M_2 = 0.7(15) + 0.3(12.40) = 14.22$
$M_4 = 0.7Y_4 + 0.3M_3 = 0.7(14) + 0.3(14.22) = 14.07$
The forecast for week five is just the average for week 4 = M_4 = 14.07 = 14 (as we cannot have fractional demand).
We can now formulate the LP for week 5 using the two demand figures (37 for product 1 and 14 for product 2) derived above.
Let
x_1 be the number of units of product 1 produced
x_2 be the number of units of product 2 produced
where $x_1, x_2 >= 0$
The constraints are:
$15x_1 + 7x_2 <= 20(60)$ machine X
$25x_1 + 45x_2 <= 15(60)$ machine Y
$x_1 <= 37$ demand for product 1
$x_2 <= 14$ demand for product 2
The objective is to maximise profit, i.e.
maximise $10x_1 + 4x_2 - 3(37- x_1) - 1(14-x_2)$
i.e. maximise $13x_1 + 5x_2 - 125$
The graph is shown below, from the graph we have that the solution occurs on the horizontal axis (x_2=0) at x_1=36 at which point the maximum profit is 13(36) + 5(0) - 125 = £343

EXAMPLE:3

A company is involved in the production of two items (X and Y). The resources need to produce X and Y are twofold, namely machine time for automatic processing and craftsman time for hand finishing. The table below gives the number of minutes required for each item:

	Machine time	Craftsman time
Item X	13	20
Y	19	29

The company has 40 hours of machine time available in the next working week but only 35 hours of craftsman time. Machine time is costed at £10 per hour worked and craftsman time is costed at £2 per hour worked. Both machine and craftsman idle times incur no costs. The revenue received for each item produced (all production is sold) is £20 for X and £30 for Y.

The company has a specific contract to produce 10 items of X per week for a particular customer.

- Formulate the problem of deciding how much to produce per week as a linear program.
- Solve this linear program graphically.

Solution

Let

- x be the number of items of X
- y be the number of items of Y

then the LP is:

maximise

- 20x + 30y - 10(machine time worked) - 2(craftsman time worked)

subject to:

- 13x + 19y <= 40(60) machine time
- 20x + 29y <= 35(60) craftsman time
- x >= 10 contract
- x,y >= 0

so that the objective function becomes

maximise

- 20x + 30y - 10(13x + 19y)/60 - 2(20x + 29y)/60

i.e. maximise

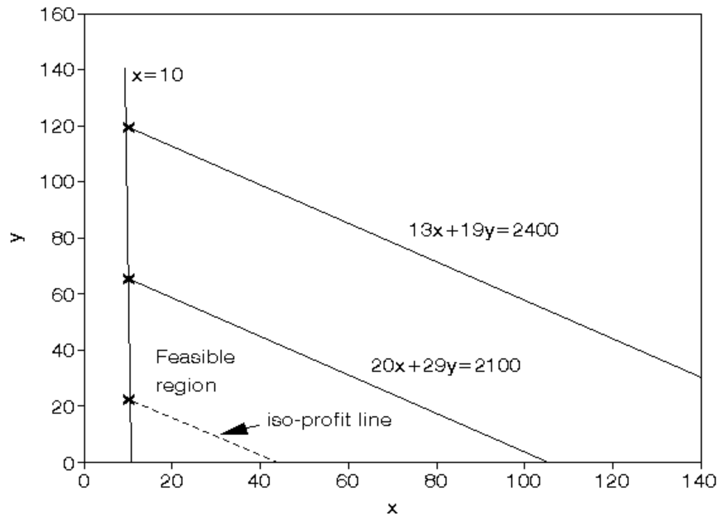
- 17.1667x + 25.8667y

subject to:

- 13x + 19y <= 2400
- 20x + 29y <= 2100
- x >= 10
- x,y >= 0

It is plain from the diagram below that the maximum occurs at the intersection of x=10 and 20x + 29y <= 2100

Solving simultaneously, rather than by reading values off the graph, we have that x=10 and y=65.52 with the value of the objective function being £1866.5



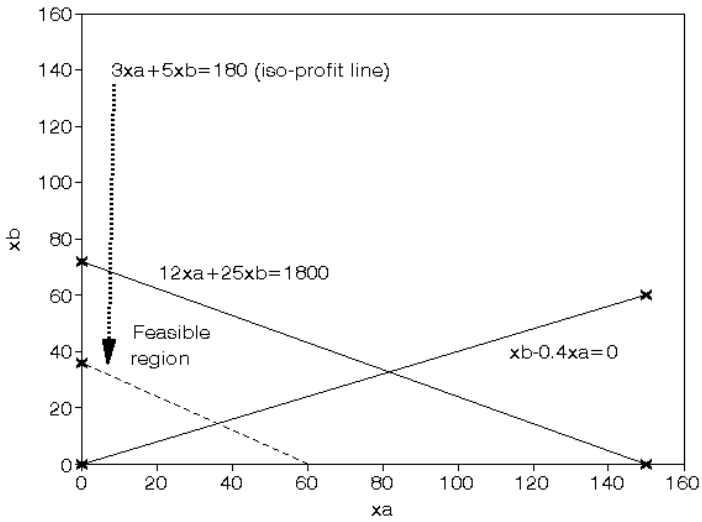
EXAMPLE:4

A company manufactures two products (A and B) and the profit per unit sold is £3 and £5 respectively. Each product must be assembled on a machine, each unit of product A taking 12 minutes of assembly time and each unit of product B 25 minutes of assembly time. The company estimates that the machine used for assembly has an effective working week of only 30 hours (due to maintenance/breakdown).

Technological constraints mean that for every five units of product A produced at least two units of product B must be produced.

- Formulate the problem of how much of each product to produce as a linear program.
- Solve this linear program graphically.
- The company has been offered the chance to hire an extra machine, thereby doubling the effective assembly time available. What is the *maximum* amount you would be prepared to pay (per week) for the hire of this machine and why?

Solution
Let
x_A = number of units of A produced
x_B = number of units of B produced
then the constraints are:
$12x_A + 25x_B <= 30(60)$ (assembly time)
$x_B >= 2(x_A/5)$
i.e. $x_B - 0.4x_A >= 0$
i.e. $5x_B >= 2x_A$ (technological)
where $x_A, x_B >= 0$
and the objective is
maximise $3x_A + 5x_B$
It is plain from the diagram below that the maximum occurs at the intersection of $12x_A + 25x_B = 1800$ and $x_B - 0.4x_A = 0$



Solving simultaneously, rather than by reading values off the graph, we have that:
$x_A= (1800/22) = 81.8$
$x_B= 0.4x_A = 32.7$
with the value of the objective function being £408.9
Doubling the assembly time available means that the assembly time constraint (currently $12x_A + 25x_B <= 1800$) becomes $12x_A + 25x_B <= 2(1800)$ This new constraint will be parallel to the existing assembly time constraint so that the new optimal solution will lie at the intersection of $12x_A + 25x_B = 3600$ and $x_B - 0.4x_A = 0$
i.e. at $x_A = (3600/22) = 163.6$
$x_B= 0.4x_A = 65.4$
with the value of the objective function being £817.8
Hence we have made an additional profit of £(817.8-408.9) = £408.9 and this is the *maximum* amount we would be prepared to pay for the hire of the machine for doubling the assembly time.
This is because if we pay more than this amount then we will reduce our maximum profit below the £408.9 we would have made without the new machine.

EXAMPLE:5
Solve
minimise
$$4a + 5b + 6c$$
subject to
$$a + b >= 11$$
$$a - b <= 5$$
$$c - a - b = 0$$
$$7a >= 35 - 12b$$

a >= 0 b >= 0 c >= 0

Solution

To solve this LP we use the equation c-a-b=0 to put c=a+b (>= 0 as a >= 0 and b >= 0) and so the LP is reduced to

minimise

$4a + 5b + 6(a + b) = 10a + 11b$

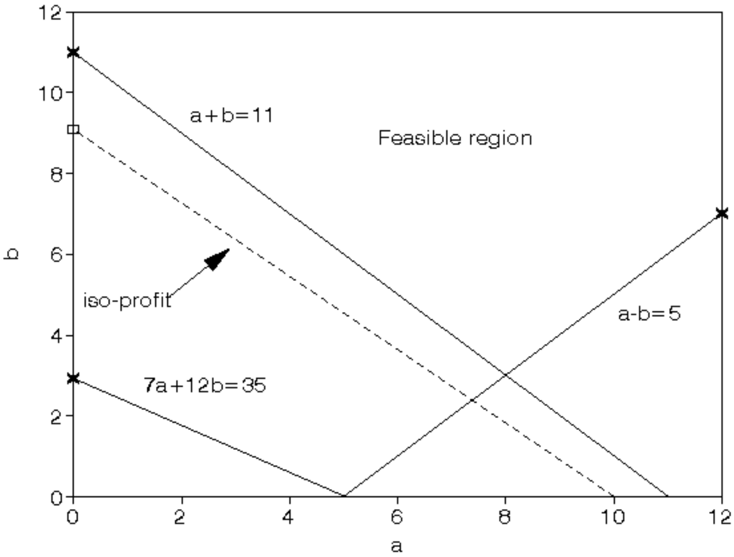
subject to

$a + b >= 11$

$a - b <= 5$

$7a + 12b >= 35$

$a >= 0$ $b >= 0$



From the diagram below the minimum occurs at the intersection of a - b = 5 and a + b = 11

i.e. a = 8 and b = 3 with c (= a + b) = 11 and the value of the objective function 10a + 11b = 80 + 33 = 113.

EXAMPLE: 6

Solve the following linear program:

maximise $5x_1 + 6x_2$

subject to

$x_1 + x_2 <= 10$

$x_1 - x_2 >= 3$

$5x_1 + 4x_2 <= 35$

$x_1 >= 0$

$x_2 >= 0$

Solution

It is plain from the diagram below that the maximum occurs at the intersection of

$5x_1 + 4x_2 = 35$ and

$x_1 - x_2 = 3$

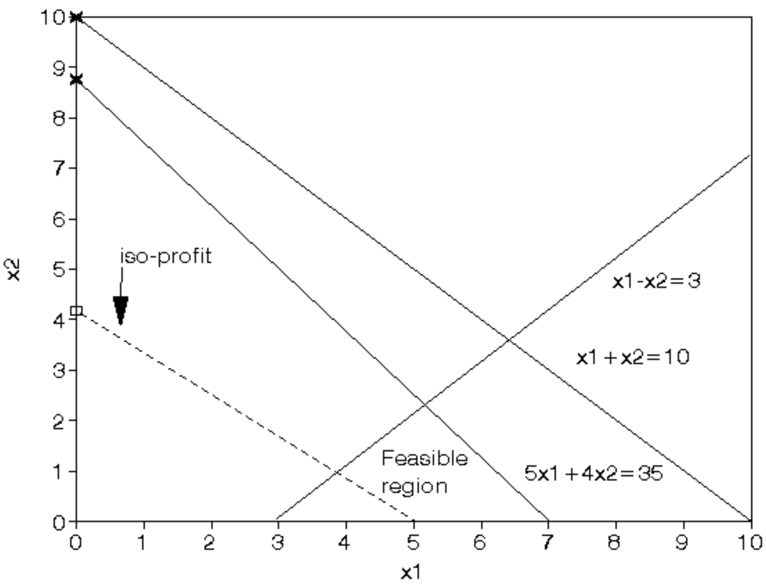
Solving simultaneously, rather than by reading values off the graph, we have that

$5(3 + x_2) + 4x_2 = 35$

i.e. $15 + 9x_2 = 35$

i.e. $x_2 = (20/9) = 2.222$ and

$x_1 = 3 + x_2 = (47/9) = 5.222$



The maximum value is 5(47/9) + 6(20/9) = (355/9) = 39.444

EXAMPLE:7

A carpenter makes tables and chairs. Each table can be sold for a profit of £30 and each chair for a profit of £10. The carpenter can afford to spend up to 40 hours per week working and takes six hours to make a table and three hours to make a chair. Customer demand requires that

he makes at least three times as many chairs as tables. Tables take up four times as much storage space as chairs and there is room for at most four tables each week.
Formulate this problem as a linear programming problem and solve it graphically.

Solution

Variables

Let

x_T = number of tables made per week

x_C = number of chairs made per week

Constraints

- total work time

$6x_T + 3x_C <= 40$

- customer demand

$x_C >= 3x_T$

- storage space

$(x_C/4) + x_T <= 4$

- all variables >= 0

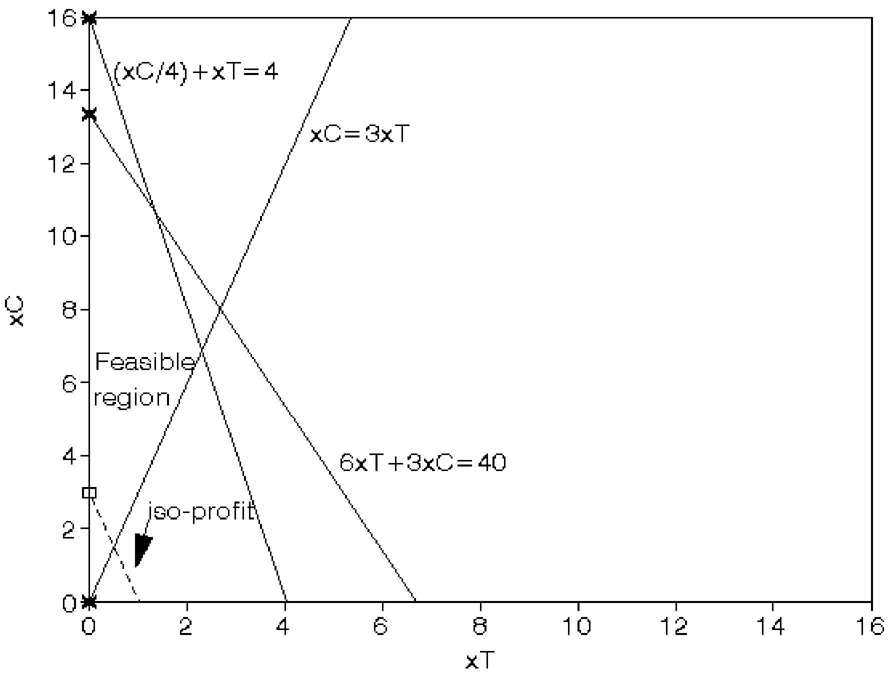
Objective

maximise $30x_T + 10x_C$

The graphical representation of the problem is given below and from that we have that the solution lies at the intersection of

$(x_C/4) + x_T = 4$ and $6x_T + 3x_C = 40$

Solving these two equations simultaneously we get $x_C = 10.667$, $x_T = 1.333$ and the corresponding profit = £146.667



UNIT II
TRANSPORTATION & ASSIGNMENT MODELS

Learning Objectives

Transportation, Assignment Models: Definition and application of the transportation model, methods for finding initial solution-tests for optimality-variations in transportation problem, the Assignment Model, Travelling Salesman Problem.

TRANSPORTATION

Introduction

Industries require to transport their products available at several sources or production centres to a number of destinations or markets. In the process of distributing to various destinations, high transportation costs are involved. Minimizing the transportation cost will benefit the organisation by increasing the profit. To analyze and minimize the cost of transportation, transportation model is used. The name "transportation model" is, however, misleading. This model can be used for a wide variety of situations such as scheduling, personnel assignment, product mix problems and many others, so that the model is really not confined to transportation or distribution only.

The origin of transportation models dates back to 1941 when F.L. Hitchcock presented a study entitled 'The Distribution of a Product from Several Sources to Numerous Localities'. The presentation is regarded as the first important contribution to the solution of transportation problems. In 1947, T.C. Koopmans presented a study called 'Optimum Utilization of the Transportation System'. These two contributions are mainly responsible for the development of transportation models which involve a number of production centres / sources and a number of destinations / markets. Each shipping source has a certain capacity and each destination has a certain requirement associated with a certain cost of transportation from the sources to the destinations. The objective is to minimize the cost of transportation while meeting the requirements at the destinations. Transportation problems may also involve movement of a product from plants to warehouses, warehouses to wholesalers, wholesalers to retailers, retailers to customers, etc.

2. Assumptions in the Transportation Model

1. Total quantity of the items available at different sources/ supply is equal to the total requirement/ demand at different destinations / markets.
2. Items can be transported conveniently from all sources to destinations.
3. The unit transportation cost of the item from all sources to destinations is known.
4. The transportation cost on a given route is directly proportional to the number of units shipped on that route.
5. The objective is to minimize the total transportation cost for the organization as a whole and not for individual supply and distribution centres.

3. Definition of the Transportation Model

Suppose that there are m sources and n destinations. Let a_i be the number of supply units available at source i(i = 1,2,3,….., m) and let b_j be the number of demand units required at destination j(j = 1,2,3,….., n). Let c_{ij} represent the unit transportation cost for transporting the units from source i to destination j. The objective is to determine the number of units to be transported from source *i* to destination *j* so that the total transportation cost is minimum. In addition, the supply limits at the sources and the demand requirements at the destinations must be satisfied exactly.

If x_{ij} ($x_{ij} \geq 0$) is the number of units shipped from source i to destination j, then the equivalent linear programming model will be

Find x_{ij} (i =1,2,3, ……… , m ; j = 1,2,3,……… , n) in order to

Destinations

		1	2	3	… j …	n	Supply
Sources or Origins	1	C_{11} x_{11}	C_{12} x_{12}	C_{13} x_{13}	C_{1j} x_{1j}	C_{1n} x_{1n}	a_1
	2	C_{21} x_{21}	C_{22} x_{22}	C_{23} x_{23}	C_{2j} x_{2j}	C_{2n} x_{2n}	a_2
	3	C_{31} x_{31}	C_{32} x_{32}	C_{33} x_{33}	C_{3j} x_{3j}	C_{3n} x_{3n}	a_3
	:	C_{i1} x_{i1}	C_{i2} x_{i2}	C_{i3} x_{i3}	C_{ij} x_{ij}	C_{in} x_{in}	: a_i :
	m	C_{m1} x_{m1}	C_{m2} x_{m2}	C_{m3} x_{m3}	C_{mj} x_{mj}	C_{mn} x_{mn}	a_m
Demand		b_1	b_2	b_3	$\ldots b_j \ldots$	b_n	

minimize

$$z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}\, x_{ij}\ ,$$

subject to

$$\sum_{j=1}^{n} x_{ij} = a_i\ , \quad i = 1,2,3, \ldots\ldots , m\ ,$$

and

$$\sum_{i=1}^{m} x_{ij} = b_j\ , \quad j = 1,2,3, \ldots\ldots , n\ ,$$

where $x_{ij} \geq 0$

The two sets of constraints will be consistent i.e., the system will be in balance if

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j\ .$$

Equality sign of the constraints causes one of the constraints to be redundant (and hence it can be deleted) so that the problem will have (m + n - 1) constraints and (m x n) unknowns.

Note that a transportation problem will have a feasible solution only if the above restriction is satisfied. Thus,

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$$ is necessary as well as a sufficient condition for a

transportation problem to have a feasible solution. Problems that satisfy this condition are called balanced transportation problems. Techniques have been developed for solving balanced or standard transportation problems only. It follows that any non – standard problem in which the supplies and demands do not balance, must be converted to a standard transportation problem before it can be solved. This conversion can be achieved by the use of a dummy source/destination.

The above information can be put in the form of a general matrix shown below:

In table , c_{ij} , i = 1,2, ….., m ; j = 1,2, …… , n , is the unit shipping cost from the ith oringin to jth destination, x_{ij} is the quantity shipped from the ith origin to jth destination, a_i is the supply available at origin i and b_j is the demand at destination j.

Definitions:

A few terms used in connection with transportation models are defined below.

1. **Feasible solution:** A feasible solution to a transportation problem is a set of non-negative allocations, x_{ij} that satisfies the rim (row and column) restrictions.
2. **Basic feasible solution:** A feasible solution to a transportation problem is said to be a basic feasible solution if it contains no more than m + n – 1 non – negative allocations, where m is the number of rows and n is the number of columns of the transportation problem.
3. **Optimal solution:** A feasible solution (not necessarily basic) that minimizes (maximizes) the transportation cost (profit) is called an optimal solution.
4. **Non -degenerate basic feasible solution:** A basic feasible solution to a (m x n) transportation problem is said to be non – degenerate if,
    1. the total number of non-negative allocations is exactly m + n – 1 (i.e., number of independent constraint equations), and
    2. these m + n – 1 allocations are in independent positions.
5. **Degenerate basic feasible solution:** A basic feasible solution in which the total number of non-negative allocations is less than m + n – 1 is called degenerate basic feasible solution.

4. Matrix Terminology

The matrix used in the transportation models consists of squares called 'cells', which when stacked form 'columns' vertically and 'rows' horizontally.

The cell located at the intersection of a row and column is designated by its row and column headings. Thus the cell located at the intersection of row A and column 3 is called cell (A, 3). Unit costs are placed in each cell.

Warehouses

		1	2	3	4		
Plants	A	2	3	11	4	15	Output
	B	5	6	8	7	20	
Demand		10	5	12	8	35 (Total)	

5. Degeneracy in Transportation Problem

In case of simplex algorithm, the basic feasible solution may become degenerate at the initial stage or at some intermediate stage of computation. In a transportation problem with m origins and n destinations if a basic feasible solution has less than m + n – 1 allocations (occupied cells), the problem is said to be a degenerate transportation problem.

While in the simplex method degeneracy does not cause any serious difficulty, it can cause computational problem in transportation technique. In stepping – stone method it will not be possible to make close paths (loops) for each and every vacant cell and hence evaluations of all the vacant cells cannot be calculated. If modified distribution method is applied, it will not be possible to find all the dual variables u_i and v_j since the number of allocated cells and their c_{ij} values is not enough. It is thus necessary to identify a degenerate transportation problem and take appropriate steps to avoid computational difficulty. Degeneracy can occur in the initial solution or during some subsequent iteration.

5.1. Degeneracy in the initial solution

Normally, while finding the initial solution (by any of the methods), any allocation made either satisfies supply or demand, but not both. If, however, both supply and demand are satisfied simultaneously, a row as well as column are cancelled simultaneously and the number of allocations become two less than m + n – 1 and so on. This degeneracy is resolved or the above degenerate solution is made non-degenerate in the following manner:

First of all the requisite number of vacant cells with least unit costs are chosen so that (incase of tie choose arbitrarily):

1. these cells plus the existing number of allocations are equal to m + n – 1.
2. these m + n – 1 cells are in independent positions i.e., no closed path (loop) can be formed among them. If a loop is formed the cells / cells with next lower cost is/are chosen so that no loop is formed among them. This can always be done if the solution we start with contains allocated cells in independent positions.

Now allocate an infinitesimally small but positive value ε (Greek letter epsilon) to each of the chosen cells. Subscripts are used when more than one such letter is required (e.g., ε1, ε2, etc.) these ε's are then treated like any other positive basic variable and are kept in the transportation array (matrix) until temporary degeneracy is removed or until the optimal solution is reached, whichever occurs first. At that point we set each ε = 0. Notice that ε is infinitesimally small and hence its effect can be neglected when it is added to or subtracted from a positive value (e.g. 10 + ε = 10, 5 – ε = 5, ε + ε = 2 ε , ε - ε = 0). Consequently, they do not appreciably alter the physical nature of the original set of allocations but do help in carrying our further computations such as optimality test.

5.2. Degeneracy during some subsequent iteration

Sometimes even if the starting feasible solution is non-degenerate, degeneracy may develop later at some subsequent iteration. This happens when the selection of the entering variable (least value in the closed path that has been assigned a negative sign), causes two or more current basic variables (allocated cell values) to become zero. In thils case we allocate ε to recently vacated cell with least cost that there are exactly m + n – 1 allocated cells in independent positions and the procedure can then be continued in the usual manner.

6. Transportation Algorithm

Transportation algorithm for a minimization problem as discussed earlier can be summarized in the following steps:

1. Construct the transportation matrix. For this enter the supply a_i from the origins, demand b_j at the destinations and the unit costs c_{ij} in the various cells.
2. Find initial basic feasible solution by Vogel's approximation method or any of the other given methods.
3. Perform optimally test using modified distribution method. For this, find dual variables u_i and v_j such that $u_i + v_j = c_{ij}$ for occupied cells. Starting with say, $v_i = 0$, all other variables can be evaluated.

4. Compute the cell evaluations = $c_{ij} - (u_i + v_j)$ for vacant cells. If all cell evaluations are positive or zero, the current basic feasible solution is optimal. In case any cell evaluation in negative, the current solution is not optimal.
5. Select the vacant cell with the most negative evaluation. This is called identified cell.
6. Make as much allocation in the identified cell as possible so that it becomes basic i.e., Reallocate the maximum possible number of units to these cells, keeping in mind the rim conditions. This will make allocation in one basic cell zero and in other basic cells the allocations will remain non-negative (≥ 0). The basic cell whose allocation becomes zero will leave the basis.
7. Return to step 3, repeat the process till optimal solution is obtained.

7. Variants in Transportation Problems

The following variations in the transportation problem will now be considered:

1. Unbalanced transportation problem.
2. Maximization problem.
3. Different production costs.
4. No allocation in a particular cell/cells.
5. Overtime production.

7.1. The unbalanced transportation problem

In the problems discussed so far, the total availability from all the origins was equal to the total demand at all the destinations i.e.,

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$$. Such problems are called balanced transportation problems.

In many real life situations, however, the total availability may not be equal to the total demand. i.e.,

$$\sum_{i=1}^{m} a_i \neq \sum_{j=1}^{n} b_j$$; such problems are called unbalanced transportation problems.

In these problems either some available resources will remain unused or some requirements will remain unfilled.

Since a feasible solution exists only for a balanced problem, it is necessary that the total availability be made equal to the total demand. If total capacity or availability is more than the demand and if there are no costs associated with the failure to use the excess capacity, we add a dummy (fictitious) destination to take up the excess capacity and the costs of shipping to this destination are set equal to zero. The zero cost cells are treated the same way as real cost cells and the problem is solved as a balanced problem. If there is, however, a cost associated with unused capacity (e.g., maintenance cost) and it is linear, it too can be easily treated.

In case the total demand is more than the availability, we add a dummy origin (source) to "fill" the balance requirement and the shipping costs are again set to equal to zero. However, in real life, the cost of unfilled demand is seldom zero since it may involve lost sales, lesser profits, possibility of losing the customer or even business or the use of a more costly substitute. Solution of the problem under such situations may be more involved.

7.2. The maximization problem

The transportation problem may involve maximization of profit rather than minimization of cost. Such a problem may be solved in one of the following ways:

1. As maximization of a function is equivalent to minimization of negative of that function, the given problem may be converted into a minimization problem by multiplying the profit matrix by – 1. Minimization of this negative

profit matrix by the usual method will be equivalent to the maximization of the given problem.

2. It may be converted into a minimization problem, by subtracting all the profits from the highest profit in the matrix. The problem can then be solved by the usual methods.
3. It may be solved as a maximization problem itself. However, while finding the initial basic feasible solution, allocations are to be made in highest profit cells, rather than in lowest cost cells. Also solution will be optimal when all cell evaluations are non-positive (≤ 0).

7.3. Different production costs

In some industries a particular product may be manufactured and transported from different production locations. The production cost could be different in different units due to various reasons, like higher labour cost, higher cost of transportation of raw materials, higher overhead charges, etc. Under this situation the production cost is added to the transportation cost while finding the optimal solution. While solving the transportation problems, if the variable production costs and the fixed costs are given for various production plants, no consideration is given for the fixed cost.

7.4. No allocation in particular cell/cells

In the transportation of goods from sources to the destinations, some routes may be banned, blocked, affected by flood, etc. To avoid allocations in a particular cell/ cells, a heavy penalty cost is assigned to the cells/ cell and the problem is solved in the usual manner.

7.5. Overtime production

In the production units, overtime production is taken up to increase the production. This will add the cost of production due to the higher wages paid to the employees involved in overtime. Such wages paid also included in the transportation cost.

TRANSPORTATION PROBLEMS

Least Time Transportation Problems

There are some transportation problems where the objective is to minimize time rather than transportation cost. Such problems are usually encountered in hospital management, military

	1	2	3	.....	n	a_i
1	t_{11}	t_{12}	t_{13}	...	t_{1n}	a_1
2	t_{21}	t_{22}	t_{23}	...	t_{2n}	a_2
3	t_{31}	t_{32}	t_{33}	...	t_{3n}	a_3
:	:	:	:	:	:	:
m	t_{m1}	t_{m2}	t_{m3}	...	b_n	a_m
b_j	b_1	b_2	b_3	...	b_n	

services, fire services, etc. where the speed of delivery or time of supply is more important than the transportation cost.

Now, while solving problems where the objective is to minimize time for each route, the cost per unit is replaced by the time required to ship the quantity x_{ij} from origin i to destination j, where i=1,2,3, ….. ,m and j= 1,2,3, ….. ,n. The corresponding transportation matrix is given below.

and $$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$$

Note that the time of shipment is independent of the number of units shipped. Also since shipments from origins to destinations can be done at the same time on different routes, the shipment time of the total plan is not the sum total of the times of the individual routes. In fact, the shipment of a feasible plan will be complete. Such problems, therefore, require a different solution procedure.

Let, T_k be the largest time associated with k^{th} feasible plan. Our objective is, therefore, to find out a plan for which T_k is minimum of all values of k. The procedure for getting minimum T_k consists of the following steps:

Step I: Find an initial basic feasible solution. This is obtained by using the same method as for the normal transportation technique.

Step II: Find T_k corresponding to the current feasible solution and cross out all the non-basic cells for which $t_{ij} \geq T_k$.

Step III: Draw a closed path (as in the normal transportation technique) for the basic variable associated with T_k such that when the values at the corner elements are shifted around, this basic variable reduces to zero and no variable becomes negative. This procedure ends if no such closed path can be traced out, otherwise repeat step II.

1.1.Post optimality analysis in transportation

The transportation models studied above will normally be valid for a limited period only. In actual practice, the resource capacities and/or destination requirements may vary with time. Likewise, there may be some changes in the transportation cost. Such changes may affect the optimal allocation and the associated transportation cost. One way to study the effect of these changes is to solve the problem a new. Many a times, however, it may not be necessary to do so and the new optimal solution may be obtained by simply incorporating the changes in the current optimal solution, seeing the effects of these changes and carrying out iterations if required.

1.2. Changes in transportation costs

An increase in the costs of empty cells will not change the current optimal solutions as the current cost itself is too high, that is why there have been no allocations in these cells.

However, a reduction in costs of empty cells or an increase in costs of allocated cells is likely to change the transportation schedule. In this case the problem is re-considered with the current optimal solution. New values of u_i and v_j numbers are computed and evaluations of the empty cells are determined. If they are all non-negative, the current solution still remains optimal. If not, a new solution is obtained which is then tested for optimality.

1.3. The trans-shipment problem

The transportation problem assumes that direct routes exist from each source to each destination. However, there are situations in which units may be shipped from one source to another or to other destinations before reaching their final destination. This is called a trans-shipment problem. For example, movement of material involving two different modes of transport – road and railways or between stations connected by broad gauge to metre gauge lines will necessarily require transshipment. For the purpose of transshipment the distinction between a source and destination is dropped so that a transportation problem with m sources n destinations gives rise to a transshipment problem will involve [(m+n)+(m+n)–1] or [2m+2n–1] basic variables and if we omit the variables appearing in the (m+n) diagonal cells, we are left with [m+n–1] basic variables.

In the trans-shipment problem, as each source or destination is a potential point of supply as well as demand, the total supply, say of N units, is added to the actual supply of each source, as well as to the actual demand at each destination. Also the 'demand' at each source and 'supply' at each destination are set equal to N.

Therefore, we may assume the supply and demand of each location to be fictitious one. These qualities (N) may be regarded as buffer stocks and each of these buffer stocks should at least be equal to the total supply/demand in the given problem.

The given trans-shipment problem can, therefore, be regarded as the extended transportation problem and can hence be solved by the transportation technique. In the final solution, units transported from a point to itself i.e., in diagonal cells are ignored as they do not have any physical meanings as there is no transportation involved.

1.4. Dual of the transportation problem

We know any linear programming problem has its dual. Since the transportation problem is a special type of linear programming problem, it also has its dual with usual interpretations and applications. To illustrate let us consider example along with the transportation cost table. The mathematical model (Primal) for this problem is rewritten below:

Minimize, $Z = 2x_{11} + 3x_{12} + 11x_{13} + 7x_{14} + x_{21} + 0x_{22} + 2x_{11} + 6x_{23} + x_{24} + 5x_{31} + 8x_{32} + 15x_{33} + 9x_{34}$,

Subject to constraints

$x_{11} + x_{12} + x_{13} + x_{14} = 6$,

$x_{21} + x_{22} + x_{23} + x_{24} = 1$,

$x_{31} + x_{32} + x_{33} + x_{34} = 10$,

$x_{11} + x_{21} + x_{31} = 7$,

$x_{21} + x_{22} + x_{33} = 5$,

$x_{13} + x_{23} + x_{33} = 3$,

$x_{14} + x_{24} + x_{34} = 2$,

where, $x_{ij} \geq 0$; $i = 1,2,3$; $j = 1,2,3,4$.

Now the dual of this linear problem can be written as,

maximize

$Z' = 6u_1 + u_2 + 10u_3 + 7v_1 + 5v_2 + 3v_3 + 2v_4$,

Subject to

$u_1+v_1 \leq 2$,

$u_1+v_2 \leq 3$,

$u_1+v_3 \leq 11$,

$u_1+v_4 \leq 7$,

$u_2+v_1 \leq 1$,

$u_2+v_2 \leq 0$,

$u_2+v_3 \leq 6$,

$u_2+v_4 \leq 1$,

$u_3+v_1 \leq 5$,

$u_3+v_2 \leq 8$,

$u_3+v_3 \leq 15$,

$u_3+v_4 \leq 9$,

where the dual variables u_i and v_j are unrestricted in sign,

$i = 1,2,3$; $j = 1,2,3,4$.

1.5. Interpretation of the dual

1. The dual variables u_i and v_j are the row numbers and column numbers respectively in table, used in solving the problem by the modified distribution method. In dual, u_i may be interpreted as the value of the product, free on board, at the i^{th} origin and, therefore, may be called location rent and v_j can be interpreted as its value (delivered) at the j^{th} destination and, therefore, may be termed market prize. Hence Z' which represents the sum of these two factors is to be maximized. The constraints of the dual indicate that to allocate in a cell, the transportation cost in that cell should not be more than the sum of these two factors for that cell.
2. We know that in case of linear programming, the final (optimal) simplex table also represents the optimal solution of the dual without actually solving it. Likewise, the optimal (final) transportation table of the primal represents the optimal solution of the associated dual. For the problem under consideration, the optimal dual solution as given by table.

$u_1=1$, $u_2=-4$, $u_3=5$, $v_1=0$, $v_2=2$, $v_3=10$, $v_4=4$.

value of

Z'_{max} = Rs. 100 [6 x 1 – 1 x 4 + 10 x 5 + 7 x 0 +5 x 2 + 3 x 10 + 2 x 4]

= Rs. 100 [6 – 4 + 50 + 0 + 10 + 30 + 8]

= Rs. 10,000 , which is same as Z'_{min} .

2. Minimization of Transportation Cost Using Distribution Linear Programming

If the problem can be formulated (modeled) as one of minimizing some given cost, such as transportation expense, the methods of distribution linear programming are useful techniques for minimizing the cost function subject to supply and demand constraints.

Distribution linear programming methods are widely used for minimizing transportation costs and are indeed useful in numerous other maximization or minimization situations, such as maximizing revenue available from various alternative locations, minimizing unit production costs, and minimizing materials handling costs. The demand requirements and supply availabilities (demand-supply constraints) are typically formulated in a rectangular arrangement (matrix) with the transported amounts (cell loadings) being governed by the cost or profit for the particular supply- demand route. Several methods of obtaining initial and final solutions have been developed, some of which include the following.

I. Initial solutions

1. North west corner method
2. Minimum matrix method (minimum cost)
3. Vogel's approximation method

II. Optimal solutions

1. Stepping- stone method
2. Modified distribution method (MODI)

The following example will illustrate the use of an initial allocation via the North West corner method and a final solution via the stepping- stone method. These are not usually the most expedient methods to follow when the problem has any degree of complexity, but they have intuitive value and quickly convey the basic methodology. An optional problem in the solved problem section at the end of this chapter illustrates the use of the minimum cost method.

The solution procedure necessitates that only unused transportation paths (vacant cells) be evaluated, and there is only one available pattern of moves to evaluate each vacant cell. This is because moves are restricted to occupied cells. Every time a vacant cell is filled, one previously occupied cell must become vacant. The initial (and continuing) number of entries is always maintained at R+C-1, that is, number of rows plus number of columns minus one. When a move happens to cause fewer entries (for example, when two cells become vacant at the same time but only one is filled), a "zero" entry must be retained in one of the cells to avoid what is termed a "degeneracy" situation. The zero entry should be assigned to an independent cell, that is, to one that cannot be reached by a closed path involving only filled cells. The cell with the zero entry is then considered to be an occupied and potentially usable cell.

A second potentially troublesome situation may arise when supply and demand are unequal. In this situation a "dummy" supply plant or absorption location can be created either to produce the additional needed supply or to absorb the excess supply.

If demand > supply: create a dummy supply and assign zero transportation cost to it so excess demand is satisfied.

If supply> demand: create a dummy demand and assign zero transportation cost to it so excess supply is absorbed.

- This corresponds to a fundamental linear programming rule which holds that the number of variables in solution must equal the number of constraints that are binding.

Problem No.1:

Obtain an initial basic feasible solution to the following distribution of products to various destinations from the sources.

Sources, b	Destinations, a				Supply, Sb_j
	a_1	a_2	a_3	a_4	
b_1	11	13	17	14	250
b_2	16	18	14	10	300
b_3	21	24	13	10	400
Demand, Sa_i	200	225	275	250	

Solution : Since, Sa_i=Sb_j=950, there exists a feasible solution to the transportation problem. The initial feasible solution can be obtained as given below.

North West Corner Rule (NWC)

Step 1: Starting with the cell at the upper left (north-west) corner of the transportation matrix, allocate as much as possible so that either the capacity of the first row is exhausted or the destination requirement of the first column is satisfied. i.e x_{11}=min(a_1,b_1).

Step 2: If b_1> a_1, the move down vertically to the second row and make the second allocation of magnitude x_{12} = min (a_2,b_1- x_{11}) in the cell (2,1).

If b_1<a_1, then move right horizontally to the second column and make the second allocation of magnitude x_{12} = min (a_1 - x_{11},b_2) in the cell (1,2).

If, b_1=a_1, there is a tie for the second allocation of magnitude x_{12} = min (a_1 - a_1 ,b_1) = 0 in the cell (1,2), or x_{21}=min(a_2,b_1- b_1) = 0 in the cell (2,1).

Step 3: Repeat steps 1 and 2 moving down towards the lower right corner of the transportation table until all the rim requirements are satisfied.

The transportation table of the given problem has 12 cells. Following north-west corner rule, the first allocation is made to the cell (1, 1), the magnitude being x_{11} = min. (250, 200) = 200. The second allocation is made to the cell (1, 2) and the magnitude of the allocation is given by,

$$x_{12} = \text{min. } (250 – 200, 50) = 50.$$

The third allocation is made in the cell (2, 2), the magnitude being x_{22}= min. (300, 225-50)=175. In the cell (2, 3) is given by x_{23}=min. (300 – 175, 275) = 125. The fifth allocation is made in the cell (3, 3), the magnitude being x_{34} = min. (400-150, 250)=250. Hence an initial basic feasible solution to the given transportation problem is obtained and given below.

200 11	50 13	17	14	250
16	175 18	125 14	10	300
21	24	150 13	250 10	
200	225	275	250	400

Table – 1

The transportation cost according to the above route is given by,

$z = (200x11) + (50x13) + (175x18) + (125x14) + (150x13) + (20x10)$

$= Rs.12,200.$

Least cost or Matrix Minima method

Step 1 : Determine the smallest cost in the cost matrix of the transportation table. Let it be C_{ij}. Allocate $x_{ij} = \min (a_i, b_j)$ in the cell (i,j).

Step 2 : If, $x_j = a_i$, cross off the i^{th} row of the transportation table and decrease b_j by a_i. Go to step3.

If $x_{ij} = b_j$, cross off the j^{th} column of the transportation table and decrease a_i by b_j. Go to step3.

If $x_j = a_i = b_j$, cross off either the i^{th} row or j^{th} column but not both.

Step 3 : Repeat steps1 and 2 for the resulting reduced transportation table until all the rim requirements are satisfied. Whenever the minimum cost is not unique, make an arbitrary choice among the minima.

Problem: Obtain an initial basic feasible solution to the following distribution of products to various destinations from the sources as given in Problem No.1.

Sources, b	Destinations, a				Supply, Sb_j
	a_1	a_2	a_3	a_4	
b_1	11	13	17	14	250
b_2	16	18	14	10	300
b_3	21	24	13	10	400
Demand, Sa_i	200	225	275	250	

Solution : Since, $Sa_i=Sb_j=950$, there exists a feasible solution to the transportation problem. The initial feasible solution can be obtained as given below following Least-Cost Method.

Following the least cost method, the first allocation is made in the cell (3,4), the magnitude being x_{34} = min. (400, 250)=250. This satisfies the requirement at a_4 and thus we cross off the fourth column from the table. The second allocation is made in the cell (1, 1) of magnitude x_{11} = min. (250, 200) = 200 which satisfied the required at a_1 and therefore, we cross off the first column from the table. This yields table 2.

11	13	17	14	250
16	18	14	10	300
21	24	13	$\boxed{250}$ 10	400
200	225	275	250	

Table – 1

$\boxed{200}$ 11	13	17	14	250
16	18	14	10	300
21	24	13	$\boxed{250}$ 10	150
200	225	275		

Table – 2

Now there is a tie for third allocation. We choose arbitrarily the cell (1, 2) and allocate x_{12}= min. (50, 225) = 50. Cross off the first row. Fourth allocation is made in the cell (3, 3) with magnitude x_{33}= min.(150, 275)=150. After crossing off the third row, it results in table 3.

$\boxed{200}$ 11	$\boxed{50}$ 13	17	14	50
16	18	14	10	300
21	24	$\boxed{150}$ 13	$\boxed{250}$ 10	150
	225	275		

Table – 3

$\boxed{200}$ 11	$\boxed{50}$ 13	17	14	
16	$\boxed{175}$ 18	$\boxed{125}$ 14	10	300
21	24	$\boxed{150}$ 13	$\boxed{250}$ 10	
	175	125		

Table – 4

Fifth allocation is made to the cell (2, 3) of magnitude x_{23}=min. (300, 125) = 125 and the sixth allocation is given by x_{22}=min (175, 175) = 175.

$\boxed{200}$ 11	$\boxed{50}$ 13	17	14
16	$\boxed{175}$ 18	$\boxed{125}$ 14	10
21	24	$\boxed{150}$ 13	$\boxed{250}$ 10

Table - 5

Thus we get the initial basic feasible solution as displayed in table 5. The transportation cost according to the above route is given by,

z = (200x11) + (50x13) + (175x18) + (125x14) + (150x13) + (250x10) = Rs.12,200.

The Row Minima Method

Step 1 : The smallest cost in the first row of the transportation table is determined. Let it be C_{ij}. Allocate the maximum feasible x_{ij}=min (a_1,b_j) in the cell (1,j).

Step 2: If $x_{1j}=a_1$, cross off the 1st row of the transportation table and move down to the second row. If $x_{1j}=b_j$, cross off the j^{th} column of the transportation table and reconsider the first row with the remaining availability.

If $x_{1j}=a_1=b_j$, cross off the 1st column and make the second allocation $x_{1k}=0$ in the cell (1,k) with C_{1k} being the new minimum cost in the first row. Cross of the first row and move down to the second row.

Step 3: Repeat steps1 and 2 for the resulting reduced transportation table until all the rim requirements are satisfied.

Problem:

Determine an initial basic feasible solution to the following transportation problem following row minima method.

Sources	Destination			Supply, Units
	A	B	C	
I	50	30	220	1

II	90	45	170	3
III	250	200	50	4
Requirements	4	2	2	

Solution:

50	$\boxed{1}$ 30	220	1
90	45	170	3
250	200	50	4
4	2	2	

90	$\boxed{1}$ 45	170	3
250	200	50	4
4	1	2	

$\boxed{2}$ 90	170	2
$\boxed{2}$ 250	$\boxed{2}$ 50	4
4	2	

50	$\boxed{1}$ 30	220	1
$\boxed{2}$ 90	$\boxed{1}$ 45	170	3
$\boxed{2}$ 250	200	$\boxed{2}$ 50	4
4	2	2	

The Column Minima Method

Step 1: Determine the smallest cost in the first column of the transportation table. Let it be C_{i1}. Allocate the maximum feasible units to $x_{i1}=\min(a_i,b_1)$ in the cell (i,1).

Step 2: If $x_{i1}=a_i$, cross off the i^{th} row of the transportation table and reconsider the first column with the remaining availability.

If, $x_{i1}=b_1$, cross off the i^{th} column and move towards right to the second column.

If, $x_{i1}=a_i=b_1$, cross off the i_{th} row and make the second allocation $x_{k1}=0$ in the cell (k,1) with C_{k1} being the new minimum cost in the first column. Cross of the first column and move down to the second column.

Step 3: Repeat steps1 and 2 for the resulting reduced transportation table until all the rim requirements are satisfied.

Problem:

Determine an initial basic feasible solution to the following transportation problem using column minima method.

Sources	Destination			Supply, Units
	A	B	C	
I	16	19	12	14
II	22	13	19	16

III	14	28	8	12
Requirements	10	15	17	

Solution :

16	19	12	14
22	13	19	16
$\boxed{10}$ 14	28	8	12
10	15	17	

19	12	14
$\boxed{15}$ 13	19	16
28	8	2
15	17	

$\boxed{14}$ 12	14
$\boxed{1}$ 19	1
$\boxed{2}$ 8	2
17	

16	19	$\boxed{14}$ 12	14
22	$\boxed{15}$ 13	$\boxed{1}$ 19	16
$\boxed{10}$ 14	28	$\boxed{2}$ 8	12
10	15	17	

Vogels approximation method (VAM or penalty method):

Step 1: Calculate the penalties by taking differences between the minimum and next to minimum unit transportation costs in each row and each column.

Step 2: Circle the largest row difference or column difference. In the event of a tie, choose either.

Step 3: Allocate as much as possible in the lowest cost cell of the row(or column) having a circled row (or column) difference.

Step 4: In case the allocation is made fully to a row (or column) , ignore that row(or column) for further consideration, by crossing it.

Step 5: Revise the differences again and cross out the earlier figures. Go to step2.

Step 6: Continue the procedure until all rows and columns have been crossed out, i.e distribution is complete.

Problem

Obtain an initial basic feasible solution to the following transportation.

Sources, b	Destinations, a	Supply, Sb_j

	a_1	a_2	a_3	a_4	
b_1	11	13	17	14	250
b_2	16	18	14	10	300
b_3	21	24	13	10	400
Demand, Sa_i	200	225	275	250	

Solution: Since, $Sa_i = Sb_j = 950$, there exists a feasible solution to the transportation problem. The initial feasible solution can be obtained as given below.

Following the Vogels approximation method, the differences between the smallest and next to smallest costs in each row and each column are computed and displayed inside the parenthesis against the respective rows and columns. The largest of these differences is (5) and is associated with the first column of the transportation table.

Since the minimum cost in the first column is c_{11}= 11, allocate x11=min. (250, 200)=200 in the cell (1,1). This exhausts the requirement of the first column and, therefore, cross off the first column. The row and column differences are now computed for the resulting reduced transportation table 1, the largest of these is (5) which is associated with the second column. Since c_{12} (=13) is the minimum cost, allocate x_{12} = min. (50, 225) = 50.

200 / 11	13	17	14	250 (2)
16	18	14	10	300 (4)
21	24	13	10	400 (3)
200 (5)	225 (5)	275 (1)	250 (0)	

Table – 1

50 / 13	17	14	50 (1)
18	14	10	300 (4)
24	13	10	400 (3)
225 (5)	275 (1)	250 (0)	

Table – 2

Thus exhausts the availability of first row and, therefore, we cross off the first row. Continuing in this manner, the subsequent reduced transportation tables and the differences for the surviving rows and columns are shown in table:

$\boxed{175}$ 18	14	10	300 (4)
24	13	10	400 (3)
175 (6)	275 (1)	250 (0)	

14	$\boxed{125}$ 10	125 (4)
13	10	400 (3)
275 (1)	250 (0)	

$\boxed{275}$ 13	$\boxed{125}$ 10	400
275	125	

Eventually, the basic feasible solution obtained is shown in the following table:

$\boxed{200}$ 11	$\boxed{50}$ 13	17	14
16	$\boxed{175}$ 18	14	$\boxed{125}$ 10
21	24	$\boxed{275}$ 13	$\boxed{125}$ 10

The transportation cost according to this route is given by

z = (200x11) + (50x13) + (175x18) + (125x10) + (275x13) + (125x10) = Rs.12,075

TRANSPORTATION PROBLEMS

The Stepping-Stone Method

Consider the matrix giving the initial feasible solution for the problem under consideration. Let us start with any arbitrary empty cell (a cell without allocation), say (3, 2) and allocate + 1 unit to this cell. As already discussed, in order to keep up the column 2 restriction, -1 must be allocated to cell (1, 2) and to keep up the row 1 restriction, +1 must be allocated to cell (1,1) and consequently -1 must be allocated to cell (3, 1); this is shown in the matrix below.

	1	2	3	4	
1	2 (1)+1	3 -1(5)	11	7	6
2	1	0	6	1 (1)	1
3	5 (6) (-1)	8 +1	15 (3)	9 (1)	10
	7	5	3	2	

Table

The net change in transportation cost as a result of this perturbation is called the evaluation of the empty cell in question.

Therefore,

Evaluation of cell (3,2) = Rs.100 x (8 x 1 – 5 x 1 + 2 x 1 – 5 x 1)
= Rs. (0 x 100)
= Rs.0.

Thus the total transportation cost increases by Rs.0 for each unit allocated to cell (3,2). Likewise, the net evaluation (also called opportunity cost) is calculated for every empty cell. For this the following simple procedure may be adopted.

Starting from the chosen empty cell, trace a path in the matrix consisting of a series of alternate horizontal and vertical lines. The path begins and terminals in the chosen cell. All other corners of the path lie in the cells for which allocations have been made. The path may skip over any number of occupied or vacant cells. Mark the corner of the path in the chosen vacant cell as positive and other corners of the path alternatively –ve, +ve, -ve and so on. Allocate I unit to the chosen cell; subtract and add 1 unit from the cells at the corners of the path, maintaining the row and column requirements. The net change in the total cost resulting from this adjustment is called the evaluation of the chosen empty cell. Evaluation of the various empty cells (in hundreds of rupees) are:

Cell (1, 3) = c_{13}-c_{33}+c_{31}-c_{11} = 11-15+5 = -1,
Cell (1, 4) = c_{14}-c_{34}+c_{31}-c_{11} = 7-9+5-2 = +1,
Cell (2, 1) = c_{21}-c_{24}+c_{34}-c_{31} = 1-1+9-5 = +4,
Cell (2, 2) = c_{22}-c_{24}+c_{34}-c_{31}+c_{11}-c_{12} = 0-1+9-5+2-3 = +2,
Cell (2, 3) = c_{23}-c_{24}+c_{34}-c_{33} = 6-1+9-15 = -1,
Cell (3, 2) = c_{32}-c_{31}+c_{11}-c_{12} = 8-5+2-3 = +2.

If any cell evaluation is negative, the cost can be reduced so that the solution under consideration can be improved i.e., it is not optimal. On the other hand, if all cell evaluations are positive or zero, the solution in question will be optimal. Since evaluations of cells (1, 3) and (2, 3) are –ve, initial basic feasible solution given in table 3.15 is not optimal.

Now in a transportation problem involving m rows and n columns, the total number of empty cells will be,

$$m.n – (m+n-1) = (m-1)(n-1).$$

Therefore, there are (m-1) (n-1) such cell evaluations which must be calculated and for large problems, the method can be quite inefficient. This method is named 'stepping-stone' since only occupied cells or 'stepping stones' are used in the evaluation of vacant cells.

2. Transportation Algorithm (Modified Distribution Method - MODI method)

Various steps involved in solving any transportation problem may be summarized in the following iterative procedure:

Step1: Find the initial basic feasible solution by using any of three methods discussed above.

Step2: Check the number of occupied cells. If there are less than m+n-1, there exists degeneracy and it is possible to introduce a very small positive assignment of Î(»0) in suitable independent positions, so that the number of occupied cells is exactly equal to m+n+1.

Step 3: For each occupied cell in the current solution, solve the system of equations $u_i+v_j=c_{ij}$.

Starting initially with some $u_i=0$ or $v_j=0$ and entering the successively the values of u_i and v_j in the transportation table margins.

Step 4: Compute the net evaluation $z_{ij}-c_{ij}=u_i+v_j-c_{ij}$ for all unoccupied basic cells and enter them in the upper right corners of the corresponding cells.

Step 5: Examine the sign of each $z_{ij}-c_{ij.}$ If all $z_{ij}-c_{ij}$ £ 0 , then the current basic feasible solution is an optimum one. If atleast one $z_{ij}-c_{ij} > 0$, select the unoccupied cells, having the largest positive net evaluation to enter the basis.

Step 6: Let the unoccupied cell (r,s) enter the basis. Allocate an unknown quantity, say q, to the cell (r,s). Identify a loop that starts and ends at the cell (r,s) and connects some of the basic cells. Add and subtract interchangeably, q to and from the transition cells of the loop in such a way that the rim requirements remain satisfied.

Step 7: Assign a maximum value to q in such a way that the value of one basic variable becomes zero and the other basic variables remain non-negative. The basic cell whose allocation has been reduced to Zero, leaves the basis.

Step8: Return to step 3 and repeat the process until an optimum basic feasible solution has been obtained.

3. Unbalanced Transportation Problem

For a feasible solution to exist in a transportation problem, it is necessary that the total supply must equal demand. That is, . But a situation may arise when the total available supply is not equal to the total requirement. Such type of transportation problem is called unbalanced transportation problem.

Problem:

Find the starting solution in the following transportation problem by any one of the methods (North-west Corner Method, Least–Cost Method, Vogel's Approximation Method, Row minima method or column minima method). Also obtain the optimum solution by using the best starting solution:

Source	Destination				Supply
	D_1	D_2	D_3	D_4	
S_1	3	7	6	4	5
S_2	2	4	3	2	2
S_3	4	3	8	5	3
Demand	3	3	2	2	

Step 1: In this problem we find the basic feasible solution by using Vogel's approximation method.

By Vogel's Approximation Method, the differences between the two successive lowest costs for each row and each column are computed. These are written besides the corresponding rows or columns under the heading 'row difference' or 'column difference'':

				Row Difference		
$\boxed{3}$ 3	7	6	$\boxed{2}$ 4	(1)	(1)	2
2	4	$\boxed{2}$ 3	$\boxed{\epsilon_1}$ 2	(1)	(2)	-
4	$\boxed{3}$ 3	8	$\boxed{\epsilon_2}$ 5	(1)	(1)	1
Column Difference						
(1)	(1)	3	(2)			
(1)	(1)	_	1			
1	-	-	(1)			

The largest row difference or column difference is 3 which corresponds to third column. Allocating as much as possible to the lowest cost cell in the third column give x_{23} = 2. This satisfies the supply at S_2 and the requirement at D_3. So arbitrarily we cross off the third column and consider e_1 (0) as the small quantity to be supplied form S_2. The row and column differences are now recomputed for the reduced cost matrix, and choose the largest of these, i.e., 2 which corresponds to fourth column. Therefore allocate $x_{24} = e_1$ and second row is now eliminated. Again, compute the row and column differences for the reduced 2x3 transportation table and choose the largest of these, i.e., 4 corresponding to the second column. Thus allocate x_{32} = 3 and eliminate second column while consider e_2 (0) being the negligible positive quantity to be supplied form S_3. Again, computing the column and row differences in the reduced 2x2 cost matrix, we assign x_{11} = 3 and thus eliminate first column. This leaves now only fourth column, which can be filled by inspection and thus we assign x_{14} = 2 and $x_{34}e_2$.

The transportation cost according to above allocation is given by

Total cost $= (3x3) + (2x4) + (2x3) + (\varepsilon_1 x2) + (3x3) + (\varepsilon_2 x5)$

$= 32 + 2\varepsilon_1 + 5\varepsilon_2$

$=32$ as $\varepsilon_1 \longrightarrow 0$ and $\varepsilon_1 \longrightarrow 0$.

Step 2: From the initial basic feasible solution obtained in step 1, it is observed that the best starting solution is obtained using Vogel's approximation method having the least transportation cost of 32. Also, observe that unknown quantities e_1, e_2 (> 0) have been allocated in the unoccupied cells (2,4) and (3,4) respectively to overcome the danger of degeneracy.

Step 3: Now compute the numbers u_i (i = 1,2,3) and v_j (j = 1,2,3,4) using successively the equations $u_i + v_j = c_{ij}$ for all the occupied cells. For this it may be arbitrarily assigned as, v_4 = 0. Thus we have

$$u_1 + v_4 = c_{14} \Rightarrow u_1 + 0 = 4 \Rightarrow u_1 = 4$$

$u_2 + v_4 = c_{24} \Rightarrow u_2 + 0 = 2 \Rightarrow u_2 = 2$

$u_3 + v_4 = c_{34} \Rightarrow u_3 + 0 = 5 \Rightarrow u_3 = 5$

Given u_1,u_2 and u_3, values of v_1,v_2 and v_3 can be calculated as shown below :

$u_1 + v_1 = c_{11} \Rightarrow 4 + v_1 = 3 \Rightarrow v_1 = -1$

$u_2 + v_2 = c_{23} \Rightarrow 2 + v_3 = 3 \Rightarrow u_3 = 1$

$u_3 + v_3 = c_{32} \Rightarrow 5 + v_2 = 3 \Rightarrow v_2 = -2$

The net evaluations for each of the unoccupied cells are now determined:

$z_{12} – c_{12} = u_1 + v_2 – c_{12} = 4 + (-2) -7 = -5$

$z_{13} – c_{13} = u_1 + v_3 – c_{13} = 4 + 1 -6 = -1$

$z_{21} – c_{21} = u_2 + v_1 – c_{21} = 2 + (-1) - 2 = -1$

$z_{22} – c_{22} = u_2 + v_2 – c_{22} = 2 + (-2) -4 = - 4$

$z_{31} – c_{31} = u_3 + v_1 – c_{31} = 5 + (-1) –4 = 0$

$z_{33} – c_{33} = u_3 + v_3 – c_{33} = 5 + 1 -8 = -2$

Since all z_y - c_y £ 0, the current basic feasible solution is an optimum one.

It is possible to present a more compact form for computing the unknown u_t's and v_j's and then evaluate the net evaluations for each of the unoccupied cells in a more convenient way by working in the transportation table margins:

					u_i
	$\boxed{3}$ 3	(-5) 7	-1 6	$\boxed{2}$ 4	4
	(-1) 2	(-4) 4	$\boxed{2}$ 3	$\boxed{\in_1}$ 2	2
	(0) 4	$\boxed{3}$ 3	(-2) 8	$\boxed{\in_2}$ 5	5
v_j	-1	-2	1	0	

The optimum solution is:

x_{11}=3; x_{14} = 2; x_{23}= 2; x_{24}=Î$_1$; x_{32}=3; x_{34}=Î$_2$.

The transportation cost associated with the optimum schedule is given by:

Total cost = (3x3) + (2x4) + (2x3) + (2x $\in_1$) + (3x3) + (5x $\in_2$)

= 32 + 2 $\in_1$ + 5 $\in_2$

= 32, since $\in_1 \rightarrow 0$ and $\in_2 \rightarrow 0$.

ASSIGNMENT MODELS

The assignment problem is defined as assigning each facility to one and only one job so as to optimize the given measures of effectiveness, when *n* facilities and *n* jobs are available and given the effectiveness of each facility for each job.

Let there be *n* facilities (machines) to be assigned to *n* jobs. Let c_{ij} is cost of assigning i^{th} facility to j^{th} job and x_{ij} represents the assignment of i^{th} facility to j^{th} job. If i^{th} facility can be assigned to j^{th} job, x_{ij}=1, otherwise zero. The objective is to make assignments that minimize the total assignment cost or maximize the total associated gain.

Jobs

		1	2	....	n	a_1 *Supply*
	1	C_{11}	C_{12}	...	C_{1n}	1
	2	C_{21}	C_{22}	...	C_{2n}	1
Facilities	:	:	:	:	:	:
	n	C_{n1}	C_{n2}	...	C_{nn}	*L*
demand b_j		1	1	...	1	

Thus an assignment problem can be represented by n x n matrix which continues n! possible ways of making assignments. One obvious way to find the optimal solution is to write all the n! possible arrangements, evaluate the cost of each and select the one involving the minimum cost.

However, this enumeration method is extremely slow and time consuming even for small values of n. For example, for n = 10, a common situation, the number of possible arrangements is 10! = 3,628,800. Evaluation of so large a number of arrangements will take a probability large time. This confirms the need for an efficient computational technique for solving such problems.

2. Mathematical Representation of Assignment Model

Mathematically, the assignment model can be expressed as follows:

Let x_{ij} denote the assignment of facility i to job j such that

$x_{ij} = 0$, if the i^{th} facility is not assigned to j^{th} job,

$x_{ij} = 1$, if the i^{th} facility is assigned to j^{th} job.

Then, the model is given by

$$Minimize\ z' = \sum_{j=1}^{n}\sum_{i=1}^{n} C_{ij}X_{ij} \quad \left\{= \sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}X_{ij}\right\}$$

subject to constraints

$$\sum_{j=1}^{n} X_{ij} = 1, i = 1,2,3, \dots \dots . n,$$

(one job is assigned to the i^{th} facility)

$$\sum_{i=1}^{n} X_{ij} = 1, j = 1,2,3, \dots \dots n,$$ (one job is assigned to the j^{th} facility)

and $x_{ij} = 0$ or 1 (or $x_{ij} = x_{ij}^2$).

If the last condition is replaced by $x_{ij} \geq 0$, this will be a transportation model with all requirements and available resources equal to 1.

3. Comparison with the transportation model

An assignment model may be regarded as a special case of the transportation model. Here, facilities represent the 'sources' and jobs represent the 'destinations'. Number of sources is equal to the number of destinations, supply at each source is unity ($a_i = 1$ for all i) and demand at each destination is also unity ($b_j = 1$, for all j). The cost of 'transporting' (assigning) facility i to job j is c_{ij} and the number of units allocated to a cell can be either one or zero. i.e., they are non-negative quantities.

However the transportation algorithm is not very useful to solve this model, when an assignment is made, the row as well as column requirements are satisfied simultaneously (rim conditions being always unity) resulting in degeneracy. Thus the assignment problem is a completely degenerate form of the transportation problem. In n x n problem, there will be n assignments instead of $n+n–1$ or $2n–1–n = n–1$ epsilons which will make the computations quite cumbersome. However, the special structure of the assignment model allows a more convenient and simple method of solution.

Difference between the transportation problem and the assignment problem

	Transportation Problem	Assignment Problem
(a)	Supply at any source may be any positive quantity a_i	Supply at any source (machine) will be 1 i.e., $a_i = 1$.
(b)	Demand at any destination may be any positive quantity b_j	Demand at any destination (job) will be 1 i.e., $b_j = 1$.
(c)	One or more source to any number of destinations	One source (machine) to only one destination (job).

4. Theorems

The technique used for solving assignment model makes use of the following two theorems:

4.1. Theorem I

It states that in an assignment problem, if we add or subtract a constant to every element of a row (or column) in the cost matrix, then an assignment which minimizes the total cost on one matrix also minimizes the total cost on the other matrix".

Let, c_{ij} represent the original cost elements of the matrix. If constants u_i and v_j are subtracted from the i^{th} row j^{th} column respectively, the new cost elements will be

$$C'_{ij} = c_{ij} - u_i - v_j$$

If Z is the original objective function, the new objective function will be

$$Z' = \sum_{i=1}^{n}\sum_{j=1}^{n}(c_{ij} - u_i v_j)\, X_{ij}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}X_{ij} - \sum_{i=1}^{n} u_i \sum_{j=1}^{n} X_{ij} - \sum_{j=1}^{n} v_j \sum_{i=1}^{n} x_{ij}$$

Now with reference,

$$z = \sum_{j=1}^{n}\sum_{i=1}^{n} C_{ij}X_{ij}$$

and

$$\sum_{i=1}^{n} X_{ij} = \sum_{j=1}^{n} X_{ij} = 1$$

$$\therefore \quad Z' = Z - \sum_{i=1}^{n} u_i - \sum_{j=1}^{n} v_j = Z - Constant$$

or z' is minimum when z is minimum. This proves the theorem.

Likewise, if in an assignment problem some cost elements are negative, we may convent them into an equivalent assignment problem where all the cost elements are non-negative by adding a suitably large constant to the elements of the relevant row.

4.2. Theorem II

It states "If all $c_{ij} \geq 0$ and we can find a set $x_{ij}=x'_{ij}$, such that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}X'_{ij} = 0$$

then this solution is optimal.

The result follows automatically since as neither of c_{ij} is negative, the value of

$$Z = \sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}X_{ij}$$

cannot be negative.

Hence, its minimum value is zero which is attained when $x_{ij}=x_{ij}'$.

Thus the present solution is optimal solution.

The above two theorems indicate that if one can create a new c_{ij}' matrix with zero entries, and if these zero elements, or a subset thereof, constitute feasible solution, then this feasible solution is the optimal solution.

Thus the method of solution consists of adding and subtracting constant from rows and columns until sufficient number of c_{ij}'s become zero to yield a solution with a value of zero.

Solution of the Assignment Models

The cost of any action consists of opportunities that are sacrificed in taking that action. Consider the following table which contains the cost in rupees of processing each of jobs, *A*, *B* and *C* on machines *X*,*Y* and *Z*.

		Machines		
		X	*Y*	*Z*
Jobs	*A*	25	15	22
	B	31	20	19
	C	35	24	17

If job *A* is assigned to machine *X*, the cost of this assignment is Rs. 25. Since machine *Y* can also process job *A* for Rs. 15, clearly assigning job *A* to machine *X* is not the best decision. Therefore, when job *A* is arbitrarily assigned to machine *X*, it is done by sacrificing the opportunity to save Rs. 10 (Rs. 25 – Rs. 15). This sacrifice is referred to as an opportunity cost. The decision to process job *A* on machine *X* precludes the assignment of this job to machine *Y*, given the constraint that one and only one job can be assigned to a machine. Thus opportunity cost of assigning job *A* to machine *X* is Rs.10 with respect to the lowest cost assignment for job *A*. Likewise, a decision to assign job *A* to machine *Z* would involve an opportunity cost of Rs. 7 (Rs. 22 – Rs. 15). Theassignment of job *A* to machine *Y* is the best assignment as the opportunity cost of this assignment is zer (Rs.15-Rs.15). This is called the machine – opportunity costs with regard to job *A*. Similarly, if the lowest cost of row *B* is subtracted from all the costs in this row, the machine-opportunity costs for job *B* can be obtained. By following the same step, the machine opportunity cost for job *C* can be obtained . This is given in the following table.

		Machines		
		X	*Y*	*Z*
Jobs	*A*	10	0	7
	B	12	1	0
	C	18	7	0

Machine-opportunity cost table

In addition to these machine-opportunity costs, there are job-opportunity costs also. Job *A*, *B* and *C*, for instance, could be assigned to machine *X*. The assignment of job *B* to machine *X* involves a cost of Rs. 31, while the assignment of job *A* to machine *X* costs only Rs. 25. Therefore, the opportunity costs of assigning job *B* to machine *X* is Rs. 6 (Rs. 31 – Rs. 25). Similarly, the opportunity cost is involved in the assignment of job *A* to machine *X* is Rs. 10 (Rs. 35 – Rs. 25). A zero opportunity cost is involved in the assignment of job *A* to machine *X*, since this is the best assignment for machine *X* (column *X*). Hence job-opportunity costs for each column (each machine) are obtained by subtracting the lowest cost entry in each column from all cost entries in that column, if the lowest entry in each column of table is subtracted from all the cost entries of that column, the resulting table is called total opportunity cost table.

Machines

Jobs	X	Y	Z
A	10 – 10 = 0	0 – 0 = 0	7 – 0 = 7
B	12 – 10 = 2	1 – 0 = 1	0 – 0 = 0
C	18 – 10 = 8	7 – 0 = 7	0 – 0 = 0

Total opportunity cost table

The objective is to assign the jobs to the machines to manimize total costs. With the total opportunity cost table this objective will be achieved if the jobs are assigned to the machines in such a way as to obtain a total opportunity cost of zero. Four cells in the total opportunity cost table contains zeros, indicating a zero opportunity cost for these cells (assignment). Hence job *A* could be assigned to machine *X* or *Y* and job *B* to machine *Z*, all assignments having zero opportunity costs. This way job *C*, however, could not be assigned to any machine with a zero opportunity cost since assignment of job *B* to machine Z precludes the assignment of job *C* to this machine. Clearly, to make an optimal assignment of the three jobs to the three machines, there must be three zero cells in the table such that a complete assignment to these cells can be made with a total opportunity cost of zero.

Drawing minimum number of lines covering all zero cells in the total opportunity cost table with minimum number of lines equals the number of rows (or columns) in the table is a convenient method for determining whether an optimal assignment is made If, however, the minimum number of lines is less than the number of rows (or columns), an optimal assignment cannot be made. In this case there is need to develop a new total opportunity cost table. In the present example, since it requires only two lines to cross (cover) all zeros, and there are three rows, an optimal assignment is not possible. Clearly, there is a need to modify the total opportunity cost table by including some assignment not in the rows and columns covered by the lines. Of course, the assignment chosen should have the least opportunity cost. In the present case it is the assignment of job B to machine Y with an opportunity cost of 1. In other words, we would like to change the opportunity cost for this assignment from 1 to zero.

Machines

		X	Y	Z
	A	0	0	7
Jobs	B	2	1	0
	C	8	7	0

Line 1

Line 2

To accomplish this (a) choose the smallest elements in the table not covered by a straight line and subtract this element from all other elements not having a line through them (b) add this smallest element to all elements lying at the intersection of any two lines. The revised total opportunity cost table is shown below.

		Machines		
		X	Y	Z
	A	0	0	7 + 1 = 8
Jobs	B	2 – 1 = 1	1 – 1 = 0	0
	C	8 – 1 = 7	7 – 1 = 6	0

Revised opportunity cost table

The test for optimal assignment described above is applied again to the revised opportunity cost table. As the minimum number of lines covering all zeros is three and there are three rows (or columns), an optimal assignment can be made. The optimal assignments are A to X, B to Y and C to Z.

In larger problems there is need for more systematic procedure, as the assignments may not be readily apparent.

Assignment Algorithm (or) Hungarian Method

First check whether the number of rows is equal to the number of columns, If it is so, the assignment problem is said to be **balanced.** Then proceed to step 1. If it is not balanced, then it should be balanced before applying the algorithm.

Step 1: Subtract the smallest cost element of each row from all the elements in the row of the given cost matrix. See that each row contains atleast one zero.

Step 2: Subtract the smallest cost element of each column from all the elements in the column of the resulting cost matrix obtained by step 1.

Step 3: (Assigning the zeros)

(a) Examine the rows successively until a row with exactly one unmarked zero is found. Make an assignment to this single unmarked zero by encircling it. Cross all other zeros in the column of this encircled zero, as these will not be considered for any future assignment. Continue in this way until all the rows have been examined.

(b) Examine the columns successively until a column with exactly one unmarked zero is found. Make an assignment to this single unmarked zero by encircling it and cross any other zero in its row. Continue until all the columns have been examined.

Step 4: (Apply optimal Test)

(a) If each row and each column contain exactly one encircled zero, then the current assignment is optimal.

(b) If atleast one row/column is without an assignment (i.e., if there is atleast one row/column is without one encircled zero), then the current assignment is not optimal. Go to step 5.

Step 5: Cover all the zeros by drawing a minimum number of straight lines as follows.

(a) Mark (√) the rows that do not have assignment.

(b) Mark (√) the columns (not already marked) that have zeros in marked rows.

(c) Mark (√) the rows (not already marked) that have assignments in marked columns.

(d) Repeat (b) and (c) until no more marking is required.

(e) Draw lines through all unmarked rows and marked columns. If the number of these lines is equal to the order of the matrix then it is an optimum solution otherwise not.

Step 6: Determine the smallest cost element not covered by the straight lines. Subtract this smallest cost element from all the uncovered elements and add this to all those elements which are lying in the intersection of these straight lines and do not change the remaining elements which lie on the straight lines.

Step 7: Repeat step (1) to (6), until an optimum assignment is attained.

Problem:

A works manager has to allocate four different jobs to four workmen. Depending on the efficiency and the capacity of the individual the times taken by each differ as shown in the Table 1. How should the tasks be assigned one job to a worker so as to minimize the total man-hours?

Table 1

	Worker			
Job	A	B	C	D
1	10	20	18	14
2	15	25	9	25
3	30	19	17	12
4	19	24	20	10

Solution:

The following steps are followed to the find an optimal solution.

STEP 1

Consider each row. Select the minimum element in each row. Subtract this smallest element form all the elements in that row. This results in the table 2.

Table 2

	Worker			
Job	A	B	C	D
1	0	10	8	4
2	6	16	0	16
3	18	7	5	0
4	9	14	10	0

STEP 2

Subtract the minimum element in each column from all the elements in its column. This will result in Table 3.

Table 3

	Worker			
Job	*A*	*B*	*C*	*D*
1	0	3	8	4
2	6	9	0	16
3	18	0	5	0
4	9	7	10	0

STEP 3

In this way we make sure that in the matrix each row and each column has atleast one zero element. Having obtained atleast one zero in each row and each column, we assign starting from first row. In the first row, we have a zero in (1,*A*). Hence we assign job 1 to the worker *A*. This assignment is indicated by a square. All other zeros in the column are crossed (*X*) to show that the other jobs cannot be assigned to worker *A*, who has already been assigned. In the above problem we do not have other zeros in column A.

Proceed to the second row where there is a zero in (2,*C*) . Hence the job 2 can be assigned to worker *C*, indicating by a square. Any other zero in this column is crossed (*X*).

Proceed to the third row. Here we have two zeros corresponding to (3,*B*) and (3, *D*). Since there is a tie for the job 3, go to the next row deferring the decision for the present. Proceeding to the fourth row, there IS only one Zero in (4, *D*). Hence assign job 4 to worker *D*. Now the column *D* has a zero in the third row and cross (3, *D*). All the assignments made in this way are as shown in Table 4.

Table 4

Job	Worker			
	A	B	C	D
1	$\boxed{0}$	3	8	4
2	6	9	$\boxed{0}$	14
3	18	$\boxed{0}$	5	$\cancel{0}$
4	9	7	10	$\boxed{0}$

STEP 4

Now having assigned certain jobs to certain workers we proceed to the column 1. Since there is an assignment in this column, we proceed to the second column. There is only one zero in the cell (3, *B*); We assign the jobs 3 to worker *B*. Thus all the four jobs have been assigned to four workers. Thus we obtain the solution to the problem as shown in the Table 5

Table 5

Job	Worker			
	A	B	C	D
1	$\boxed{0}$	3	8	4
2	6	9	$\boxed{0}$	16
3	18	$\boxed{0}$	5	$\cancel{0}$
4	9	7	10	$\boxed{0}$

The assignments are

Job to Worker

1	A
2	B
3	C
4	D

The above procedure is summerised as a set of following rules:

1. Subtract the minimum element in each row from all the elements in its row to make sure that at least there is one zero in that row.
2. Subtract the minimum element in each column from all the elements in its column in the above reduced matrix, to make sure that atleast one zero exists in each column.
3. Having obtained atleast one zero in each row and atleast one zero in each column, examine rows successively until a row with exactly one unmarked zero is found and mark (X) this zero, indicating that assignment in made there. Mark (X) all other zeros in the same column, to show that they cannot be used to make other assignments. Proceed in this way until all rows have been examined. If there is a tie among zeros defer the decision.
4. Next consider columns, for single unmarked zero, mark them ( ) and mark (X) any other unmarked zero in their rows.
5. Repeat (c) and (d) successively until one of the two occurs

(1) There are no zeros left unmarked.

(2) The remaining unmarked zeros lie atleast two in each row and column. i.e., they occupy corners of a square.

If the outcome is (1), we have a maximal assignment. In the outcome (2) we use arbitrary assignments. This process may yield multiple solutions.

Solving problem following Hungarian Method

Problem:

Solve the following assignment problem to minimize the total cost represented as elements in the matrix (cost in thousand rupees).

Building	Contractor			
	1	2	3	4
A	48	48	50	44
B	56	60	60	68
C	96	94	90	85
D	42	44	54	46

Solution:

STEP 1:

Choose the least element in each row of the matrix and subtract the same from all the elements in each row so that each row contains atleast one zero. This results in Table 6

	Contractor			
Building	1	2	3	4
A	4	4	6	0
B	0	4	4	12
C	11	9	5	0
D	0	2	12	4

Table 6

STEP 2:

Choose the least element from each column in Table 6 and subtract the same form all the elements in that column to ensure that there is atleast one zero in each column. This results in the following table (Table 7)

Table 7

Building	Contractor			
	1	2	3	4
A	4	2	2	$\boxed{0}$
B	$\boxed{0}$	2	$\cancel{0}$	12
C	11	7	1	$\cancel{0}$
D	$\cancel{0}$	$\boxed{0}$	8	4

STEP 3:

Make the assignment in each row and column as explained previously. This results in table 8

Table 8

Building	Contractor			
	1	2	3	4
A	4	2	2	$\boxed{0}$
B	$\boxed{0}$	2	$\cancel{0}$	12
C	11	7	1	$\cancel{0}$
D	$\cancel{0}$	$\boxed{0}$	8	4

Here there are only three assignments, but four assignments are required. With this maximal assignment, minimum number of lines to cover all the zeros are needed to draw. This is carried out as explained in steps 4 to 9. Refer Table 9

Table 9

Building	Contractor			
	1	2	3	4
A	4	2	2	$\boxed{0}$
B	$\boxed{0}$	2	$\cancel{0}$	12
C	11	7	1	$\cancel{0}$
D	$\cancel{0}$	$\boxed{0}$	8	4

STEP 4:

Mark (X) the unassigned row (row C).

STEP 5:

Against the marked column 4, look for any (X) element and mark that column (column 4).

STEP 6:

Against the marked column 4, look for any assignment and mark that row (row A.)

STEP 7:

Repeat steps 6 and 7 until the chain of markings ends.

STEP 8:

Draw lines through all unmarked rows (row B and Row D) and through all marked columns (column 4). (Check: There should be only three lines to cover all the zeros).

STEP 9:

Select the minimum from the elements that do not have a line through them. In this example we have 1 as the minimum element, subtract the same from all the elements that do not have a line through them and add this smallest element at the intersection of two lines. Thus we have the matrix shown in Table 10

Table 10

Building	Contractor			
	1	2	3	4
A	3	1	1	$\boxed{0}$
B	$\boxed{0}$	2	$\cancel{0}$	13
C	10	6	$\boxed{0}$	$\cancel{0}$
D	$\cancel{0}$	$\boxed{0}$	8	5

STEP 10:

Go ahead with the assignment with the usual procedure. This is carried out in the Table 10. Thus we have four assignments.

Building A is allotted to contractor 4
Building B is allotted to contractor 1
Building C is allotted to contractor 3
Building D is allotted to contractor 2
Total cost is 44 + 56 + 90 + 44 = Rs.234 thousands.

Example 2

There are five machines and five jobs are to be assigned and the associated cost matrix is as follows. Find the proper assignment.

		Machines				
		I	*II*	*III*	*IV*	*V*
	A	6	12	3	11	15
	B	4	2	7	1	10
Jobs	*C*	8	11	10	7	11
	D	16	19	12	23	21
	E	9	5	7	6	10

Solution:
In order to find the proper assignment, we apply the Hungarian method as follows:
Step 1: Row reduction

Machines

		I	II	III	IV	V
	A	3	9	0	8	12
	B	3	1	6	0	9
Jobs	C	1	4	3	0	4
	D	4	7	0	11	9
	E	4	0	2	1	5

Step 2: (Column reduction)

Machines

		I	II	III	IV	V
	A	2	9	0	8	8
	B	2	1	6	0	5
Jobs	C	0	4	3	0	0
	D	3	7	0	11	5
	E	3	0	2	1	1

Step 3: (Zero Assignment)

Machines

		I	II	III	IV	V
	A	2	9	$\boxed{0}$	8	8
	B	2	1	6	$\boxed{0}$	5
Jobs	C	$\boxed{0}$	4	3	~~0~~	~~0~~
	D	3	7	~~0~~	11	5
	E	3	$\boxed{0}$	2	1	1

From the last table we see that all the zeros are either assigned or crossed out, but the total number of assignment,i.e., 4<5 (number of jobs to be assigned to machines). Therefore, we have to follow step 4 and onwards as follows:

Step 4:

Machines

		I	II	III	IV	V	
	A	2	9	$\boxed{0}$	8	8	√
	B	2	1	6	$\boxed{0}$	5	
Jobs	C	$\boxed{0}$	4	3	$\cancel{0}$	$\cancel{0}$	
	D	3	7	$\cancel{0}$	11	5	√
	E	3	$\boxed{0}$	2	1	1	

Step 5:
Here, the smallest element among the uncovered elements is 2.
(i) Subtract 2 from all those elements which are not covered.
(ii) Add 2 to those entries which are at the junction of two lines.
Complete the table as under:

Machines

		I	II	III	IV	V
	A	0	7	0	6	6
	B	2	1	8	0	5
Jobs	C	0	4	5	0	0
	D	1	5	0	9	3
	E	3	0	4	1	1

Step 6: using step 3 again

Machines

		I	II	III	IV	V
	A	$\boxed{0}$	7	$\cancel{0}$	6	6
	B	2	1	8	$\boxed{0}$	5
Jobs	C	$\cancel{0}$	4	5	$\cancel{0}$	$\boxed{0}$
	D	1	5	$\boxed{0}$	9	3
	E	3	$\boxed{0}$	4	1	1

Thus, we have got five assignments as required by the problem.

The assignment is as follows:

A ⟶ I, B ⟶ IV,C ⟶ V, D ⟶ III and E ⟶ II.

Thus from the cost matrix the minimum cost = 6+1+11+12+5=Rs.35.
Note:
If we are given a maximization problem then convert it into minimization problem, simply, multiplying by -1 to each entry in the effectiveness matrix and then solve it in the usual manner.

TESTS FOR OPTIMALITY

Optimality test can be performed if two conditions are satisfied i.e.
1. There are m + n – 1 allocations, whose m is number of rows, n is number of columns. Here m + n – 1 =6. But number of allocation is five.
2. These m + n – 1 allocations should be at independent positions. I.e. it should not be possible to increase or decrease any allocation without either changing the position of the allocations or violating the row or coloumn restrictions.

A simple rule for allocations to be in independent positions is that it is impossible to travel from any allocation, back to itself by a series of horizontal and vertical steps form one occupied cell to another, without a direct reversal of route. It can be seen that in present example, allocation are at independent positions as no closed loop can be formed at the allocated cells.
Therefore first condition is not satisfied and therefore in order to satisfy first condition, we will have to allocate a small amount E at the vacant cells having lowest cost of transportation. It can be seen that t can be allocated at cell (2, 2) having cost of 7 units and still the allocations will remain at independent position as described below:

Table 2

10	$2_{(15)}$	20	11
12	$7_{(t)}$	$9_{(15)}$	$20_{(10)}$
$4_{(5)}$	14	16	$18_{(5)}$

Allocations

Now the number of allocation is m + n- =6 and they are at independent positions.

Write down cost matrix at allocated cells.

Table 3

v_j	1	2	4	15
u_i 0		2		
5		7	9	20
3	4			18

Initial cost matrix for allocated cells.
Also write the values of u_i and v_j as explained earlier.

Table 3

1		4	15
6			
	5	7	

$u_i + v_j$ matrix for non–allocated cells

Table 5

10–1=9		20–4=16	11–15=–4
12–6=6			
	14–5=9	16–7=9	

Cell evaluation matrix

It can be seen from table 5 that cell evaluation at cell (1, 4) is negative i.e. -4, therefore by allocating at cell (1, 4) transportation cost be further reduced. Let us write down the original allocations and the proposed new allocation.

Table 6

	− 15		+
	+ t	15	− 10
5			5

It can be seen from table 6 that if we allocate at cell (1, 4) a loop is formed as shown and we allocate 10 units so that allocation at cell (2, 4) vanishes as shown below in table 7.

Table 7

15 – 10 = 5	+10
t + 10 = 10	10 – 10 = 0

New allocation Table will become

Table 8

	5		10
	10	15	
5			5

Transportation cost = $5X_2+10X_11+10X_7+15X9+5X_4+18+5$ =435 units. i.e. Transportation cost has come down from 475 units to 435 units.

Check for Optimaiity:

Let us see whether this solution is optima! or not? For that again two conditions have to be checked i.e.

No. of allocation = m + n – 1 =6 (satisfied)

Allocation at independent position (satisfied since closed loop for allocated cells is not formed)

Write cost at the allocated alls and values of u_i and v_j

Table 9

	-3	2	4	11
0		2		11
5		7	9	
7	4			18

Table 10

-3		4	
2			16
	9	11	

Cell evaluation metrix

Table 11

13		16	
10			4
	5	5	

Cell evaluation matrix

Since all the cell evaluations are the +ve therefore the solution was optimal *i.e.* 435 units cost of tran sportation.

Example 2:

(Unbalanced Supply and Demand). Solve the following transportation problem

		Stores 1	2	3	4	Supply
Factories	1	4	6	8	13	50
	2	13	11	10	8	70
	3	14	4	10	13	30
	4	9	11	13	8	50
Demand		25	35	105	20	

Total supply =200 units, Demand =185 units.

Solution:

Since supply and demand are not equal therefore problem is unbalanced. In order to balance the problem a dummy coloumn has to be added as shown below. Demand at that dummy coloumn (store) will be 15 units.

Table 1

Stores

		1	2	3	4	d	Supply
	A	4	6	8	13	0	50
Factories	B	13	11	10	8	0	70 Total supply =200
	C	14	4	10	13	0	30 Total demand = 200
	D	9	11	13	8	0	50
Demand		25	35	105	20	15	

Basic Feasible Solution:

We shall use Vogel's approximation method to find the initial feasible solution.

Table 2

Stores

		1	2	3	4	d	Supply
	A	$4_{(25)}$	$6_{(5)}$	$8_{(20)}$	13	0	50/25/20 [4] [2] [2] [2] [5]←
Factories	B	13	11	$10_{(70)}$	8	0	70 [8] [2] [2] [2] [2] [2]
	C	14	$4_{(30)}$	10	13	0	30 [4] [6]←
	D	9	11	$13_{(15)}$	$8_{(20)}$	$0_{(15)}$	50/35/15 [8] [1] [1] [3] [5] [5]←
Demand		25	35/5	105/85	20	15	
		[5]	[2]	[2]	[0]	[0]	
		[5]	[2]	[2]	[0]		
		[5]	[5]	[2]	[0]		
		↑	[5]	[2]	[0]		
			↑	[2]	[0]		
				[3]	[0]		

The initial feasible solution is given by the following matrix:

Table 3
Stores

		1	2	3	4	d	Supply
	A	$4_{(25)}$	$6_{(5)}$	$8_{(20)}$	13	0	50
Facotries	B	13	11	$10_{(70)}$	8	0	70
	C	14	$4_{(30)}$	10	13	0	30
	D	9	11	$13_{(15)}$	$8_{(20)}$	$0_{(15)}$	50
Demand		25	35	105	20	15	

Optimality Test:

From the above matrix we find that:

(a) Number of allocations = m + n – 1 = 4 + 5-1 = 8

(b) These m + n – 1 allocations are in independent positions.

Therefore optimality test can be performed. This consists of the sub steps explained earlier as shown in Tables below:

Table 4

	v_j	0	2	4	–1	–9
u_j	4	4	6	8		
	6			10		
	2		4			
	9			13	8	0

Matrix of *ui* + *vj* for occupied cells

Table 5

	v_j	0	2	4	–1	–9
u_j	4				3	–5
	6	6	8		5	–3
	2	2		6	1	–7
	9	9	11			

Matrix with cell values of *ui* + *vj* for empty cells

Table 6

			10	5
7	3		3	3
12		4	12	7
0	0			

Cell evaluation matrix

Since cell values are +ve the first feasible solution is optimal. Since table 6 contains zero entries, there exist alternate optimal solutions. The practical significance of demand being 15 units less than supply is that the company may cut down the production of 15 units at the factory where it is uneconomical.

The optimum (minimum) transportation plus production cost.
Z = Rs. (4 x 25 + 6 x 5 + 8 x 20 +10 x 70 + 4 x 30 + 13 x 15 + 8 x 20 + 0 x 15)
= Rs. (100 + 30 + 160 + 170 + 120 + 195 + 160 + 0) = Rs. 1,465.

Example 3:
Solve the following transportation problem for maximizing the profit. Due to difference in raw material cost and transportation cost, the profit for unit in rupees differs which is given in the table below:

Desinations

		1	2	3	4	
Origins	*A*	90	90	100	110	200
	B	50	70	130	85	100
		75	100	100	30	

Solve the problem for maximizing the profit.

Solution:
Problem is unbalanced and therefore a dummy row has to be added to make it balanced.

Table 1

	1	2	3	4	Supply
A	90	90	100	110	200
B	50	70	130	85	100
Dumny source	0	0	0	0	5
Demand	75	100	100	30	

Find Initial Basic Feasible solution:
We shall use vogel's approximation method to determine the initial feasible solution.
Note that we are dealing with maximization problem. Hence we shall enter the difference between the highest and second highest elements in each row to the right of the row and the difference between the highest and the second highest elements in each column below the corresponding column.

Each of these differences represents the unit profit lost for not allocating to the highest profit cell. Thus, while making allocations, at first we select cell (2, 3) with highest entry in row 2 which corresponds to the highest difference of [45].

Table 2

	1	2	3	4	Supply
A	$90_{(70)}$	$90_{(100)}$	100	$110_{(30)}$	200/170/70 [10] [20] [0]
B	50	70	$130_{(100)}$	85	100 [45]←
Dumny source	$0_{(5)}$	0	0	0	5 [0] [0] [0]
Demand	75/5	100	100	30	
	[40]	[20]	[30]	[25]	
	[90]	[90]		[110]	
	[90]	[90]		↑	
	↑	↑			

Optimality Test:
Required number of allocations = m + n – 1 = 3 + 4 – 1 = 6
Actual number of allocation = 5.
Therefore we allocate small positive number € to cell (1, 3) (cell having maximum profit out of vacant cells) so that the number of allocations becomes 6. These 6 allocations are in independent positions. Therefore optimality test can be performed.

Table 3

	1	2	3	4	Supply
A	$90_{(70)}$	$90_{(100)}$	$100_{(e)}$	$110_{(30)}$	200
B	50	70	$130_{(100)}$	85	100
Dumny source	$0_{(5)}$	0	0	0	5
Demand	75	100	100	30	

Table 4

v_j	0	0	10	20
u_i 90	90	90	100	110
120			130	
0	0			

Matrix of $u_i + v_j$ for allocated cells

Table 5

v_j	0	0	10	20
u_i 90				
120	120	120		140
0		0	10	20

Matrix of $u_i + v_j$ for vacant cells

Table 6

–70	–50		–55
	0	–10	–20

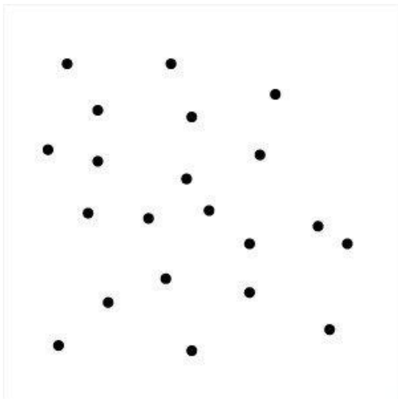
Cell evaluation matrix

Since all cell values are either negative or zero (maximization problem), the initial basic feasible solution is optimal. The demand at first destination is ‘ left unsatisfied by 5units. The profit is

TRAVELLING SALESMAN PROBLEM.

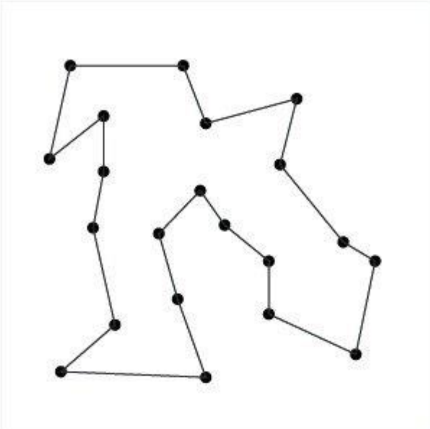
The Travelling Salesman Problem (TSP) is the challenge of finding the shortest yet most efficient route for a person to take given a list of specific destinations. It is a well-known algorithmic problem in the fields of computer science and operations research.

There are obviously a lot of different routes to choose from, but finding the best one—the one that will require the least distance or cost—is what mathematicians and computer scientists have spent decades trying to solve for.



TSP has commanded so much attention because it's so easy to describe yet so difficult to solve. In fact, TSP belongs to the class of combinatorial optimization problems known as NP-complete. This means that TSP is classified as NP-hard because it has no "quick" solution and the complexity of calculating the best route will increase when you add more destinations to the problem.

The problem can be solved by analyzing every round-trip route to determine the shortest one. However, as the number of destinations increases, the corresponding number of roundtrips surpasses the capabilities of even the fastest computers. With 10 destinations, there can be more than 300,000 roundtrip permutations and combinations. With 15 destinations, the number of possible routes could exceed 87 billion.



Real-life TSP and VRP solvers use route optimization algorithms that find near-optimal solutions in a fraction of the time, giving delivery businesses the ability to plan routes quickly and efficiently. Check out our solutions!

Popular Travelling Salesman Problem Solutions

Here are some of the most popular solutions to the Traveling Salesman Problem:

The Brute-Force Approach

The Brute Force approach, also known as the Naive Approach, calculates and compares all possible permutations of routes or paths to determine the shortest unique solution. To solve the TSP using the Brute-Force approach, you must calculate the total number of routes and then draw and list all the possible routes. Calculate the distance of each route and then choose the shortest one—this is the optimal solution.

The Branch and Bound Method

This method breaks a problem to be solved into several sub-problems. It's a system for solving a series of sub-problems, each of which may have several possible solutions and where the solution selected for one problem may have an effect on the possible solutions of subsequent sub-problems. To solve the TSP using the Branch and Bound method, you must choose a start node and then set bound to a very large value (let's say infinity). Select the cheapest arch between the unvisited and current node and then add the distance to the current distance. Repeat the process while the current distance is less then the bound. If the current distance is less than the bound, you're done. You may now add up the distance so that the bound will be equal to the current distance. Repeat this process until all the arcs have been covered.

The Nearest Neighbor Method

This is perhaps the simplest TSP heuristic. The key to this method is to always visit the nearest destination and then go back to the first city when all other cities are visited. To solve the TSP using this method, choose a random city and then look for the closest unvisited city and go there. Once you have visited all cities, you must return to the first city.

Academic Solutions to TSP

Academics have spent years trying to find the best solution to the Travelling Salesman Problem The following solutions were published in recent years:

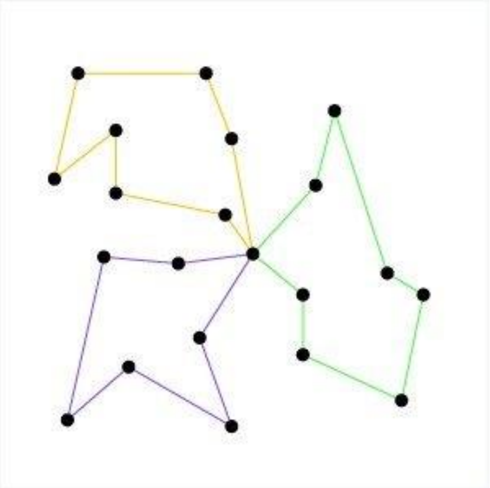
- Zero Suffix Method: Developed by Indian researchers, this method solves the classical symmetric TSP.
- Biogeography-based Optimization Algorithm: This method is designed based on the animals' migration strategy to solve the problem of optimization.

- Meta-Heuristic Multi Restart Iterated Local Search (MRSILS): The proponents of this research asserted that the meta-heuristic MRSILS is more efficient than the Genetic Algorithms when clusters are used.
- Multi-Objective Evolutionary Algorithm: This method is designed for solving multiple TSP based on NSGA-II.
- Multi-Agent System: This system is designed to solve the TSP of N cities with fixed resource.

Real-world TSP Applications

Despite the complexity of solving the Travelling Salesman Problem, it still finds applications in all verticals.

For instance, efficient solutions found through the TSP are being used in the last mile delivery. Last mile delivery refers to the movement of goods from a transportation hub, such as a depot or a warehouse, to the end customer's choice of delivery. Last mile delivery is the leading cost driver in the supply chain. Companies usually shoulder some of the costs to better compete in the market. In fact, a last mile delivery costs the company an average of $10.1, but the customer only pays an average of $8.08. This is the reason why businesses strive to minimize the cost of last mile delivery.



The minimization of costs in last mile delivery is essentially a **Vehicle Routing Problem (VRP)**. VRP is a generalized version of the TSP and is one of the most widely studied problems in mathematical optimization. It deals with finding a set of routes or paths to reduce delivery costs. The problem domain may involve a set of depot locations, hundreds of delivery locations, and several vehicles. As with TSP, determining the best solution to VRP is NP-hard, so the number of problems that can be solved, optimally, using combinatorial optimization or mathematical programming may be limited. Thus, commercial solvers usually use heuristics—these are like shortcuts for our brain, eliminating a lot of math or calculations for a quick and easy solution—due to the frequency and size of real world VRPs they need to solve.

Real-life TSP and VRP Solvers

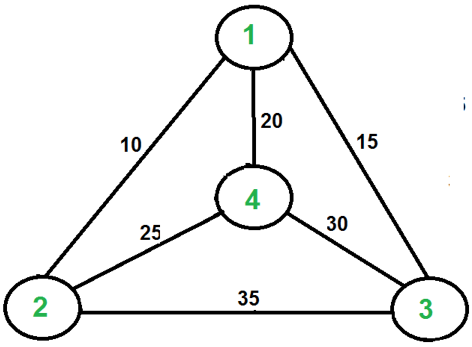
While academic solutions to TSP and VRP aim to provide ***the*** optimal solution to these NP-hard problems, many of them aren't practical when solving real world problems, especially when it comes to solving last mile logistical challenges.

That's because academic solvers strive for perfection and thus take a long time to compute the optimal solutions – hours, days, and sometimes years. If a delivery business needs to plan daily routes, they need a route solution within a matter of minutes. Their business depends on those route solutions so they can get their drivers and their goods out the door as soon as possible.

Real-life TSP and VRP solvers use route optimization algorithms that find a near-optimal solutions in a fraction of the time, giving delivery businesses the ability to plan routes quickly and efficiently.

Travelling Salesman Problem (TSP):

Given a set of cities and distance between every pair of cities, the problem is to find the shortest possible route that visits every city exactly once and returns to the starting point.



Note the difference between Hamiltonian Cycle and TSP. The Hamiltoninan cycle problem is to find if there exist a tour that visits every city exactly once. Here we know that Hamiltonian Tour exists (because the graph is complete) and in fact many such tours exist, the problem is to find a minimum weight Hamiltonian Cycle.

For example, consider the graph shown in figure on right side. A TSP tour in the graph is 1-2-4-3-1. The cost of the tour is 10+25+30+15 which is 80.
The problem is a famous NP hard problem. There is no polynomial time know solution for this problem.
Following are different solutions for the traveling salesman problem.

Naive Solution:

1) Consider city 1 as the starting and ending point.
2) Generate all (n-1)! Permutations of cities.
3) Calculate cost of every permutation and keep track of minimum cost permutation.
4) Return the permutation with minimum cost.

Time Complexity: $\Theta(n!)$

Dynamic Programming:

Let the given set of vertices be {1, 2, 3, 4,….n}. Let us consider 1 as starting and ending point of output. For every other vertex i (other than 1), we find the minimum cost path with 1 as the starting point, i as the ending point and all vertices appearing exactly once. Let the cost of this path be cost(i), the cost of corresponding Cycle would be cost(i) + dist(i, 1) where dist(i, 1) is the distance from i to 1. Finally, we return the minimum of all [cost(i) + dist(i, 1)] values. This looks simple so far. Now the question is how to get cost(i)?
To calculate cost(i) using Dynamic Programming, we need to have some recursive relation in terms of sub-problems. Let us define a term *C(S, i) be the cost of the minimum cost path visiting each vertex in set S exactly once, starting at 1 and ending at i*.
We start with all subsets of size 2 and calculate C(S, i) for all subsets where S is the subset, then we calculate C(S, i) for all subsets S of size 3 and so on. Note that 1 must be present in every subset.
If size of S is 2, then S must be {1, i},
C(S, i) = dist(1, i)
Else if size of S is greater than 2.
C(S, i) = min { C(S-{i}, j) + dis(j, i)} where j belongs to S, j != i and j != 1.
For a set of size n, we consider n-2 subsets each of size n-1 such that all subsets don't have nth in them.
Using the above recurrence relation, we can write dynamic programming based solution. There are at most $O(n*2^n)$ subproblems, and each one takes linear time to solve. The total running time is therefore $O(n^2*2^n)$. The time complexity is much less than O(n!), but still exponential. Space required is also exponential. So this approach is also infeasible even for slightly higher number of vertices.

UNIT – III
DYNAMIC PROGRAMMING

Learning Objectives
Dynamic Programming – Applications of D.P. (Capital Budgeting, Production Planning, Solving Linear Programming Problem) – Integer Programming – Branch and Bound Method.

DYNAMIC PROGRAMMING

Dynamic Programming (DP) is an algorithmic technique for solving an optimization problem by breaking it down into simpler subproblems and utilizing the fact that the optimal solution to the overall problem depends upon the optimal solution to its subproblems.

Insertion sort is an **example** of **dynamic programming**, selection sort is an **example** of greedy algorithms, Merge Sort and Quick Sort are **example** of divide and conquer. So, different categories of algorithms may be used for accomplishing the same goal - in this case, sorting.

Dynamic programming (DP) is an effective method for finding an optimum solution to a multi-stage decision problem, as long as the solution can be defined recursively. DP roots in the **principle of optimality** proposed by Richard Bellman:

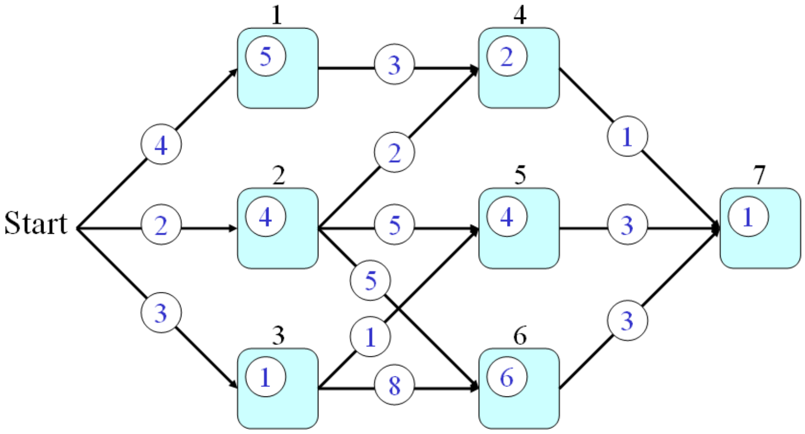
An optimal policy has the property that whatever the initial state and the initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

In terms of path finding problems, the principle of optimality states that any partial path of the optimum path is also an optimum path given the starting and ending nodes. This is an obvious principle that can be proved by contradiction. In practice, a number of seemingly unrelated applications can be effectively solved by DP.

To employ the procedure of DP to solve a problem, usually we need to specify the following items:

I. Define the optimum-value function.
II. Derive the recursive formula for the optimum-value function, together with its initial condition.
III. Specify the answer of the problem in term of the optimum-value function.

We shall give an example of DP for a path finding problem shown in the following graph:



In the above graph, we assume that:

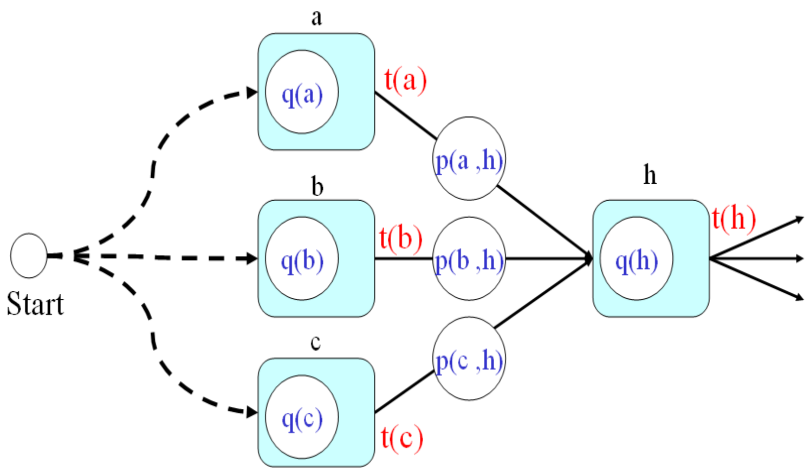
- Every node is a city and we use q(a) to represent the time required to go through node a, where a is in {1, 2, 3, 4, 5, 6, 7}.
- Every link is a route connecting two city. We use p(a, b) to represent the time required to go from nodes a to b, where a and b are in {1, 2, 3, 4, 5, 6, 7}.

From the start point, how can we find a path that requires the minimum time to pass through node 7? This is a typical problem of DP which can be solved systematically and efficiently by the following three steps:

I. First of all, we can define the optimum-value function t(h) as the minimum time from the start point to node h (including the time for passing node h).
II. Secondly, the optimum-value function should satisfy the following recursive formula:

$$t(h) = \min\{t(a)+p(a,h),\ t(b)+p(b,h),\ t(c)+p(c,h)\} + q(h)$$

In the above equation, we assume the fan-ins of node h are a, b, and c. Please refer to the next figure.



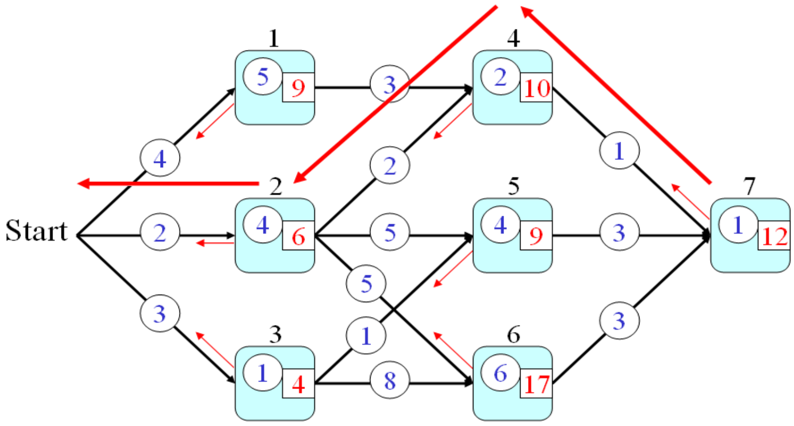
- And the initial condition is t(0)=0 where 0 is the start node.

- And finally, the answer to the original problem should be t(7). By using the recursion, we can have the following steps for computing the time required by the optimum path:

1. t(0) = 0
2. t(1) = t(0)+4+5 = 9
3. t(2) = t(0)+2+4 = 6
4. t(3) = t(0)+3+1 = 4
5. t(4) = min(9+3, 6+2)+2 = 10

6. t(5) = min(6+5, 4+1)+4 = 9
7. t(6) = min(6+5, 4+8)+6 = 17
8. t(7) = min(10+1, 9+3, 17+3)+1 = 12

The value of the optimum-value function is shown as the red number in each node in the following figure:

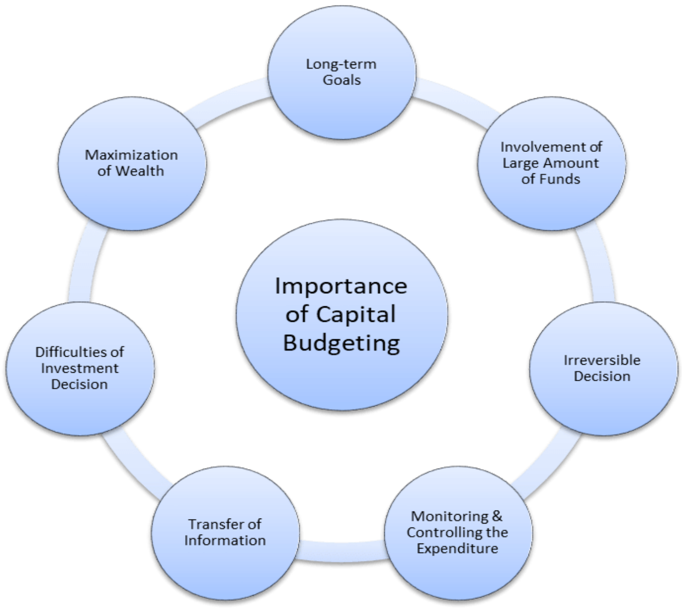


CAPITAL BUDGETING

Capital budgeting is the process that a business uses to determine which proposed fixed asset purchases it should accept, and which should be declined. This process is used to create a quantitative view of each proposed fixed asset investment, thereby giving a rational basis for making a judgment.

Capital budgeting is a process that helps in planning the investment projects of an organization in long run. It takes all possible consideration into account so that the company can evaluate the profitability of the project. It is useful for evaluating capital investment project such as purchasing equipment, the rebuilding of equipment etc. The benefit from an investment may be in form of a reduction in cost or in form of increased revenue. Importance of capital budgeting can be understood from its impact on the business.

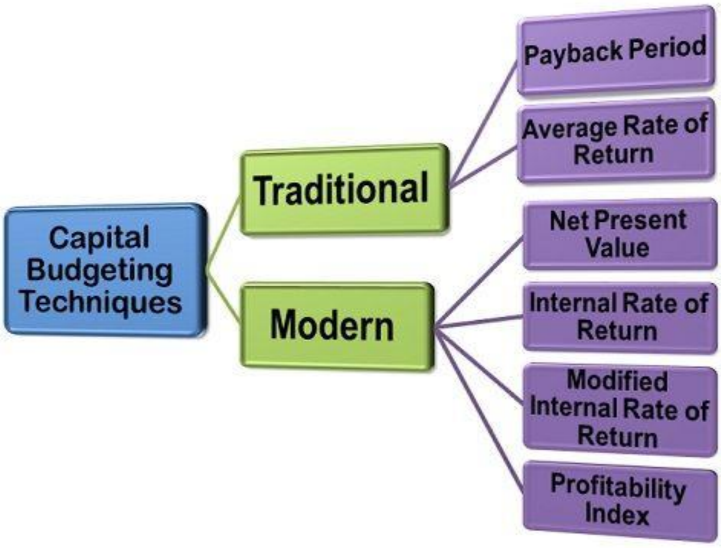
Businesses exist to earn profit except for non-profit organization. Capital budgeting is very important for any business as it impacts the growth & prosperity of the business in the long term. It creates accountability & measurability. Some of the popular techniques are net present value, internal rate of return, payback period, accounting rate of return & profitability index.



Since the capital budgeting is related to the long-term investments whose returns will be fetched in the future, certain **traditional and modern** capital budgeting techniques are employed by the firm to judge the feasibility of these projects.

The traditional method relies on the non-discounting criteria that do not consider the time value of money, whereas the modern method includes the discounting criteria where the time value of money is taken into the consideration.

Capital Budgeting Techniques



Traditional methods
The traditional methods comprise of the following evaluation techniques:

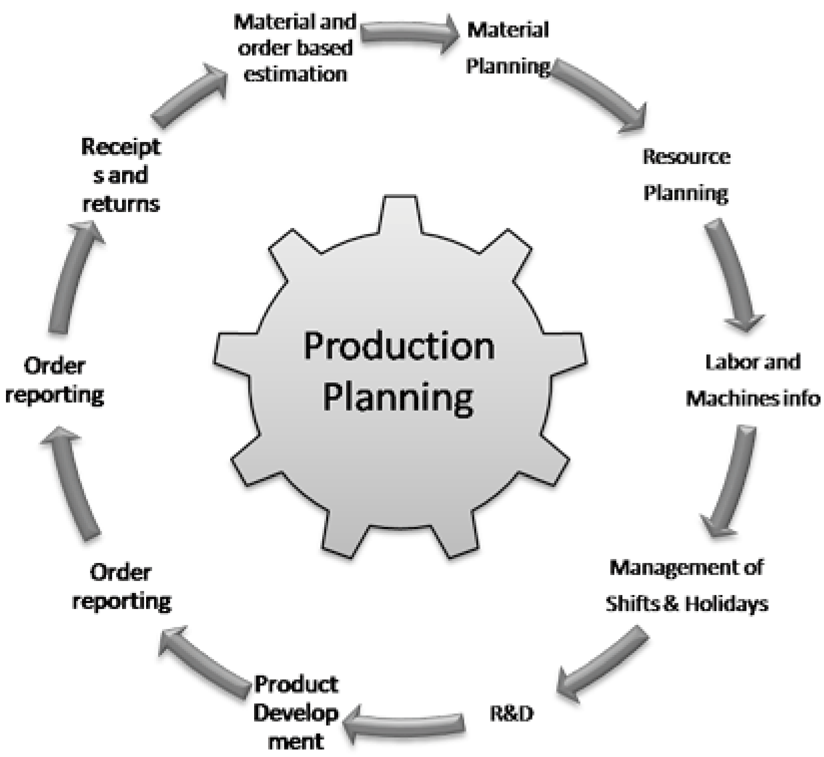
1. Payback Period Method
2. Average Rate of Return or Accounting Rate of Return Method

Modern Methods
The modern methods comprise of the following evaluation techniques:

1. Net Present Value Method
2. Internal Rate of Return
3. Modified Internal Rate of Return
4. Profitability Index

The common thing about both these methods (Traditional and Modern) is that these are based on the cash inflows and the outflows of the project.

PRODUCTION PLANNING



Production planning is the **planning** of **production** and **manufacturing** modules in a company or industry. It utilizes the resource allocation of activities of employees, materials and **production** capacity, in order to serve different customers.

Production planning is the act of developing a guide for the design and production of a given product or service. Production planning helps organizations make the production process as efficient as possible. Production planning originated to optimize the manufacturing process, and today its general logic is applied in various forms to design, production and delivery of software as well.

Why is production planning important?

Production planning is important because it creates an efficient process for production according to customer and organizational needs. It optimizes both customer-dependent processes -- such as on-time delivery -- and customer-independent processes, such as production cycle time.

A good production plan minimizes lead time, which is the amount of time that passes between the placing of an order and the completion and delivery of that order. Depending on the company and the type of production planning necessary, the definition of lead time varies slightly. In supply chain management, for example, lead time includes the amount of time it takes for parts to ship from a supplier. This is included because the manufacturing business needs to know when the parts will arrive to properly execute material requirements planning (MRP). This is especially important with tight manufacturing constraints or just-in-time (JIT) manufacturing.

Production planning process

The production planning process involves the following steps:

- Estimate product demand -- This will a give a rough outline of how many products should be produced in a given time period. This estimate is generated by combining analysis of historical production trends with new potentially relevant trends in the market.
- Weigh production options -- This involves accounting for the resources on hand and exploring ways to most effectively use them based on projected demand estimates.

- Choose the most efficient option -- The use of resources that is the least costly and most time-efficient should be chosen.
- Monitoring and evaluation -- As the plan is carried out, companies monitor what is happening compared to what should be happening according to the plan, and evaluate how well those two match up.
- Adjust plan -- This involves altering the plan so that future production plans meet customer goals more efficiently and are more successful in their execution.

Types of production planning

There are many types of production planning that focus on various particulars of the production process. Some of these include:

- Master production schedule (MPS) -- These are schedules for individual, specific commodities to be produced in a given time period. They are often generated by software, and then adjusted by users.
- Material requirements planning -- MRP is a system used for production planning, scheduling and inventory control. MRP ensures the availability of raw materials, maintains the lowest possible material and product levels in-house, and plans manufacturing and purchasing activities. It is often automated to some extent by software, but can be performed completely manually as well.
- Capacity planning -- This is the process of determining what capacity an organization has to meet changing demands.
- Workflow planning -- This is the planning of a sequence of operations performed by an employee or group of employees.

There are also various planning types that apply the logic of production planning to areas other than manufacturing, or complementary areas. For example, human resources planning involves optimizing processes that allow a company to meet their hiring and talent demands. Other examples include:

- Enterprise resource planning (ERP) -- This is the integration of main business processes into one unified system, often through the use of software.
- Sales and operations planning (S&OP) -- This is the process for more accurately matching a manufacturer's supply with existing demand.

Production scheduling

Production scheduling is like production control. Production scheduling is the allocation of available resources to production processes and events. It is essentially the mapping of actual resources to the production plan built for them. Scheduling is used to plan use of factory equipment and resources, human resources, and to plan processes and material purchasing. Scheduling is necessary to create a production plan. Production plans aim to ultimately deliver on customer demand. The goal of a production schedule is to create the most efficient production plan possible.

History of production planning

Modern production planning has its roots in the first half of the 19th century. It developed out of a need for information around internal planning in control. Entities like railroads, textile mills and other factories needed internal administrative frameworks to guide the multiple processes involved in providing their basic product or service at a large scale.

The first production plans were simple. Factories were relatively small and produced a limited number of products in large batch sizes. Factory foremen were technical experts in their field, and handled all planning and scheduling, which sometimes would include no more than a list of production orders and the date at which they were to be completed.

As the production line and manufacturing efforts as a whole became bigger and more complex, more involved production planning was necessary. By the beginning of the 20th century, plans began focusing on not just delivering orders, but optimizing the processes required to do so, so that production process flow could be as even as possible at the minimum possible production cost.

Today, as the nature of production methods and manufacturing has changed, so has production planning. The technology surrounding production has evolved, enabling more precise communication and monitoring of and around production. The products themselves, and customer expectations, have also evolved. Now, there is more information available than ever before for organizations to weigh when creating their production plans.

INTEGER PROGRAMMING

When formulating LP's we often found that, strictly, certain variables should have been regarded as taking integer values but, for the sake of convenience, we let them take fractional values reasoning that the variables were likely to be so large that any fractional part could be neglected. Whilst this is acceptable in some situations, in many cases it is not, and in such cases we must find a numeric solution in which the variables take integer values.

Problems in which this is the case are called *integer programs* (*IP's*) and the subject of solving such programs is called *integer programming* (also referred to by the initials *IP*).

IP's occur frequently because many decisions are essentially discrete (such as yes/no, go/no-go) in that one (or more) options must be chosen from a finite set of alternatives.

Note here that problems in which some variables can take only integer values and some variables can take fractional values are called *mixed-integer programs (MIP's).*

As for formulating LP's the key to formulating IP's is *practice*. Although there are a number of standard "tricks" available to cope with situations that often arise in formulating IP's it is probably true to say that formulating IP's is a much harder task than formulating LP's.

We consider an example integer program below.

Capital budgeting

There are four possible projects, which each run for 3 years and have the following characteristics.

```
                         Capital requirements (£m)
Project   Return (£m)    Year   1      2      3
```

1	0.2	0.5	0.3	0.2
2	0.3	1.0	0.8	0.2
3	0.5	1.5	1.5	0.3
4	0.1	0.1	0.4	0.1
Available capital (£m)		3.1	2.5	0.4

We have a **decision problem** here: **Which projects would you choose in order to maximise the total return?**

Capital budgeting solution

We follow the same approach as we used for formulating LP's - namely:

- variables
- constraints
- objective.

We do this below and note here that the only significant change in formulating IP's as opposed to formulating LP's is in the definition of the variables.

Variables

Here we are trying to decide whether to undertake a project or not (a "go/no-go" decision). One "trick" in formulating IP's is to introduce variables which take the integer values 0 *or* 1 and represent *binary* decisions (e.g. do a project or not do a project) with typically:

- the positive decision (do something) being represented by the value 1; and
- the negative decision (do nothing) being represented by the value 0.

Such variables are often called *zero-one* or *binary* variables

To define the variables we use the *verbal* description of

x_j = 1 if we decide to do project j (j=1,...,4)
   = 0 otherwise, i.e. not do project j (j=1,...,4)

Note here that, by definition, the x_j are *integer* variables which must take one of two possible values (zero or one).

Constraints

The constraints relating to the availability of capital funds each year are

$0.5x_1 + 1.0x_2 + 1.5x_3 + 0.1x_4 <= 3.1$ (year 1)
$0.3x_1 + 0.8x_2 + 1.5x_3 + 0.4x_4 <= 2.5$ (year 2)
$0.2x_1 + 0.2x_2 + 0.3x_3 + 0.1x_4 <= 0.4$ (year 3)

Objective

To maximise the total return - hence we have maximise $0.2x_1 + 0.3x_2 + 0.5x_3 + 0.1x_4$

This gives us the complete IP which we write as

maximise $0.2x_1 + 0.3x_2 + 0.5x_3 + 0.1x_4$

subject to

$0.5x_1 + 1.0x_2 + 1.5x_3 + 0.1x_4 <= 3.1$
$0.3x_1 + 0.8x_2 + 1.5x_3 + 0.4x_4 <= 2.5$
$0.2x_1 + 0.2x_2 + 0.3x_3 + 0.1x_4 <= 0.4$
$x_j = 0$ or 1 j=1,...,4

Note:

- in writing down the complete IP we include the information that $x_j = 0$ or 1 (j=1,...,4) as a reminder that the variables are integers
- you see the usefulness of defining the variables to take zero/one values - e.g. in the objective the term $0.2x_1$ is zero if $x_1=0$ (as we want since no return from project 1 if we do not do it) and 0.2 if $x_1=1$ (again as we want since get a return of 0.2 if we do project 1). Hence effectively the zero-one nature of the decision variable means that we always capture in the single term $0.2x_1$ what happens both when we do the project and when we do not do the project.
- you will note that the objective and constraints are linear (i.e. any term in the constraints/objective is either a constant or a constant multiplied by an unknown). In this course we deal only with linear integer programs (IP's with a linear objective and linear constraints). It is plain though that there do exist non-linear integer programs - these are, however, outside the scope of this course.
- whereas before in formulating LP's if we had integer variables we assumed that we could ignore any fractional parts it is clear that we cannot do so in this problem e.g. what would be the physical meaning of a numeric solution with $x_1=0.4975$ for example?

Extensions to this basic problem include:

- projects of different lengths
- projects with different start/end dates
- adding capital inflows from completed projects
- projects with staged returns
- carrying unused capital forward from year to year
- mutually exclusive projects (can have one or the other but not both)
- projects with a time window for the start time.

How to amend our basic IP to deal with such extensions is given here.

In fact note here that integer programming/quantitative modelling techniques are increasingly being used for financial problems.

Solving IP's

For [solving LP's]() we have *general purpose* (independent of the LP being solved) and *computationally effective* (able to solve large LP's) algorithms (simplex or interior point).

For solving IP's *no* similar general purpose and computationally effective algorithms exist.

Indeed theory *suggests* that no general purpose computationally effective algorithms will ever be found. This area is known as *computational complexity* and concerns *NP-completeness*. It was developed from the early 1970's onward and basically is a theory concerning "how long it takes algorithms to run". This means that IP's are a lot harder to solve than LP's.

Solution methods for IP's can be categorised as:

- *general purpose* (will solve any IP) but potentially computationally ineffective (will only solve relatively small problems); or
- *special purpose* (designed for one particular type of IP problem) but potentially computationally more effective.

Solution methods for IP's can also be categorised as:

- *optimal*
- *heuristic*

An optimal algorithm is one which (mathematically) *guarantees* to find the optimal solution.

It may be that we are not interested in the optimal solution:

- because the size of problem that we want to solve is beyond the computational limit of known optimal algorithms within the computer time we have available; or
- we could solve optimally but feel that this is not worth the effort (time, money, etc) we would expend in finding the optimal solution.

In such cases we can use a heuristic algorithm - that is an algorithm that should hopefully find a feasible solution which, in objective function terms, is close to the optimal solution. In fact it is often the case that a well-designed heuristic algorithm can give good quality (near-optimal) results.

For example a heuristic for our capital budgeting problem would be:

- consider each project in turn
- decide to do the project if this is feasible in the light of previous decisions

Applying this heuristic we would choose to do just project 1 and project 2, giving a total return of 0.5, which may (or may not) be the optimal solution.

Hence we have four categories that we potentially need to consider:

- general purpose, optimal
- general purpose, heuristic

- special purpose, optimal
- special purpose, heuristic.

Note here that the methods presented below are suitable for solving both IP's (all variables integer) and MIP's (mixed-integer programs - some variables integer, some variables allowed to take fractional values).

General purpose optimal solution algorithms

We shall deal with just two general purpose (able to deal with any IP) optimal solution algorithms for IP's:

- enumeration (sometimes called complete enumeration)
- branch and bound (tree search).

We consider each of these in turn below. Note here that there does exist another general purpose solution algorithm based upon *cutting planes* but this is beyond the scope of this course.

Enumeration

Unlike LP (where variables took continuous values (>=0)) in IP's (where all variables are integers) each variable can only take a finite number of discrete (integer) values.

Hence the obvious solution approach is simply to *enumerate* all these possibilities - calculating the value of the objective function at each one and choosing the (feasible) one with the optimal value.

For example for the capital budgeting problem considered above there are 2^4=16 possible solutions. These are:

x_1	x_2	x_3	x_4	
0	0	0	0	do no projects
0	0	0	1	do one project
0	0	1	0	
0	1	0	0	
1	0	0	0	
0	0	1	1	do two projects
0	1	0	1	
1	0	0	1	
0	1	1	0	
1	0	1	0	
1	1	0	0	
1	1	1	0	do three projects
1	1	0	1	
1	0	1	1	
0	1	1	1	
1	1	1	1	do four projects

Hence for our example we merely have to examine 16 possibilities before we know precisely what the best possible solution is. This example illustrates a general truth about integer programming:

What makes solving the problem easy when it is small is precisely what makes it become very hard very quickly as the problem size increases

This is simply illustrated: suppose we have 100 integer variables each with 2 possible integer values then there are 2x2x2x ... x2 = 2^{100} (approximately 10^{30}) possibilities which we have to enumerate (obviously many of these possibilities will be infeasible, but until we generate one we cannot check it against the constraints to see if it is feasible or not).

This number is plainly too many for this approach to solving IP's to be computationally practicable. To see this consider the fact that the universe is around 10^{10} years old so we would need to have considered 10^{20} possibilities per year, approximately 4×10^{12} possibilities per second, to have solved such a problem by now if we started at the beginning of the universe.

Be clear here - **conceptually** there is not a problem - simply enumerate all possibilities and choose the best one. But **computationally** (numerically) this is just impossible.

IP nowadays is often called "*combinatorial optimisation*" indicating that we are dealing with optimisation problems with an extremely large (combinatorial) increase in the number of possible solutions as the problem size increases.

Branch and bound (tree search)

The most effective general purpose optimal algorithm is LP-based tree search (tree search also being called branch and bound). *This is a way of systematically enumerating feasible solutions such that the optimal integer solution is found.*

Where this method differs from the enumeration method is that *not all* the feasible solutions are enumerated but only a fraction (hopefully a small fraction) of them. However we can still *guarantee* that we will find the optimal integer solution. The method was first put forward in the early 1960's by Land and Doig.

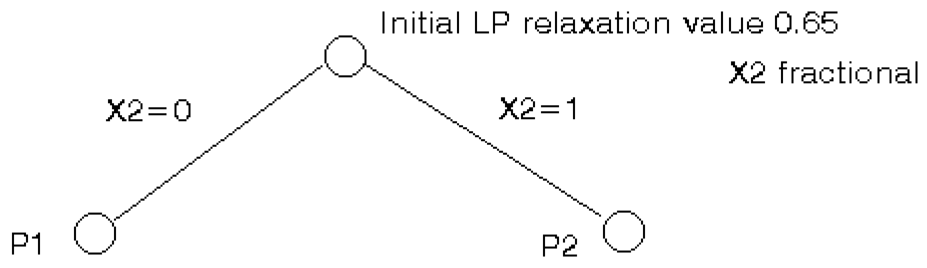
Consider our example capital budgeting problem. What made this problem difficult was the fact that the variables were restricted to be integers (zero or one). If the variables had been allowed to be fractional (takes all values between zero and one for example) then we would have had an LP which we could easily solve. Suppose that we were to solve this LP relaxation of the problem [replace x_j = 0 or 1 j=1,...,4 by 0 <= x_j <= 1 j=1,...,4]. Then using the package we get x_2=0.5, x_3=1, x_1=x_4=0 of value 0.65 (i.e. the objective function value of the optimal linear programming solution is 0.65).

As a result of this we now know something about the optimal integer solution, namely that it is <= 0.65, i.e. this value of 0.65 is a (upper) *bound* on the optimal integer solution. This is because when we relax the integrality constraint we (as we are maximising) end up with a solution value at least that of the optimal integer solution (and maybe better).

Consider this LP relaxation solution. We have a variable x_2 which is fractional when we need it to be integer. **How can we rid ourselves of this troublesome fractional value?** To remove this troublesome fractional value we can generate two new problems:

- original LP relaxation plus $x_2=0$
- original LP relaxation plus $x_2=1$

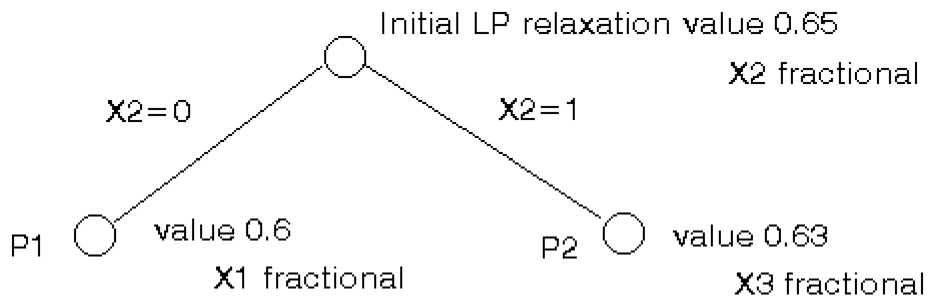
then we will claim that the optimal integer solution to the original problem is contained in one of these two new problems. This process of taking a fractional variable (a variable which takes a fractional value in the LP relaxation) and explicitly constraining it to each of its integer values is known as ***branching***. It can be represented diagrammatically as below (in a tree diagram, which is how the name ***tree search*** arises).



We now have two new LP relaxations to solve. If we do this we get:

- P1 - original LP relaxation plus $x_2=0$, solution $x_1=0.5$, $x_3=1$, $x_2=x_4=0$ of value 0.6
- P2 - original LP relaxation plus $x_2=1$, solution $x_2=1$, $x_3=0.67$, $x_1=x_4=0$ of value 0.63

This can be represented diagrammatically as below.

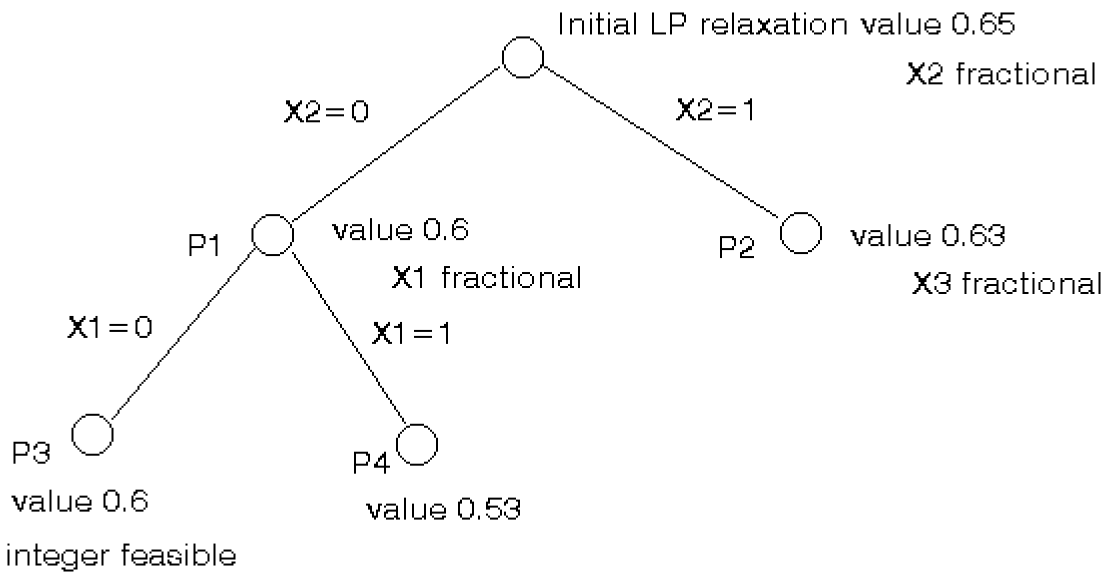


To find the optimal integer solution we just repeat the process, choosing one of these two problems, choosing one fractional variable and generating two new problems to solve.

Choosing problem P1 we branch on x_1 to get our list of LP relaxations as:

- P3 - original LP relaxation plus $x_2=0$ (P1) plus $x_1=0$, solution $x_3=x_4=1$, $x_1=x_2=0$ of value 0.6
- P4 - original LP relaxation plus $x_2=0$ (P1) plus $x_1=1$, solution $x_1=1$, $x_3=0.67$, $x_2=x_4=0$ of value 0.53
- P2 - original LP relaxation plus $x_2=1$, solution $x_2=1$, $x_3=0.67$, $x_1=x_4=0$ of value 0.63

This can again be represented diagrammatically as below.



At this stage we have identified a integer feasible solution of value 0.6 at P3. There are no fractional variables so no branching is necessary and P3 can be dropped from our list of LP relaxations.

Hence we now have new information about our optimal (best) integer solution, namely that it lies between 0.6 and 0.65 (inclusive).

Consider P4, it has value 0.53 and has a fractional variable (x_3). However if we were to branch on x_3 any objective function solution values we get after branching can never be better (higher) than 0.53. As we already have an integer feasible solution of value 0.6 P4 can be dropped from our list of LP relaxations since branching from it could never find an improved feasible solution. This is known as ***bounding*** - using a known feasible solution to identify that some relaxations are not of any interest and can be discarded.

Hence we are just left with:

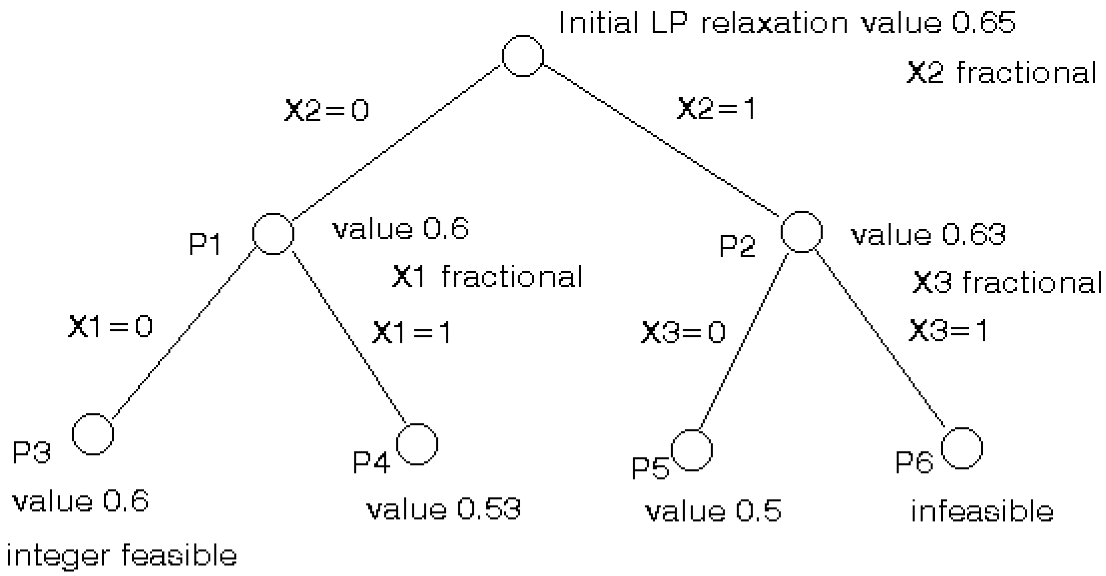
- P2 - original LP relaxation plus x_2=1, solution x_2=1, x_3=0.67, x_1=x_4=0 of value 0.63

Branching on x_3 we get

- P5 - original LP relaxation plus x_2=1 (P2) plus x_3=0, solution x_1=x_2=1, x_3=x_4=0 of value 0.5
- P6 - original LP relaxation plus x_2=1 (P2) plus x_3=1, problem infeasible

Neither of P5 or P6 lead to further branching so we are done, we have discovered the optimal integer solution of value 0.6 corresponding to x_3=x_4=1, x_1=x_2=0.

The entire process we have gone through to discover this optimal solution (and to prove that it is optimal) is shown graphically below.



You should be clear as to why 0.6 is the optimal integer solution for this problem, simply put if there were a better integer solution the above tree search process would (logically) have found it.

Note here that this method, like complete enumeration, also involves powers of two as we progress down the (binary) tree. However also note that we did not enumerate all possible integer solutions (of which there are 16). Instead here we solved 7 LP's. This is an important point, and indeed why tree search works at all. **We do not need to examine as many LP's as there are possible solutions.** Whilst the computational efficiency of tree search differs for different problems it is this basic fact that enables us to solve problems that would be completely beyond us were we to try complete enumeration.

You may have noticed that in the example above we never had more than one fractional variable in the LP solution at any tree node. This arises due to the fact that in constructing the above example I decided to make the situation as simple as possible. In general we might well have more than one fractional variable at a tree node and so we face a decision as to which variable to choose to branch on. A simple rule for deciding might be to take the fractional variable which is closest in value to 0.5, on the basis that the two branches (setting this variable to zero and one respectively) may well perturb the situation significantly.

Good computer packages (solvers) exist for finding optimal solutions to IP's/MIP's via LP-based tree search. Many of the computational advances in IP optimal solution methods (e.g. constraint aggregation, coefficient reduction, problem reduction, automatic generation of valid inequalities) are included in these packages. Often the key to making successful use of such packages for any particular problem is to put effort into a good formulation of the problem in terms of the variables and constraints. By this we mean that for any particular IP there may be a number of valid formulations. Deciding which formulation to adopt in a solution algorithm is often a combination of experience and trial and error.

Constraint logic programming (CLP), also called constraint programming, which is essentially branch and bound but without the bound, can be of use here if:

- the problem cannot be easily expressed in linear mathematics

- the gap between the LP relaxation solution and the IP optimal solution is so large as to render LP-based tree search impracticable.

Currently there is a convergence between CLP and LP-based solvers with ILOG and CPLEX merging.

General purpose heuristic solution algorithms

Essentially the only effective approach here is to run a general purpose optimal algorithm and terminate early (e.g. after a specified computer time).

Special purpose optimal solution algorithms

If we are dealing with one specific type of IP then we might well be able to develop a special purpose solution algorithm (designed for just this one type of IP) that is more effective computationally than the general purpose branch and bound method given earlier. These are typically tree search approaches based upon generating bounds via:

- dual ascent
- lagrangean relaxation and:
    - subgradient optimisation; or
    - multiplier adjustment.

Such algorithms, being different for different problems, are really beyond the scope of this course but suffice to say:

- such algorithms draw upon the concepts, such as branch and bound, outlined previously
- such algorithms often use linear programming via LP relaxation
- such algorithms take advantage of the ***structure*** of the constraints of the IP they are solving.
- different methods have different computational performance and their general behaviour is that each method is computationally effective up to a certain size of problem and then becomes computationally ineffective (as the effort needed to obtain an optimal solution begins to increase *exponentially* in terms of the size of problem considered).

A large amount of academic effort in this field is devoted to generating methods that out-perform previous methods for the same problem.

On a personal note this is an area with which I am familiar and special purpose algorithms can be very effective (e.g. a problem dealing with the location of warehouses and involving some 500,000 continuous variables, 500 zero-one variables and 500,000 constraints solved in some 20 minutes). They can be very successful compared with general purpose optimal algorithms (perhaps an order of magnitude or more in terms of the size of problem that can be solved).

You should be clear though why a special purpose optimal algorithm can be so computationally more effective than a general purpose optimal algorithm, it is because you have to put **intellectual effort** into designing the algorithm!

Special purpose heuristic solution algorithms

With regard to heuristics we have a number of **generic** approaches in the literature, for example:

- greedy
- interchange
- bound based heuristics (e.g. lagrangean heuristics)
- tabu search
- simulated annealing
- population heuristics (e.g. genetic algorithms)

By generic here we mean that there is a general framework/approach from which to build an algorithm. All of these generic approaches however must be tailored for the particular IP we are considering. In addition we can design heuristics purely for the particular problem we are considering (problem-specific heuristics).

Heuristics for IP's are widespread in the literature and applied quite widely in practice. Less has been reported though in terms of heuristics for MIP's.

General IP application areas

There are many areas in which IP has been applied, below we briefly mention some of them, but only in words, no more maths!

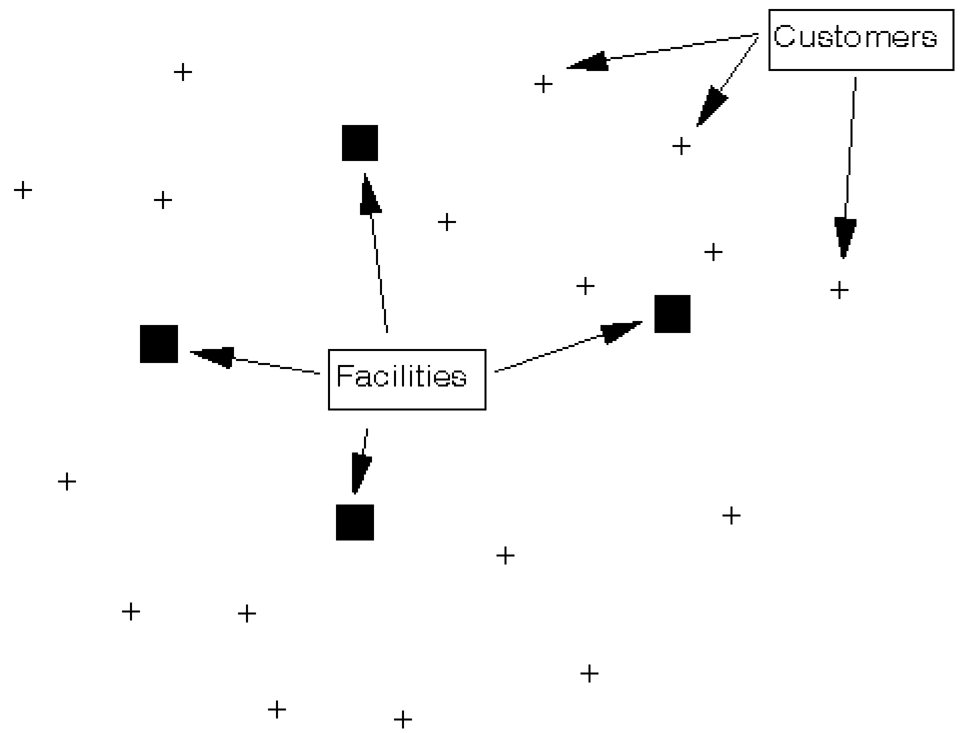
Facility location

Given a set of facility locations and a set of customers who are served from the facilities then:

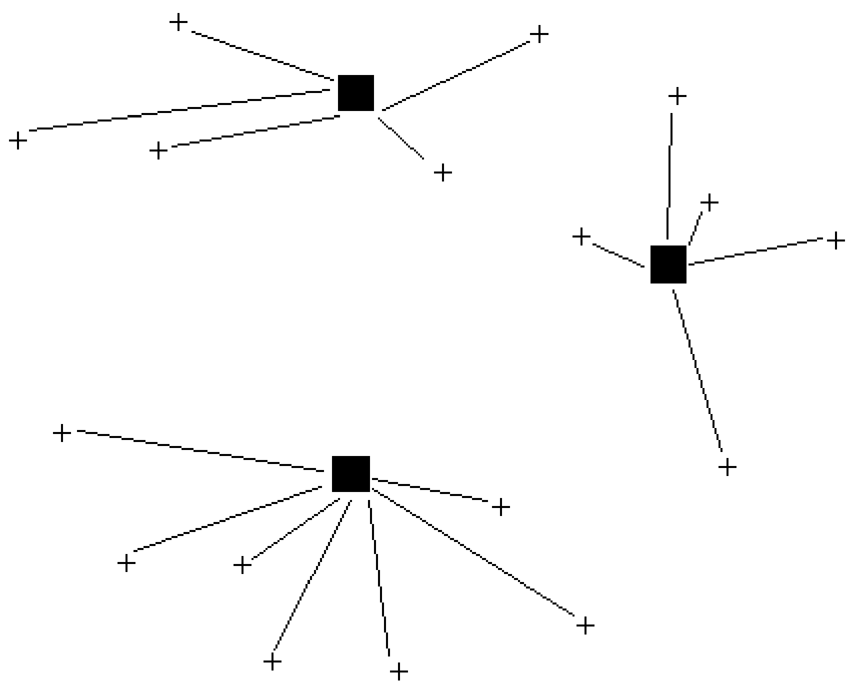
- which facilities should be used
- which customers should be served from which facilities so as to minimise the total cost of serving all the customers.

Typically here facilities are regarded as "open" (used to serve at least one customer) or "closed" and there is a fixed cost which is incurred if a facility is open. Which facilities to have open and which closed is our decision (hence an IP with a zero-one variable representing whether the facility is closed (zero) or open (one)).

Below we show a graphical representation of the problem.



One possible solution is shown below.



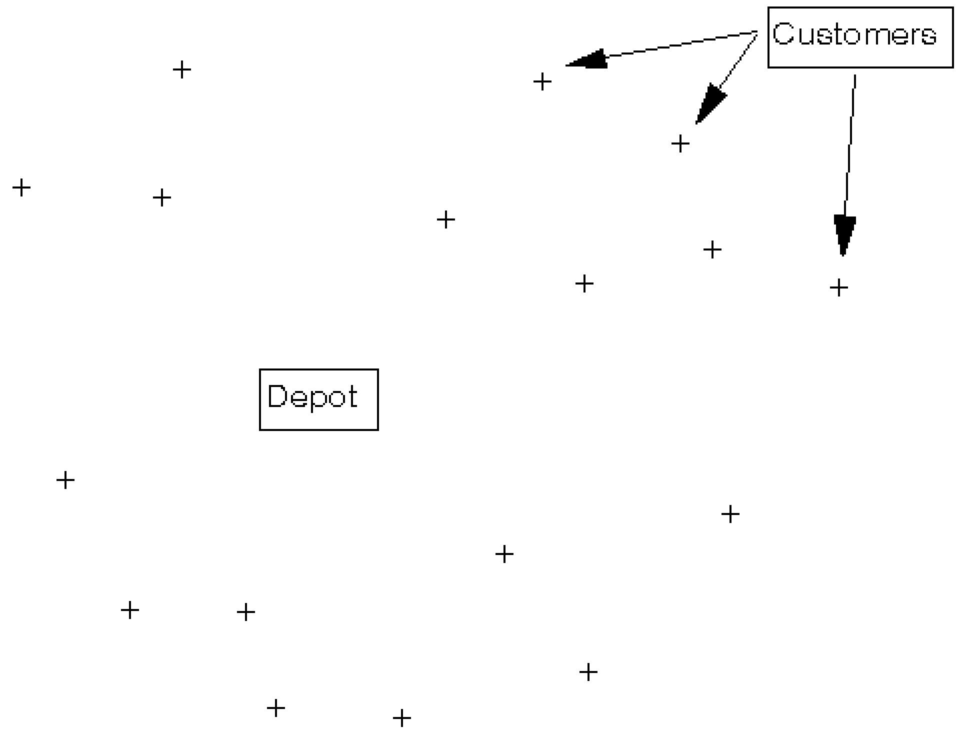
Other factors often encountered here:

- customers have an associated demand with capacities (limits) on the total customer demand that can be served from a facility
- customers being served by more than one facility.

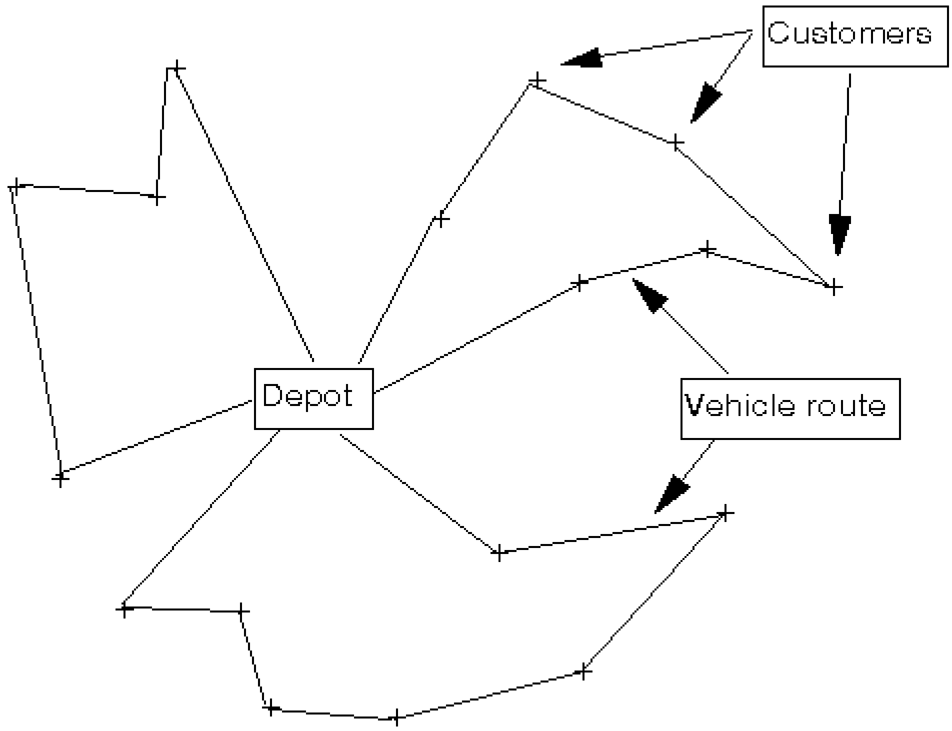
Vehicle routing

Given a set of vehicles based at a central depot and a set of geographically dispersed customers which require visits from the vehicles (e.g. to supply some goods) which vehicles should visit which customers and in what order?

The problem is shown graphically below.



One possible solution is:



Other factors that can occur here are:

- time windows for customer visits
- deliveries and collections
- compartmentalised vehicles (e.g. tankers).

For more about this problem see the vehicle routing notes.

How do I recognise an IP problem?

As a final thought, any decision problem where any of the decisions that have to be made are essentially discrete (such as yes/no, go/no-go), in that one option must be chosen from a finite set of alternatives, can potentially be formulated and solved as an integer programming problem.

How do I choose which IP solution method is appropriate?

Recall the four categories we had:

- general purpose, optimal
- general purpose, heuristic
- special purpose, optimal
- special purpose, heuristic.

One point to note here, although we did not stress it above, is that often an heuristic algorithm can be built for an IP problem without ever having a mathematical formulation of the problem. After all one starts from a verbal description of a problem to construct a mathematical formulation, equally one can construct (i.e. design and code) a heuristic algorithm from a verbal description of the problem (e.g. see here to see an example of this).

Factors which come into play in choosing which IP solution method is appropriate are:

- size of the IP (variables and constraints)
- time available to build the model (formulation plus solution algorithm)
- time available for computer solution once the model has been built
- experience.

Note too that typically IP is used (if applicable) when:

- the problem is a strategic one with large amounts of money involved
- the problem is a tactical one that requires repeated solutions.

BRANCH AND BOUND METHOD.

The branch and bound approach is based on the principle that the total set of feasible solutions can be partitioned into smaller subsets of solutions. These smaller subsets can then be evaluated systematically until the best **solution** is found.

Basic Idea

Branch and Bound algorithm, as a method for global optimization for discrete problems, which are usually NP-hard, searches the complete space of solutions for a given problem for the optimal solution. By solving a relaxed problem of the original one, fractional solutions are recognized and for each discrete variable, B&B will do branching and creating two new nodes, thus dividing the solution space into a set of smaller subsets and obtain the relative upper and lower bound for each node. Since explicit enumeration is normally impossible due to the exponentially increasing number of potential solutions, the use of bounds for the function to be optimized combined with the value of the current best solution found enables this B&B algorithm to search only parts of the solution space implicitly.

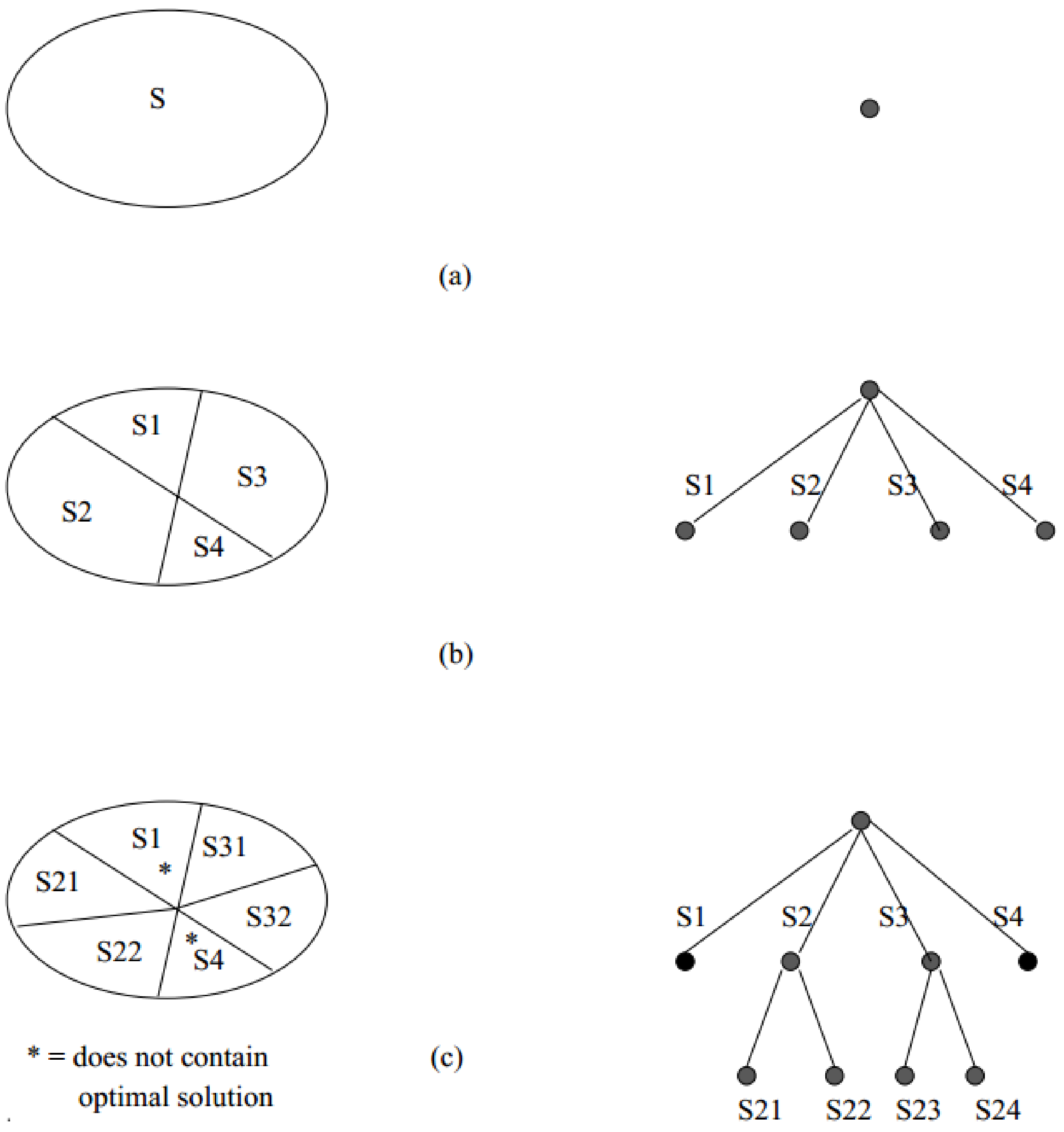


Figure 1: Illustration of the search space of B&B

Branching Strategy

According to the work of Gupta and Ravindran, Generally there are two ways to do branching:

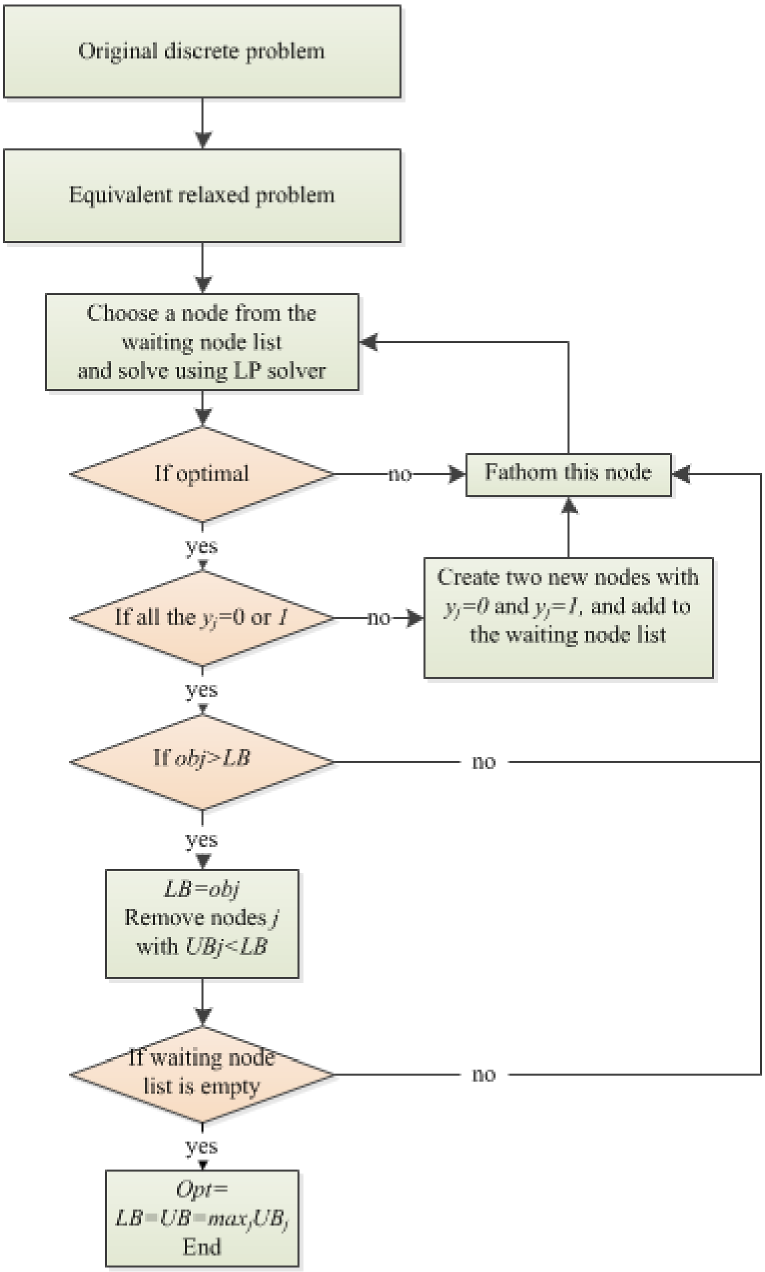
- Branching on the node with the smallest bound

Search all the nodes and find the one with the smallest bound and set it as the next branching node. Advantage: Generally it will inspect less subproblems and thus saves computation time. Disadvantage: Normally it will require more storage.

- Branching on the newly created node with the smallest bound

Search the newly created nodes and find the one with the smallest bound and set it as the next branching node. Advantage: Saves storage space. Disadvantage: Require more branching computation and thus less computational efficiently.

The flow chart for Branch and Bound algorithm is as below:



A Numerical Example

The original mixed integer linear programming problem is as follows:

$$\begin{array}{lccccccccccc} \max & 23x_1 & + & 19x_2 & + & 28x_3 & + & 14x_4 & + & 44x_5 & & \\ & 8x_1 & + & 7x_2 & + & 11x_3 & + & 6x_4 & + & 19x_5 & \leq & 25 \\ & & & & x_1, x_2, x_3, x_4, x_5 = 0 \text{ or } 1 & & & & & & & \end{array}$$

Because this problem is difficult to solve, so we will solve the relaxed problem instead, which is as below:

$$\begin{array}{lccccccccccc} \max & 23x_1 & + & 19x_2 & + & 28x_3 & + & 14x_4 & + & 44x_5 & & \\ & 8x_1 & + & 7x_2 & + & 11x_3 & + & 6x_4 & + & 19x_5 & \leq & 25 \\ & & & & 0 \leq x_1, x_2, x_3, x_4, x_5 \leq 1\,. & & & & & & & \end{array}$$

The set of feasible solution is donated as R_0, which is shown below:

$$\mathcal{R} = \{\mathbf{x} \mid 8x_1 + 7x_2 + 11x_3 + 6x_4 + 19x_5 \leq 25;\ 0 \leq x_1, x_2, x_3, x_4, x_5 \leq 1\}.$$

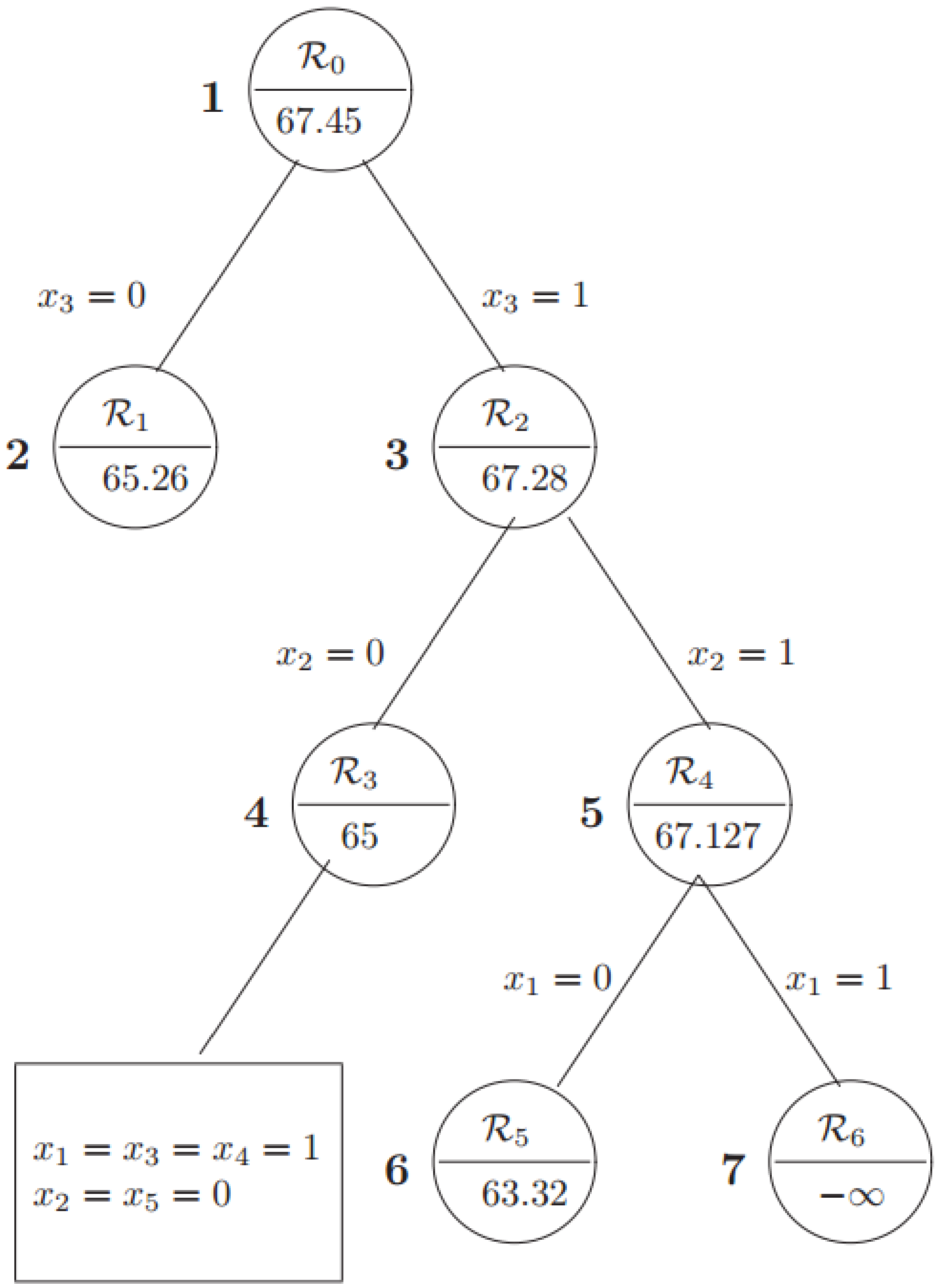
and the solution to the relaxed problem is as follows:

$$x_1^* = x_2^* = 1, x_3^* = \frac{10}{11}, x_4^* = x_5^* = 0\,.$$

Based on this solution, next step we will do branching on x_3, and the resulting new solution subsets is as below:

$$\mathcal{R}_0 = \mathcal{R},\ \mathcal{R}_1 = \mathcal{R}_0 \cap \{\mathbf{x} \mid x_3 = 0\},\ \mathcal{R}_2 = \mathcal{R}_0 \cap \{\mathbf{x} \mid x_3 = 1\}\,.$$

In this way, the branch tree is as follows:



UNIT – IV
GAME THEORY

Learning Objectives

Game Theory: Introduction – Two Person Zero-Sum Games, Pure Strategies, Games with Saddle Point, Mixed strategies, Rules of Dominance, Solution Methods of Games without Saddle point – Algebraic, matrix and arithmetic methods. Simulation – Simulation Inventory and Waiting Lines.

GAME THEORY - INTRODUCTION

Game theory is the process of modelling the strategic interaction between two or more players in a situation containing set rules and outcomes. While used in several disciplines, **game theory** is most notably used as a tool within the study of economics

Game theory is the process of modeling the strategic interaction between two or more players in a situation containing set rules and outcomes. While used in a number of disciplines, game theory is most notably used as a tool within the study of economics. The economic application of game theory can be a valuable tool to aide in the fundamental analysis of industries, sectors and any strategic interaction between two or more firms.

Here, we'll take an introductory look at game theory and the terms involved, and introduce you to a simple method of solving games, called backwards induction.

Game Theory Definitions

Any time we have a situation with two or more players that involves known payouts or quantifiable consequences, we can use game theory to help determine the most likely outcomes.

Let's start out by defining a few terms commonly used in the study of game theory:

- **Game**: Any set of circumstances that has a result dependent on the actions of two of more decision-makers (players).
- **Players**: A strategic decision-maker within the context of the game.
- **Strategy**: A complete plan of action a player will take given the set of circumstances that might arise within the game.
- **Payoff**: The payout a player receives from arriving at a particular outcome. The payout can be in any quantifiable form, from dollars to utility.
- **Information set**: The information available at a given point in the game. The term information set is most usually applied when the game has a sequential component.
- **Equilibrium**: The point in a game where both players have made their decisions and an outcome is reached.

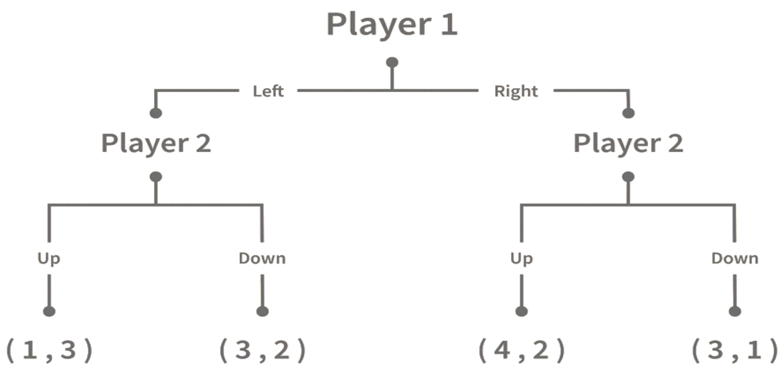
Assumptions in Game Theory

As with any concept in economics, there is the assumption of rationality. There is also an assumption of maximization. It is assumed that players within the game are rational and will strive to maximize their payoffs in the game.

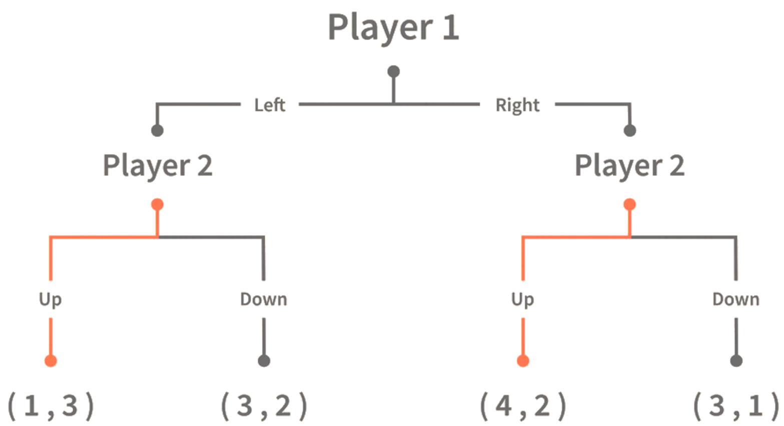
When examining games that are already set up, it is assumed on your behalf that the payouts listed include the sum of all payoffs associated with that outcome. This will exclude any "what if" questions that may arise.

The number of players in a game can theoretically be infinite, but most games will be put into the context of two players. One of the simplest games is a sequential game involving two players.

Solving Sequential Games Using Backwards Induction

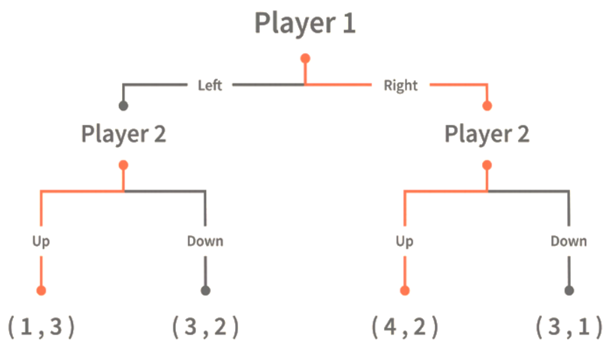


Below is a simple sequential game between two players. The labels with Player 1 and Player 2 within them are the information sets for players one or two, respectively. The numbers in the parentheses at the bottom of the tree are the payoffs at each respective point. The game is also sequential, so Player 1 makes the first decision (left or right) and Player 2 makes its decision after Player 1 (up or down).



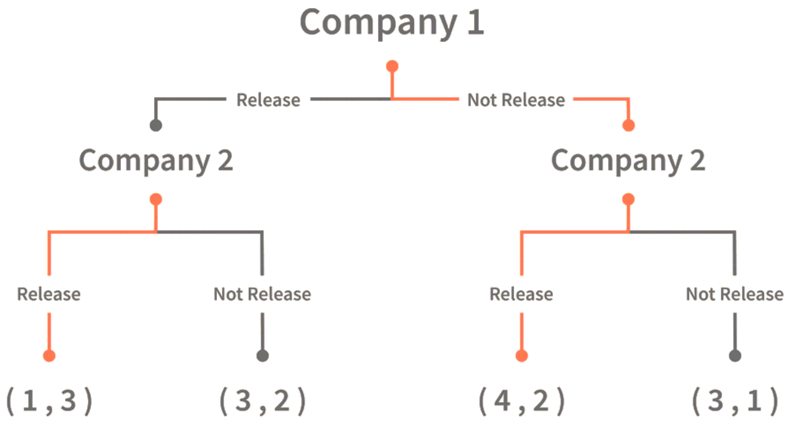
Backward induction, like all game theory, uses the assumptions of rationality and maximization, meaning that Player 2 will maximize his payoff in any given situation. At either information set, we have two choices, four in all. By eliminating the choices that Player 2 will not choose, we can narrow down our tree. In this way, we will bold the lines that maximize the player's payoff at the given information set.

After this reduction, Player 1 can maximize its payoffs now that Player 2's choices are made known. The result is an equilibrium found by backward induction of Player 1 choosing "right" and Player 2 choosing "up." Below is the solution to the game with the equilibrium path in bold.



For example, one could easily set up a game similar to the one above using companies as the players. This game could include product release scenarios. If Company 1 wanted to release a product, what might Company 2 do in response? Will Company 2 release a similar competing product?

By forecasting sales of this new product in different scenarios, we can set up a game to predict how events might unfold. Below is an example of how one might model such a game.



The Bottom Line

By using simple methods of game theory, we can solve for what would be a confusing array of outcomes in a real-world situation. Using game theory as a tool for financial analysis can be very helpful in sorting out potentially messy real-world situations, from mergers to product releases.

TWO PERSON ZERO-SUM GAMES

A two player game is called a *zero-sum* game if the sum of the payoffs to each player is constant for all possible outcomes of the game. More specifically, the terms (or coordinates) in each payoff vector must add up to the same value for each payoff vector. Such games are sometimes called *constant-sum* games instead.

The simplest type of competitive situations are **two-person**, **zero-sum games**. These **games** involve only **two** players; they are called **zero-sum games** because one **player** wins whatever the other **player** loses.

Two-Person Zero-Sum Games: Basic Concepts

Game theory provides a mathematical framework for analyzing the decision-making processes and strategies of adversaries (or *players*) in different types of competitive situations. The simplest type of competitive situations are **two-person, zero-sum games**. These games involve only two players; they are called *zero-sum* games because one player wins whatever the other player loses.

Payoff Matrix

Strategy		Player 2: 1	Player 2: 2
Player 1	1	2	-2
	2	-2	2

Example: Odds and Evens

Consider the simple game called **odds and evens**. Suppose that player 1 takes evens and player 2 takes odds. Then, each player simultaneously shows either one finger or two fingers. If the number of fingers matches, then the result is *even*, and player 1 wins the bet ($2). If the number of fingers does not match, then the result is *odd*, and player 2 wins the bet ($2). Each player has two possible strategies: show one finger or show two fingers. The *payoff matrix* shown below represents the payoff to player 1.

Basic Concepts of Two-Person Zero-Sum Games

This game of odds and evens illustrates important concepts of simple games.

- A two-person game is characterized by the strategies of each player and the payoff matrix.
- The payoff matrix shows the gain (positive or negative) for player 1 that would result from each combination of strategies for the two players. *Note that the matrix for player 2 is the negative of the matrix for player 1 in a zero-sum game.*
- The entries in the payoff matrix can be in any units if they represent the *utility (or value)* to the player.
- There are two key assumptions about the behaviour of the players. The first is that both players are *rational*. The second is that both players are *greedy* meaning that they choose their strategies in their own interest (to promote their own wealth).

Definition of two-person zero sum game

A game with only two players, say player A and player B, is called a two-person zero sum game if the gain of the player A is equal to the loss of the player B, so that the total sum is zero.

Payoff matrix

When players select their strategies, the payoffs (gains or losses) can be represented in the form of a payoff matrix.

Since the game is zero sum, the gain of one player is equal to the loss of other and vice-versa. Suppose A has m strategies and B has n strategies. Consider the following payoff matrix.

Player B's strategies

$$\text{Player A's strategies}\quad \begin{array}{c} \\ A_1 \\ A_1 \\ \vdots \\ Am \end{array} \begin{array}{c} \begin{array}{cccc} B_1 & B_2 & \cdots & B_n \end{array} \\ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \end{array}$$

Player A wishes to gain as large a payoff a_{ij} as possible while player B will do his best to reach as small a value a_{ij} as possible where the gain to player B and loss to player A be ($- a_{ij}$).

PURE STRATEGIES

A **pure strategy** provides a complete definition of how a player will play a **game**. ... In particular, it determines the move a player will make for any situation they could face. A player's **strategy** set is the set of **pure strategies** available to that player.

Pure vs. Mixed Strategies

The stadium lights are blinding, and the murmuring of the crowd in the stands is amplified into a deafening roar. Yet, your senses have never been more acute. The date is January 29, 2017, and are playing to win your fifth Australian Open championship title in tennis. Millions of people are watching your every movement from across the world. You are Roger Federer. Where do you place your serves across the net?

We are often unaware of the different dimensions that everyday games are comprised of. Something seemingly as simple as a serve in tennis can be dissected into many parts, both physical and mental. In this article, we are going to explore pure and mixed strategies in game theory, using tennis as an example.

<u>What is a pure strategy?</u>

A pure strategy is an unconditional, defined choice that a person makes in a situation or game. For example, in the game of Rock-Paper-Scissors,if a player would choose to only play scissors for each and every independent trial, regardless of the other player's strategy, choosing scissors would be the player's pure strategy. The probability for choosing scissors equal to 1 and all other options (paper and rock) is chosen with the probability of 0. The set of all options (i.e. rock, paper, and scissors) available in this game is known as the strategy set.

<u>What is a mixed strategy?</u>

A mixed strategy is an assignment of probability to all choices in the strategy set. Using the example of Rock-Paper-Scissors, if a person's probability of employing each pure strategy is equal, then the probability distribution of the strategy set would be 1/3 for each option, or approximately 33%. In other words, a person using a mixed strategy incorporates more than one pure strategy into a game.

The definition of a mixed strategy does not rule out the possibility for an option(s)to never be chosen (eg. $p_{scissors}= 0.5$, $p_{rock} = 0.5$, $p_{paper} = 0$). This means that in a way, a pure strategy can also be considered a mixed strategy at its extreme, with a binary probability assignment (setting one option to 1 and all others equal to 0). For this article, we shall say that pure strategies are not mixed strategies.

In the game of tennis, each point is a zero-sum game with two players (one being the server *S*, and the other being the returner *R*). In this scenario, assume each player has two strategies (forehand *F*, and backhand *B*). Observe the following hypothetical in the payoff matrix:

		Returner	
		F	B
Server	F	40, 60	70, 30
	B	90, 10	50, 50

The strategies F_S or B_S are observed for the server when the ball is served to the side of the service box closest to the returner's forehand or backhand, respectively. For the returner, the strategies F_R and B_R are observed when the returner moves to the forehand or backhand side to return the serve, respectively. This gives us the payoffs when the returner receives the serve correctly (F_S,F_R or B_S,B_R), or incorrectly (F_S,B_R or B_S,F_R). The payoffs to each player for every action are given in pure strategy payoffs, as each player is only guaranteed their payoff given the opponent's strategy is employed 100% of the time. Given these pure strategy payoffs, we can calculate the mixed strategy payoffs by figuring out the probability each strategy is chosen by each player.

So you are Roger. It is apparent to you that a pure strategy would be exploitable. If you serve to the backhand 100% of the time, it would be easy for the opponent to catch on and return from the backhand side more often than the forehand, maximizing his expected payoff. Same goes for the serve to the forehand. But how often should you mix your strategy and serve to each side to minimize your opponent's chances of winning? Calculating these probabilities would give us our mixed strategy Nash equilibria, or the probabilities that each strategy is used which would minimize the opponent's expected payoff. In the following article, we will look at how to find mixed strategy Nash equilibria, and how to interpret them.

GAMES WITH SADDLE POINT

In a zero-sum matrix **game**, an outcome is a **saddle point** if the outcome is a minimum in its row and maximum in its column. The argument that players will prefer not to diverge from the **saddle point** leads us to offer the following principle of **game theory**: ... If a matrix **game** has a **saddle point**, both players should play it.

SADDLE POINT STEPS (RULE)	
Step-1:	1. Select the minimum element from each row and write them in Row Minimum column. 2. Select the maximum element from Row Minimum column and enclose it in [ ]. It is called Row MaxiMin.
Step-2:	1. Select the maximum element from each column and write them in Column Maximum row. 2. Select the minimum element from Column Maximum row and enclose it in ( ). It is called Column MiniMax.
Step-3:	1. Find out the elements that is same in rectangle [ ] and circle ( ). 2. If Column MiniMax = Row MaxiMin then the game has saddle point and it is the value of the game.

Example-1
Find Solution of game theory problem using saddle point

Player A\Player B	B1	B2	B3	B4
A1	20	15	12	35
A2	25	14	8	10
A3	40	2	10	5
A4	-5	4	11	0

Solution:
1. Saddle point testing
Players

			Player B			
			B1	B2	B3	B4
Player A	A1		20	15	12	35
	A2		25	14	8	10
	A3		40	2	10	5
	A4		-5	4	11	0

We apply the maximin (minimax) principle to analyze the game.

			Player B					
			B1	B2	B3	B4		Row Minimum
Player A	A1		20	15	[(12)]	35		[12]
	A2		25	14	8	10		8

	A3		40	2	10	5			2
	A4		-5	4	11	0			-5
	Column Maximum		40	15	(12)	35			

Select minimum from the maximum of columns
Column MiniMax = (12)

Select maximum from the minimum of rows
Row MaxiMin = [12]

Here, Column MiniMax = Row MaxiMin = 12
∴ This game has a saddle point and value of the game is 12

The optimal strategies for both players are
The player A will always adopt strategy 1
The player B will always adopt strategy 3

Example-2

Find Solution of game theory problem using saddle point

Player A\Player B	B1	B2	B3
A1	-2	14	-2
A2	-5	-6	-4
A3	-6	20	-8

Solution:
1. Saddle point testing
Players

			Player B				
			B1	B2	B3		
Player A	A1		-2	14	-2		
	A2		-5	-6	-4		
	A3		-6	20	-8		

We apply the maximin (minimax) principle to analyze the game.

		Player B			
		B1	B2	B3	Row Minimum
Player A	A1	[(-2)]	14	-2	[-2]
	A2	-5	-6	-4	-6
	A3	-6	20	-8	-8
	Column Maximum	(-2)	20	-2	

Select minimum from the maximum of columns
Column MiniMax = (-2)
Select maximum from the minimum of rows
Row MaxiMin = [-2]
Here, Column MiniMax = Row MaxiMin = -2
∴ This game has a saddle point and value of the game is -2

The optimal strategies for both players are
The player A will always adopt strategy 1
The player B will always adopt strategy 1

MIXED STRATEGIES

A **mixed strategy** exists in a strategic **game**, when the player does not choose one definite action, but rather, chooses according to a probability distribution over a his actions. ... Note: In pure **strategies**, the player assigns 100% probability to one plan of action.

Suppose that Row believes Column plays Heads with probability *p*. Then if Row plays Heads, Row gets 1 with probability *p* and –1 with probability (1 – *p*), for an expected value of 2*p* – 1. Similarly, if Row plays Tails, Row gets –1 with probability *p* (when Column plays Heads), and 1 with probability (1 – *p*), for an expected value of 1 – 2*p*.

		Column	
		Heads	Tails
Row	Heads	(1, –1)	(–1, 1)
	Tails	(–1, 1)	(1, –1)

If 2*p* – 1 > 1 – 2*p*, then Row is better off, on average, playing Heads than Tails. Similarly, if 2*p* – 1 < 1 – 2*p*, then Row is better off playing Tails than Heads. If, on the other hand, 2*p* – 1 = 1 – 2*p*, then Row gets the same payoff no matter what Row does. In this case, Row could play Heads, could play Tails, or could flip a coin and randomize Row's play.

Column			
Row	Heads	Tails	
Heads	(1, –1)	(–1, 1)	$1p + -1(1 - p) = 2p - 1$
Tails	(–1, 1)	(1, –1)	$-1p + 1(1 - p) = 1 - 2p$

A mixed strategy Nash equilibrium involves at least one player playing a randomized strategy and no player being able to increase his or her expected payoff by playing an alternate strategy. A Nash equilibrium in which no player randomizes is called a pure strategy Nash equilibrium.

Note that randomization requires equality of expected payoffs. If a player is supposed to randomize over strategy *A* or strategy *B*, then both of these strategies must produce the same expected payoff. Otherwise, the player would prefer one of them and wouldn't play the other.

Computing a mixed strategy has one element that often appears confusing. Suppose that Row is going to randomize. Then Row's payoffs must be equal for all strategies that Row plays with positive probability. But that equality in Row's payoffs doesn't determine the probabilities with which Row plays the various rows. Instead, that equality in Row's payoffs will determine the probabilities with which Column plays the various columns. The reason is that it is Column's probabilities that determine the expected payoffs for Row; if Row is going to randomize, then Column's probabilities must be such that Row is willing to randomize.

Thus, for example, we computed the payoff to Row of playing Heads, which was $2p - 1$, where p was the probability that Column played Heads. Similarly, the payoff to Row of playing Tails was $1 - 2p$. Row is willing to randomize if these are equal, which solves for $p = ½$.

This game has two pure strategy Nash equilibria: (Baseball, Baseball) and (Ballet, Ballet). Is there a mixed strategy? To compute a mixed strategy, let the Woman go to the Baseball game with probability p, and the Man go to the Baseball game with probability q.

Woman		
Man	Baseball	Ballet
Baseball	(3, 2)	(1, 1)
Ballet	(0, 0)	(2, 3)

		Woman		
		Baseball (p)	Ballet ($1-p$)	Man's E Payoff
Man	Baseball (q)	(3, 2)	(1, 1)	$3p + 1(1-p) = 1 + 2p$
	Ballet ($1-q$)	(0, 0)	(2, 3)	$0p + 2(1-p) = 2 - 2p$
	Woman's E Payoff	$2q + 0(1-q) = 2q$	$1q + 3(1-q) = 3 - 2q$	

For example, if the Man (row player) goes to the Baseball game, he gets 3 when the Woman goes to the Baseball game (probability p), and otherwise gets 1, for an expected payoff of $3p + 1(1-p) = 1 + 2p$. The other calculations are similar, but you should definitely run through the logic and verify each calculation.

A mixed strategy in the battle of the sexes game requires both parties to randomize (since a pure strategy by either party prevents randomization by the other). The Man's indifference between going to the Baseball game and to the Ballet requires $1 + 2p = 2 - 2p$, which yields $p = ¼$. That is, the Man will be willing to randomize which event he attends if the Woman is going to the Ballet ¾ of the time, and otherwise to the Baseball game. This makes the Man indifferent between the two events because he prefers to be with the Woman, but he also likes to be at the Baseball game. To make up for the advantage that the game holds for him, the Woman has to be at the Ballet more often.

Similarly, in order for the Woman to randomize, the Woman must get equal payoffs from going to the Baseball game and going to the Ballet, which requires $2q = 3 - 2q$, or $q = ¾$. Thus, the probability that the Man goes to the Baseball game is ¾, and he goes to the Ballet ¼ of the time. These are independent probabilities, so to get the probability that both go to the Baseball game, we multiply the probabilities, which yields 3/16. Figure 16.17 "Mixed strategy probabilities" fills in the probabilities for all four possible outcomes.

Note that more than half of the time (Baseball, Ballet) is the outcome of the mixed strategy and the two people are not together. This lack of coordination is generally a feature of mixed strategy equilibria. The expected payoffs for both players are readily computed as well. The Man's payoff is $1 + 2p = 2 - 2p$, and since $p = ¼$, the Man obtains 1½. A similar calculation shows that the Woman's payoff is the same. Thus, both do worse than coordinating on their less preferred outcome. But this mixed strategy Nash equilibrium, undesirable as it may seem, is a Nash equilibrium in the sense that neither party can improve his or her own payoff, given the behavior of the other party.

		Woman	
		Baseball	Ballet
Man	Baseball	3/16	9/16
	Ballet	1/16	3/16

In the battle of the sexes, the mixed strategy Nash equilibrium may seem unlikely; and we might expect the couple to coordinate more effectively. Indeed, a simple call on the telephone should rule out the mixed strategy. So let's consider another game related to the battle of the sexes, where a failure of coordination makes more sense. This is the game of "chicken." In this game, two players drive toward one another, trying to convince the other to yield and ultimately swerve into a ditch. If both swerve into the ditch, we'll call the outcome a draw and both get zero. If one swerves and the other doesn't, the driver who swerves loses and the other driver wins, and we'll give the winner one point.Note that adding a constant to a player's payoffs, or multiplying that player's payoffs by a positive constant, doesn't affect the Nash equilibria—pure or mixed. Therefore, we can always let one outcome for each player be zero, and another outcome be one. The only remaining question is what happens when neither yield, in which case a crash results. In this version, the payoff has been set at four times the loss of swerving, as shown in Figure 16.18 "Chicken", but you can change the game and see what happens.

Column

Row	Swerve	Don't
Swerve	(0, 0)	(–1, 1)
Don't	(1, –1)	(–4, –4)

This game has two pure strategy equilibria: (Swerve, Don't) and (Don't, Swerve). In addition, it has a mixed strategy. Suppose that Column swerves with probability p. Then Row gets $0p + -1(1 - p)$ from swerving, $1p + (-4)(1 - p)$ from not swerving, and Row will randomize if these are equal, which requires $p = ¾$. That is, the probability that Column swerves in a mixed strategy equilibrium is ¾. You can verify that the row player has the same probability by setting the probability that Row swerves equal to q and computing Column's expected payoffs. Thus, the probability of a collision is 1/16 in the mixed strategy equilibrium.

The mixed strategy equilibrium is more likely, in some sense, in this game: If the players already knew who was going to yield, they wouldn't actually need to play the game. The whole point of the game is to find out who will yield, which means that it isn't known in advance. This means that the mixed strategy equilibrium is, in some sense, the more reasonable equilibrium.

"Rock, paper, scissors" is a child's game in which two children use their hands to simultaneously choose paper (hand held flat), scissors (hand with two fingers protruding to look like scissors), or rock (hand in a fist). The nature of the payoffs is that paper beats rock, rock beats scissors, and scissors beats paper.

Column

Row	Paper	Scissors	Rock
Paper	(0, 0)	(–1, 1)	(1, –1)
Scissors	(1, –1)	(0, 0)	(–1, 1)
Rock	(–1, 1)	(1, –1)	(0, 0)

Key Takeaways

- A mixed strategy Nash equilibrium involves at least one player playing a randomized strategy and no player being able to increase his or her expected payoff by playing an alternate strategy.
- A Nash equilibrium without randomization is called a pure strategy Nash equilibrium.
- If a player is supposed to randomize over two strategies, then both must produce the same expected payoff.
- The matching pennies game has a mixed strategy and no pure strategy.
- The battle of the sexes game has a mixed strategy and two pure strategies.
- The game of chicken is like the battle of the sexes and, like it, has two pure strategies and one mixed strategy.

RULES OF DOMINANCE

The principle of dominance states that if one strategy of a player dominates over the other strategy in all conditions then the later strategy can be ignored. A strategy dominates over the other only if it is preferable over other in all conditions.

A strategy dominates over the other only if it is preferable over other in all conditions. The concept of dominance is especially useful for the evaluation of **two-person zero-sum games** where a saddle point does not exist.

Dominant Strategy Rules (Dominance Principle)

- If all the elements of a column (say i_{th} column) are greater than or equal to the corresponding elements of any other column (say j_{th} column), then the i_{th} column is dominated by the j_{th} column and can be deleted from the matrix.
- If all the elements of a row (say i_{th} row) are less than or equal to the corresponding elements of any other row (say j_{th} row), then the i_{th} row is dominated by the j_{th} row and can be deleted from the matrix.

Dominance Example: Game Theory

		Player B			
		I	II	III	IV
Player A	I	3	5	4	2
	II	5	6	2	4
	III	2	1	4	0
	IV	3	3	5	2

Use the principle of dominance to solve this problem.

Solution.

	Player B					
		I	II	III	IV	Minimum
Player A	I	3	5	4	2	2
	II	5	6	2	4	2
	III	2	1	4	0	0
	IV	3	3	5	2	2
Maximum		5	6	5	4	

There is no **saddle point** in this game.

Using Dominance Property In Game Theory

	Player B	
		III	IV
Player A	I	4	2
	II	2	4
	III	4	0
	IV	5	2

If a column is greater than another column (compare corresponding elements), then delete that column. Here, I and II column are greater than the IV column. So, player B has no incentive in using his I and II course of action.

	Player B		
		III	IV
Player A	II	2	4
	IV	5	2

If a row is smaller than another row (compare corresponding elements), then delete that row. Here, I and III row are smaller than IV row. So, player A has no incentive in using his I and III course of action.

Dominance property in Operation Research

The principle of dominance states that if one strategy of a player dominates over the other strategy in all conditions then the later strategy can be ignored. A strategy dominates over the other only if it is preferable over other in all conditions. The concept of dominance is especially useful for the evaluation of two-person zero-sum games where a saddle point does not exist. In case of pay-off matrices larger than 2×2 size, the dominance property can be used to reduce the size of the pay-off matrix by eliminating the strategies that would never be selected.

Dominance Principle in Game Theory problems

The dominance principle in game theory problems are explained below

Example : Solve the game given below in Table after reducing it to 2×2 game:

Game Problem

$$\begin{array}{ccc} & & \textbf{Player B} \\ & & \begin{array}{ccc} \mathbf{1} & \mathbf{2} & \mathbf{3} \end{array} \\ \textbf{Player A} & \begin{array}{c} \mathbf{1} \\ \mathbf{2} \\ \mathbf{3} \end{array} & \begin{pmatrix} 1 & 7 & 2 \\ 6 & 2 & 7 \\ 5 & 1 & 6 \end{pmatrix} \end{array}$$

Solution: Reduce the matrix by using the dominance property. In the given matrix for player A, all the elements in Row 3 are less than the adjacent elements of Row 2. Strategy 3 will not be selected by player A, because it gives less profit for player A. Row 3 is dominated by Row 2. Hence delete Row 3, as shown in table.

Reduced the Matrix by Using Dominance Property

$$\begin{array}{ccc} & & \textbf{Player B} \\ & & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \textbf{Player A} & \begin{array}{c} \mathbf{1} \\ \mathbf{2} \end{array} & \begin{pmatrix} 1 & 7 & 2 \\ 6 & 2 & 7 \end{pmatrix} \end{array}$$

For Player B, Column 3 is dominated by column 1 (Here the dominance is opposite because Player B selects the minimum loss). Hence delete Column 3. We get the reduced 2 × 2 matrix as shown below in table.

Reduced 2 × 2 Matrix

$$\begin{array}{ccc} & & \textbf{Player B} \\ & & \begin{array}{cc} 1 & 2 \end{array} \\ \text{Player A} & \begin{array}{c} \mathbf{1} \\ \mathbf{2} \end{array} & \begin{pmatrix} 1 & 7 \\ 6 & 2 \end{pmatrix} \end{array}$$

Now, solve the 2 × 2 matrix, using the maximin criteria as shown below in table.

Maximin Procedure

$$
\begin{array}{lccc}
 & & \textbf{Player B} & \\
 & \textbf{1} & \textbf{2} & \textbf{Row Min} \\
\text{Player A} \begin{array}{c} \textbf{1} \\ \textbf{2} \end{array} & \multicolumn{2}{c}{\begin{pmatrix} 1 & 7 \\ 6 & 2 \end{pmatrix}} & \begin{array}{c} 1 \\ 2 \end{array} \\
\textbf{Column Max} & 6 & 7 &
\end{array}
$$

$$\text{Max Min} \neq \text{Min Max}$$

$$2 \neq 6$$

The optimum strategies are shown in table

Optimum Strategies

(a) $S_A = \begin{pmatrix} A_1 & A_2 & A_3 \\ {}^2/_5 & {}^3/_5 & 0 \end{pmatrix}$ and (b) $S_B = \begin{pmatrix} B_1 & B_2 & B_3 \\ ½ & ½ & 0 \end{pmatrix}$

Value of the game, $v = 4$

Example : Is the following two-person zero-sum game stable? Solve the game given below in table.

Two-person Zero-sum Game Problem

$$
\begin{array}{ll}
 & \textbf{Player B} \\
 & \begin{array}{cccc} \quad\textbf{1} & \quad\textbf{2} & \quad\textbf{3} & \quad\textbf{4} \end{array} \\
\textbf{Player A} \begin{array}{c} \textbf{1} \\ \textbf{2} \\ \textbf{3} \\ \textbf{4} \end{array} & \begin{pmatrix} 5 & -10 & 9 & 0 \\ 6 & 7 & 8 & 1 \\ 8 & 7 & 15 & 1 \\ 3 & 4 & -1 & 4 \end{pmatrix}
\end{array}
$$

Solution: Solve the given matrix using the maximin criteria as shown in table.

Table 14.25: Maximin Procedure

		Player B				
		1	2	3	4	Row Min
	1	5	–10	9	0	–10
Player A	2	6	7	8	1	1
	3	8	7	15	1	1
	4	3	4	–1	4	–1
Column Max		8	7	15	4	

Therefore, there is no saddle point and hence it has a mixed strategy.
The pay-off matrix is reduced to 2×2 size using dominance property. Compare the rows to find the row which dominates other row. Here for Player A, Row 1 is dominated by Row 3 (or row 1 gives the minimum profit for Player A), hence delete Row 1. The matrix is reduced as shown in table.

Use Dominance Property to Reduce Matrix (Deleted Row 1)

$$\text{Player B}$$

$$\text{Player A}\quad \begin{array}{c} \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 6 & 7 & 8 & 1 \\ 8 & 7 & 15 & 1 \\ 3 & 4 & -1 & 4 \end{pmatrix} \end{array}$$

When comparing column wise, column 2 is dominated by column 4. For Player B, the minimum profit column is column 2, hence delete column 2. The matrix is further reduced as shown in table.

Matrix Further Reduced to 3×3 (2 Deleted Column)

$$\text{Player B}$$

$$\text{Player A}\quad \begin{array}{c} \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 3 & 4 \end{array} \\ \begin{pmatrix} 6 & 8 & 1 \\ 8 & 15 & 1 \\ 3 & -1 & 4 \end{pmatrix} \end{array}$$

Now, Row 2 is dominated by Row 3, hence delete Row 2, as shown in table.

Reduced Matrix (Row 2 Deleted)

$$\begin{array}{ccc} & & \textbf{Player B} \\ & & \begin{array}{ccc} \mathbf{1} & \mathbf{3} & \mathbf{4} \end{array} \\ \textbf{Player A} & \begin{array}{c} \mathbf{3} \\ \mathbf{4} \end{array} & \begin{pmatrix} 8 & 15 & 1 \\ 3 & -1 & 4 \end{pmatrix} \end{array}$$

Now, as when comparing rows and columns, no column or row dominates the other. Since there is a tie while comparing the rows or columns, take the average of any two rows and compare. We have the following three combinations of matrices as shown in table.

Matrix Combinations

(a) B	(b) B	(c) B
$\begin{array}{cc} & \frac{R_1+R_3}{2} \quad R_3 \\ A & \begin{pmatrix} 11.5 & 1 \\ 1 & 4 \end{pmatrix} \end{array}$	$\begin{array}{cc} & R_1 \quad \frac{R_3+R_4}{2} \\ A & \begin{pmatrix} 8 & 8 \\ 3 & 1.5 \end{pmatrix} \end{array}$	$\begin{array}{cc} & R_2 \quad \frac{R_1+R_4}{2} \\ A & \begin{pmatrix} 15 & 4.5 \\ -1 & 3.5 \end{pmatrix} \end{array}$
×	✓	×

When comparing column 1 and the average of column 3 and column 4, column 1 is dominated by the average of column 3 and 4. Hence delete column 1. Finally, we get the 2 × 2 matrix as shown in table.

2×2 Matrix After Deleting Column 1

$$\begin{array}{ccc} & & \textbf{Player B} \\ & & \begin{array}{cc} \mathbf{3} & \mathbf{4} \end{array} \\ \textbf{Player A} & \begin{array}{c} \mathbf{3} \\ \mathbf{4} \end{array} & \begin{pmatrix} 15 & 1 \\ -1 & 4 \end{pmatrix} \end{array}$$

Optimum Mixed Strategies

$$\text{(a)} \quad S_A = \begin{bmatrix} A_1 & A_2 & A_3 & A_4 \\ 0 & 0 & {}^{5}/_{19} & {}^{14}/_{19} \end{bmatrix} \text{ and (b)} \quad S_B = \begin{bmatrix} B_1 & B_2 & B_3 & B_4 \\ 0 & 0 & {}^{3}/_{19} & {}^{16}/_{19} \end{bmatrix}$$

SOLUTION METHODS OF GAMES WITHOUT SADDLE POINT

Method of solution of a 2x2 zero-sum game without saddle point

Suppose that a 2x2 game has no saddle point. Suppose the game has the following pay-off matrix.

$$\begin{array}{cc} & \text{Player B} \\ \text{Player A} & \begin{bmatrix} a & b \\ c & d \end{bmatrix} \end{array}$$

Since this game has no saddle point, the following condition shall hold:

Max {*Min* {*a* , *b*}, *Min* {*c* , *d* }} ≠ *Min* {*Max* {*a* , *c*}, *Max* {*b*, *d*}}

In this case, the game is called a mixed game. ***No strategy of Player A can be called the best strategy for him. Therefore A has to use both of his strategies. Similarly no strategy of Player B can be called the best strategy for him and he has to use both of his strategies.***

Let p be the probability that Player A will use his first strategy. Then the probability that Player A will use his second strategy is 1-p.

<u>If Player B follows his first strategy</u>

$$\begin{aligned} &\text{Expected value of the pay-off to Player A} \\ &= \left\{ \begin{matrix} \text{Expected value of the pay-off to Player A} \\ \text{arising from his first strategy} \end{matrix} \right\} + \left\{ \begin{matrix} \text{Expected value of the pay-off to Player A} \\ \text{arising from his second strategy} \end{matrix} \right\} \\ &= ap + c(1-p) \qquad (1) \end{aligned}$$

In the above equation, note that the expected value is got as the product of the corresponding values of the pay-off and the probability.

<u>If Player B follows his second strategy</u>

$$\left.\begin{matrix} \text{Expected value of the} \\ \text{pay-off to Player A} \end{matrix}\right\} = bp + d(1-p) \qquad (2)$$

If the expected values in equations (1) and (2) are different, Player B will prefer the minimum of the two expected values that he has to give to player A. Thus B will have a pure strategy. This contradicts our assumption that the game is a mixed one. Therefore the expected values of the pay-offs to Player A in equations (1) and (2) should be equal. Thus we have the condition

$$
\begin{aligned}
ap + c(1-p) &= bp + d(1-p) \\
ap - bp &= (1-p)[d-c] \\
p(a-b) &= (d-c) - p(d-c) \\
p(a-b) + p(d-c) &= d-c \\
p(a-b+d-c) &= d-c \\
p &= \frac{d-c}{(a+d)-(b+c)} \\
1-p &= \frac{a+d-b-c-d+c}{(a+d)-(b+c)} \\
&= \frac{a-b}{(a+d)-(b+c)}
\end{aligned}
$$

$$
\left\{\begin{matrix}\text{The number of times A} \\ \text{will use first strategy}\end{matrix}\right\} : \left\{\begin{matrix}\text{The number of times A} \\ \text{will use second strategy}\end{matrix}\right\} = \frac{d-c}{(a+d)-(b+c)} : \frac{a-b}{(a+d)-(b+c)}
$$

The expected pay-off to Player A

$$
\begin{aligned}
&= ap + c(1-p) \\
&= c + p(a-c) \\
&= c + \frac{(d-c)(a-c)}{(a+d)-(b+c)} \\
&= \frac{c\{(a+d)-(b+c)\} + (d-c)(a-c)}{(a+d)-(b+c)} \\
&= \frac{ac + cd - bc - c^2 + ad - cd - ac + c^2)}{(a+d)-(b+c)} \\
&= \frac{ad-bc}{(a+d)-(b+c)}
\end{aligned}
$$

Therefore, the value V of the game is

$$
\frac{ad-bc}{(a+d)-(b+c)}
$$

To find the number of times that B will use his first strategy and second strategy:

Let the probability that B will use his first strategy be r. Then the probability that B will use his second strategy is 1-r.

When A use his first strategy

The expected value of loss to Player B with his first strategy = ar

The expected value of loss to Player B with his second strategy = b(1-r)

Therefore the expected value of loss to B = ar + b(1-r) (3)

When A use his second strategy

The expected value of loss to Player B with his first strategy = cr

The expected value of loss to Player B with his second strategy = d(1-r)

Therefore the expected value of loss to B = cr + d(1-r) (4)
If the two expected values are different then it results in a pure game, which is a contradiction. Therefore the expected values of loss to Player B in equations (3) and (4) should be equal. Hence we have the condition

$$
\begin{aligned}
&ar + b(1-r) = cr + d(1-r) \\
&ar + b - br = cr + d - dr \\
&ar - br - cr + dr = d - b \\
&r(a - b - c + d) = d - b \\
&r = \frac{d-b}{a-b-c+d} \\
&\quad = \frac{d-b}{(a+d)-(b+c)}
\end{aligned}
$$

Problem 2

Solve the following game

$$
\begin{array}{cc}
 & Y \\
X & \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix}
\end{array}
$$

Solution

First consider the row minima.

Row	Minimum Value
1	2
2	1

Maximum of {2, 1} = 2

Next consider the maximum of each column

Column	Maximum Value
1	4
2	5

Minimum of {4, 5}= 4

We see that

Max {row minima} $\neq$ min {column maxima}

So the game has no saddle point. Therefore it is a mixed game.

We have a = 2, b = 5, c = 4 and d = 1.

Let p be the probability that player X will use his first strategy. We have

$$p = \frac{d-c}{(a+d)-(b+c)}$$

$$= \frac{1-4}{(2+1)-(5+4)}$$

$$= \frac{-3}{3-9}$$

$$= \frac{-3}{-6}$$

$$= \frac{1}{2}$$

The probability that player X will use his second strategy is

$1\text{-p} = 1\text{-} \frac{1}{2} = \frac{1}{2}$.

Value of the game V = $\frac{ad-bc}{(a+d)-(b+c)} = \frac{2-20}{3-9} = \frac{-18}{-6} = 3$.

Let r be the probability that Player Y will use his first strategy. Then the probability that Y will use his second strategy is (1-r). We have

$$
\begin{aligned}
r &= \frac{d-b}{(a+d)-(b+c)} \\
&= \frac{1-5}{(2+1)-(5+4)} \\
&= \frac{-4}{3-9} \\
&= \frac{-4}{-6} \\
&= \frac{2}{3}
\end{aligned}
$$

$$1-r = 1-\frac{2}{3} = \frac{1}{3}$$

Interpretation

$$\text{p} : (1\text{-p}) = \frac{1}{2} : \frac{1}{2}$$

Therefore, out of 2 trials, player X will use his first strategy once and his second strategy once.

$\text{r} : (1\text{-r}) = \frac{2}{3} : \frac{1}{3}$

Therefore, out of 3 trials, player Y will use his first strategy twice and his second strategy once.
Tags : Operations Management - Game Theory, Goal Programming & Queuing Theory

SIMULATION

Simulation is the process of designing a model of a real system and conducting experiments with this model for the purpose of understanding the behaviour of the system and/or evaluating various strategies for the **operation** of the system.

The Leading Types of Simulation Models

1. Monte Carlo / Risk Analysis Simulation

In simple terms, a Monte Carlo simulation is a method of risk analysis. Businesses use it prior to implementing a major project or change in a process, such as a manufacturing assembly line. Built on mathematical models, Monte Carlo analyses use the empirical data of the real system's inputs and outputs (e.g., supply intake and production yield). It then identifies uncertainties and potential risks through probability distributions.

The advantage of a Monte Carlo-based simulation is that it provides awareness and a thorough understanding of potential threats to your bottom-line and time-to-market.

You can implement Monte Carlo simulations to practically any industry or field, including oil and gas, manufacturing, engineering, supply chain management, and many others.

2. Agent-Based Modeling & Simulation

An agent-based simulation is a model that examines the impact of an 'agent' on the 'system' or 'environment.' In simple terms, just think of the impact a new laser-cutter or some other factory equipment has on your overall manufacturing line.

The 'agent' in agent-based models could be people, equipment, and practically anything else. The simulation includes the agent's 'behavior,' which serve as rules of how those agents must act in the system. You then look at how the system responds to those rules.

However, you must draw your rules from real-world data — otherwise, you will not generate accurate insights. In a way, it serves as a means to examine a proposed change and identify potential risks and opportunities.

3. Discrete Event Simulation

A discrete event simulation model enables you to observe the specific events that result in your business processes. For example, the typical technical support process involves the end-user calling you, your system receiving and assigning the call, and your agent picking up the call. You would use a discrete event simulation model to examine that technical support process. You can use discrete event simulation models to study many types of systems (e.g., healthcare, manufacturing, etc), and for a diverse range of outcomes.

For example, the Nebraska Medical Center had used discrete event simulation models to see how it could remove workflow bottlenecks, increase the utilization of its operating rooms, and lower patient/surgeon travel distance and time.

4. System Dynamics Simulation Solutions

This is a very abstract form of simulation modeling. Unlike agent-based modeling and discrete event modeling, system dynamics does not include specific details about the system. So for a manufacturing facility, this model will not factor in data about the machinery and labor.

Rather, businesses would use system dynamics models to simulate for a long-term, strategic-level view of the overall system.

In other words, the priority is to get aggregate-level insights about the entire system in response to an action — e.g., a reduction in CAPEX, ending a product line, etc

SIMULATION INVENTORY AND WAITING LINES.

Waiting line systems, also called queuing systems from the underlying **modeling** basis of queuing theory, involve a population source, an arrival process, a **waiting** area, and a service area or channel. **Waiting line** systems also have costs, operating characteristics, and management response strategies.

Waiting in **lines** is a part of our everyday life. **Waiting** in **lines** may be due to overcrowded, overfilling or due to congestion. Any time there is more customer demand for a service than **can** be provided, a **waiting line** forms.

There are two main **types of waiting lines**: finite, where new customers can only be added to the **line** once others move out of the **line**, and infinite, where new customers are not affected by the number of customers already in the **system**. **Lines** can also be set up so that there is one single or multiple **lines** of service.

Structure of Waiting Line Systems

Perhaps the most significant difference between systems that produce products and systems that produce services is that product manufacturers can buffer their manufacturing processes from customers through use of inventories. In fact, a major category of inventory is called "buffer" or safety stock as we will see in the next module of this course.

On the other hand, in pure service systems, where the customer receives services directly from the service provider, the customer is in the "boundary" of the provider. In this case, when demand for service exceeds the capacity for service, one or more waiting lines form or customers leave the system un-served. Medical doctors cannot "inventory" medical care; lawyers and other professional counselors cannot "inventory" consultations; and airports cannot "inventory" aircraft parking positions at terminals.

My first experience with waiting line systems was in 1969 when I was Commander of an Aerial Port Detachment at Ramstein Air Base Germany. One of my first tasks was to design a parking plan for aircraft involved in the evacuation of troops, equipment, bombs and other explosives from a fighter training facility in Libya. Waiting line models helped me determine how many parking positions we would need (the waiting line) to avoid holding aircraft on the runway (also the waiting line), until cargo handlers (service providers) could unload and turnaround the aircraft. The models worked fairly well, although I should have done more sensitivity analysis to handle increases in the arrival rate as things got a little "hot" the last day of the evacuation in Libya.

Enough nostalgia - on to the basic structure of waiting line systems. Waiting line systems, also called queuing systems from the underlying modeling basis of queuing theory, involve a ***population source***, an ***arrival process***, a ***waiting area***, and a ***service area*** or ***channel.***

Waiting line systems also have ***costs***, ***operating characteristics***, and ***management response strategies***.

Waiting Line System Costs and Management Strategies

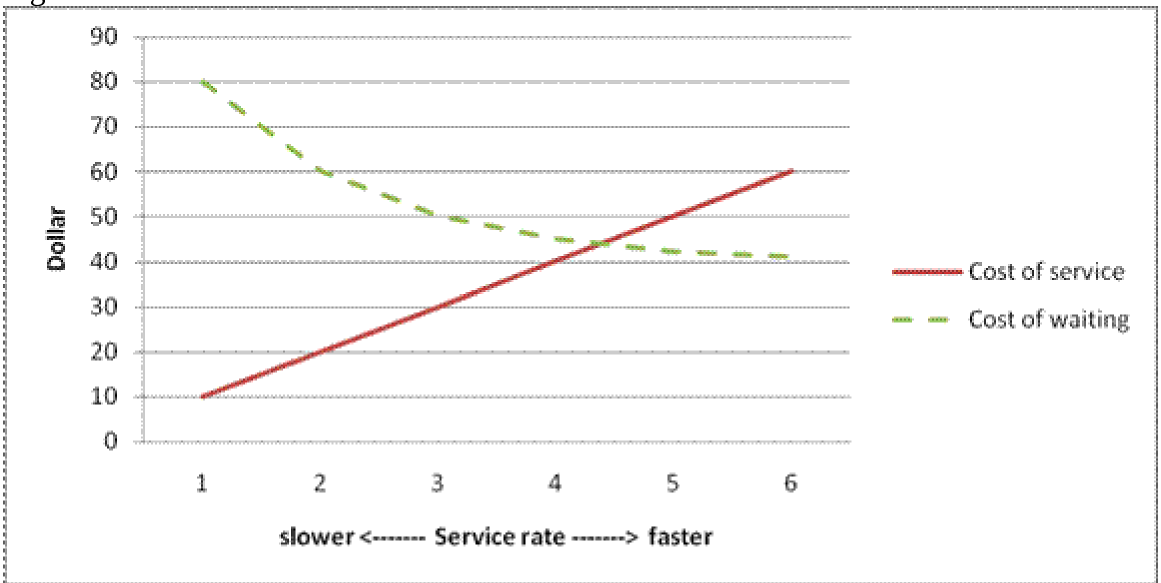
There are basically two costs that must be balanced in waiting line system - the cost of ***service*** and the cost of ***waiting***. Note that I am not considering another possible cost component - the cost of a scheduling system. Theoretically, a scheduling system is a management strategy designed to avoid waiting lines (meaning you should ***never*** wait in the doctor's office - yeah, right!) and is not covered in this module. Scheduling systems are useful when the customer is known to the system and the short and long run costs of waiting are relatively high. We will study scheduling system applications in linear programming later on in the course.

Rather, in this module, we will study the more interesting scenario where customer arrivals to the system are random - that is, the customer is not known before arriving to the system and the arrival process is random as we will discuss later. In this case, to avoid longer than desired waiting lines, the service provider may exercise management strategies to increase the service rate by using faster servers, more servers, automated service, or some combination of strategies. Each of these strategies increases the cost of service, but produces the benefit of reducing the cost of waiting.

When the cost of service and the cost of waiting are known and measurable, the waiting line models in this set of module notes help us determine the optimal, or close to optimal waiting system configuration and rate of service. This can be tricky because the cost of service has a positive relationship with the rate of service, whereas the cost of waiting has a negative relationship with the rate of service. That is, the faster the service rate (or the more service), the higher the cost of service; whereas the faster the service rate, the lower the cost of waiting. Of course, the opposite is true: the slower the service rate, the lower the cost of service but the higher the cost of waiting.

The cost relationship can be sketched out in a graph, with dollars on the vertical axis and the rate of service on the horizontal axis, as shown in Figure 3.1.1.

Figure 3.1.1



We will examine the cost of service and cost of waiting components as we examine the main waiting line models later. It should be cautioned, however, that it is often difficult to measure the cost of waiting for customers that are external to the service provider's company or organization. For example, if customers in a car repair waiting line system are defined as mechanics waiting for tools from a tool crib to repair a car, it is fairly easy to measure the cost of mechanic waiting since the repair company employs the mechanic. The cost of waiting may simply be the time spent idle in the line times the employee's salary for that time unit.

However, if the owner of the car being repaired is defined as the customer of the waiting line system, the cost of waiting may be more difficult to measure. Retired "snow birds" may have discretionary time to spend in a waiting room and their cost of waiting may be relatively low. On the other hand, for someone who is employed and depends on their car to get to work, the cost of waiting may be relatively high, depending on their income level.

If the cost of waiting also incorporates the cost of losing a customer because long waiting time drives the customer away, measurement becomes even more difficult. As an alternative to managing the waiting system by measuring, analyzing and minimizing its combined total costs, the service provider may try to manage the service system by setting threshold parameters for system operating characteristics, and then use faster servers, more servers, automation of the service activity or some combination of strategies to achieve those parameters.

For example, a consumer products store may decide to open another cash register checkout station when the number of customers in line at the first register goes over six; or a bank may decide to open another teller position when the waiting time in line exceeds five minutes. Other waiting line system operating characteristics are discussed next.

Waiting Line System Operating Characteristics

Operational characteristics of waiting lines include:

1. the probability that no customers (or units) are in the system,

2. the average number of customers in the lines,

3. the average number of customers in the system (customers in line plus those being served,

4. the average time a customer spends in the waiting line,

5. the average time a customer spends in the system (waiting time plus time in the service facility,

6. the probability that an arriving customer has to wait for service,

7. the probability of n customers in the system, where n could be any real integer such as 1 customer, 2, 3, ...

We will examine the operating characteristics of each of the main waiting line models presented in this module. Before that, we will examine the main features of the structure within a waiting line system: the population, arrival process, waiting line configuration, service area and its configuration, and exit.

The Population

The population that generates customers to waiting line systems may be infinite or finite. In most cases, populations can be considered infinite, even though they are really finite. For example, if we were to study the characteristics of waiting lines forming at the Fort Myers side of the Cape Coral Bridge at morning rush hour, we know that the population generating the arrivals is the finite population of Cape Coral, around 100,000. However, since there is no actual limit placed on the customers arriving at the toll booths, ***we assume the population is infinite***. All but one of the models we will study make this assumption. The finite model is the appropriate model to use when the population is relatively small, such as 20 total computers in a network office that feed a computer repair person server.

Arrival Process

Arrivals to the waiting line system from the population source may be on an individual or batch basis. ***We will assume arrivals are on an individual basis***. The difference is best illustrated by the arrival of a car to a parking lot at a restaurant. One driver leaving the car to enter the restaurant would represent the arrival of one unit or customer to the waiting line system. If a bus pulls in, there could be a batch arrival of 30 customers. Did you ever notice that the bus stalls are ***behind*** the Cracker Barrel Restaurants on the interstate highways - just so you can't see all those batch arrivals before you pull off!

It is also assumed that the arrivals are ***nonscheduled***, and the arrival of one unit is ***independent*** of, or does not impact, the arrival of other units. Whenever these assumptions are made, arrivals are assumed to follow the Poisson Probability Distribution, a member of the family of discrete probability distributions. The Poisson Probability Distribution is completely described by its mean, which is given the Greek symbol lambda. In a waiting line system, the mean we are referring to is the ***mean arrival rate***. For example, we may say that the mean arrival rate is 4 calls per hour to a catalog company's telephone bank.

Another way of representing the mean arrival rate is to take its inverse, which gives us the ***mean time between arrivals***. So, if I invert the mean rate of 4 calls per hour, I get ¼ hours. The mean time between arrivals is $1/4^{th}$ of an hour, or one arrival every 15 minutes.
Mean Time Between Arrivals = 1 / Mean Arrival Rate
The probability distribution that is used to describe this time between arrivals in a waiting line system is the Exponential Distribution. The Exponential Distribution is used to model the probabilities of continuous variables such as time, in the case of waiting line systems. The Greek Symbol Mu, is used to describe the mean of the Exponential Distribution.

If you have the mean time between arrivals, you can find the mean arrival rate by the similar procedure - taking the inverse. For example, what if we knew that the average time between arrivals to a bank teller was 5 minutes. The mean arrival rate would be computed as follows:
Mean Arrival Time = 5 minutes = $5/60^{th}$ hours

Mean Arrival Rate = 60/5 = 12 customers per hour
Customers arriving from a population next join the waiting line in the waiting line system.

Waiting Line Configuration

Waiting lines may be infinite or truncated. For all of the models we will examine except one, we will ***assume infinite line length***. One of the models we will examine is designed to model situations where no waiting is allowed - the ultimate of truncated systems. My telephone allows no waiting - if I am talking to someone, the next caller gets a busy signal. However, airline reservation systems allow callers to a busy reservation agent to wait in a queue.

The waiting line system may use a ***single line/single server channel*** configuration which means one line forms in front of a single service channel. The word ***service channel*** is used rather than server to avoid confusion. A single service channel may have many servers, but room for only one customer. That is called a single service channel. Grocery stores have multiple single line/single channel configurations. Banks, on the other hand, employ a ***single line/multiple channel*** or teller configuration.

Customers ***discipline*** in the waiting line configuration may vary from ***patient***, to ***balk*** (view the line, then leave), ***renege*** (join the line, then leave), ***jockey*** (join the line, then move to another line when you think it is moving faster - that's me!), or collude (give your groceries to another customer - scum of the earth to students of quantitative methods)! Waiting line models ***assume that the customer is patient*** (my wife).

Departures from the waiting line to the server are ***assumed to be first in, first out*** or first come, first served rather than last in first out, or like my deli - service in random order!

Service Channel

The main feature of a service channel in a waiting line system is the service time, also ***assumed to follow the Exponential Distribution*** when the time to perform service for one customer is independent from the time to perform service for others. If the average time to serve one customer is 10 minutes, then the mean service rate in hours may be found by converting 10 minutes to hours ($10/60^{th}$ of an hour), then taking the inverse:
Mean Service Rate = 60/10 = 6 customers per hour
Note that rates are always stated as per hour, … per minute, … or per whatever time unit; whereas service times and times between arrivals are stated simply as hours, minutes, or whatever time unit.

Once service is completed, it is assumed that customers ***exit the system*** and return to the population. Of course, they may exit one system and feed another. If the second system is independent of the first, then there may be two separate and distinct single line/single server systems.

That finishes our coverage of the basis structure of the waiting line system. Waiting line models have been designed as quantitative methods to analyze the operating characteristics and costs of waiting line systems. The models are categorized by the probability distributions that describe the arrival rate and service time processes, the number of channels, and whether the population is infinite or finite. The first model we will examine is one which follows the structure of a single service/single channel system.

3.2 Single Server/Single Channel System with Poisson Arrivals and Exponential Service Times

The title of this section is a long way of describing a very common waiting line system: a single line forming in front of a single server. The quantitative methods described in this section are used to compute the operating characteristics and costs of this system.

Pause and Reflect

The assumptions needed for using quantitative methods to analyze operating characteristics of this single server/single channel system include an infinite population, Poisson arrival rates, infinite line length, patient customer discipline, FIFO departure from the waiting line, Exponential Service times and customer departure to the population.

Let's examine the following example problem. A document clerk earns $15/hour, processes an average of 5 documents per hour throughout the workday, and receives documents at the rate of 4 per hour. There is a cost of waiting for the documents (by an office paralegal) of $25.00 per hour. For this problem, the mean or average service rate is **mu** = 5 documents per hour, and the mean or average arrival rate is **lambda** = 4 documents per hour.

The law firm is interested in knowing the operating characteristics of this waiting line system, such as average time a document spends in the waiting line, number of documents in the waiting line, the utilization rate of the document clerk, the probability of 3 documents in the system, and the cost of the system. The following formulas are used to model the operating characteristics of the single line/single channel waiting line system with Poisson Arrivals and Exponential Service time.

1. The probability of an idle system (no documents in the system, P_0):

P_0 = 1 - (lambda/mu) = 1 - (4/5) = 0.20

There is a 20% chance that the system will be idle at any one point in time.

Note: look at the formula carefully to see that we divide lambda, the mean arrival rate, by mu, the mean service rate. Recall that the arrival rate must be less than the service rate or the waiting line explodes. This can be shown mathematically in the formula. If lambda were greater than mu, the probability of an idle system would be a negative number greater than one which is infeasible.

2. The average number of documents in the waiting line (L_q):

L_q = lambda^2/ [mu(mu - lambda)] = 4^2 / [5 (5-4)] = 3.2 documents

3. The average number of units in the system (L):

L = L_q + (lambda / mu) = 3.2 + (4/5) = 4 documents

From the formula, you can see that this includes the documents in line as well as in the service channel.

4. The average time a document spends in the waiting line:

W_q = L_q / lambda = 3.2 / 4 = 0.8 hours or 48 minutes.

Note:

The initial time units of this and other time-related operating characteristics are the same as the input units. If input arrival and service times and rates are in hours, then output times will be in hours. I converted 0.8 hours to minutes for ease in interpretation.

5. The average time a document spends in the system:

$W = W_q + (1/mu) = 0.8 + (1/5) = 0.8 + 0.2 = 1$ hour

As with the number of units in the system parameter, this time includes the time a document spends in the line, plus the time in service.

6. The probability that an arriving unit has to wait for service:

$P_w = lambda / mu = 4/5 = 0.80$

This operating characteristic is also known as the utilization factor for the service channel. A critical requirement for single line/single channel service systems is that the utilization factor be less than one. Otherwise, the waiting line would explode. In fact, queuing systems are not very efficient anytime the utilization factor exceeds 75 or 80% due to the interaction of the two probability functions. We will illustrate this later.

7. The probability of n documents in the system. Let's say we are interested in knowing the probability of n = 3 documents in the system:

$P_n = (lambda/mu)^n * P_0 = (4/5)^3 * 0.2 = 0.1024$

Since we are working with discrete probabilities, to find the probability of three or less documents in the system:

$P_n \leq 3 = P_0 + P_1 + P_2 + P_3$

$= 0.20 + (4/5)^1 * 0.2 + (4/5)^2 * 0.2 + 0.1024$

$= 0.20 + 0.16 + 0.128 + 0.1024 = 0.5904$

…and we follow the law of probability that says all probabilities of the distribution must sum to one, so the probability that there will be more than 3 documents in the system is:

$P_n > 3 = 1 - 0.5904 = 0.4096$

Costs of the Waiting Line System

The total cost of this waiting line system is the sum of the cost of waiting and cost of service.

Total Cost = Cost of Waiting + Cost of Service

$= (c_w L) + (c_s k)$ where

k =number of channels

c_w = cost of waiting = $25.00 per hour for 1 paralegal

c_s = cost of service = $15.00 per hour for clerk

Total Cost = ($25 * 4) + ($15 * 1) = $115

Sensitivity Analysis

Recall that I said with purely random arrivals waiting line systems are most stable with utilization rates less than or equal to 75%. That may seem unusual to you, that we only want to work our employees at 75% utilization. The problem is that the arrival rate and service time parameters are averages. The variation is always what kills us.

For example, the average arrival rate is 4 documents per hour, and the service rate is 5 documents per hour. At any point in time, the system may experience an arrival rate of 5 documents per hour and would be at the verge of ***explosion*** unless at the next point in time, the arrival rate slows down to 3 documents per hour. And, this considers that the service rate remains constant, which it doesn't in this model. The service rate could slow down to 4 documents per hour with an arrival rate of 4, and we have the same problem. Of course, we could also have a speed up in the arrival rate and a slow down in the service rate to result in chaos.

To buffer against explosion, which occurs when the waiting grows and grows and we never catch up, waiting line systems with purely random arrivals are generally kept to a maximum utilization rate of 75%. Note that the above example illustrated a system with a utilization rate of 80%. What if the arrival rate increases to 4.5 documents per hour, which represents a 12.5% increase. The new utilization rate is:

P_w = lambda / mu = 4.5 / 5 = .90, a 12.5% increase as well.

The line length now increases to:

L_q = lambda^2/ [mu(mu - lambda)] = 4.5^2 / [5 (5-4.5)] = 8.1

documents, a 150% increase!

That's why we don't operate at utilization rates above 75%. I really find this interesting from my operational experience. People who say they are working their people at 100% utilization are not working in random service systems - they can't be or the system would be pure chaos all the time.

The formulas for the operating characteristics and costs of this model are relatively simple. As we begin to change the assumptions and develop more complicated models, it is important to understand relationships and concepts, and then rely on the software to do the number crunching, in my opinion.

UNIT – V
P.E.R.T. & C.P.M. AND REPLACEMENT MODEL

Learning Objectives

P.E.R.T. & C.P.M. and Replacement Model: Drawing networks – identifying critical path – probability of completing the project within given time- project crashing – optimum cost and optimum duration. Replacement models comprising single replacement and group replacement.

Network analysis is concerned with minimizing some measure of performance of the system such as the total completion time for the project, overall cost and so on. By preparing a network of the system, a decision maker can identify,

(i) The physical relationship (properties) of the system

(ii) The inter relationships of the system components

Network analysis is specially suited to project which are not routine or repetitive and which will be conducted only once or a few times.

Objectives:

Network analysis can be used to serve the following objectives:

1. Minimization of total time: Network analysis is useful in completing a project in the minimum possible time. A good example of this objective is the maintenance of production line machinery in a factory. If the cost of down time is very high, it is economically desirable to minimize time despite high resource costs.
2. Minimization of total cost: Where the cost of delay in the completion of the project exceeds cost of extra effort, it is desirable to complete the project in time so as to minimize total cost.
3. Minimization of time for a given cost: When fixed sum is available to cover costs, it may be preferable to arrange the existing resources so as to reduce the total time for the project instead of reducing total cost.
4. Minimization of cost for a given total time: When no particular benefit will be gained from completing the project early, it may be desirable to arrange resources in such a way as to give the minimum cost for the project in the set time.
5. Minimization of idle resources: The schedule should be devised to minimize large fluctuations in the use of limited resources. The cost of having men/machines idle should be compared with the cost of hiring resources on a temporary basis.
6. Network analysis can also be employed to minimize production delays, interruptions and conflicts.

Managerial Applications :

Network analysis can be applied to very wide range of situations involving the use of time, labour and physical resources. Some of the more common applications of network analysis in project scheduling are as follows:

1. Construction of bridge, highway, power plant etc.
2. Assembly line scheduling.
3. Installation of a complex new equipment. Eg., computers, large machinery.
4. Research and Development
5. Maintenance and overhauling complicated equipment in chemical or power plants, steel and petroleum industries, etc.
6. Inventory planning and control.
7. Shifting of manufacturing plant from one site to another.
8. Development and testing of missile system.
9. Development and launching of new products and advertising campaigns.
10. Repair and maintenance of an oil refinery.
11. Construction of residential complex.
12. Control of traffic flow in metropolitan cities.
13. Long range planning and developing staffing plans.
14. Budget and audit procedures.
15. Organization of international conferences.
16. Launching space programmes, etc.

A network is a graphic representation of a project's operations and is composed of activities and events (or nodes) that must be completed to reach the end objective of a project, showing the planning sequence of their accomplishments, their dependence and inter relationships.

Basic Components

Events (node)

A specific point in time at which an activity begins and ends is called a node. It is recognizable as a particular instant in time and does not consume time or resource. An event is generally represented on the network by a circle, rectangle, hexagon or some other geometric shape.

Activity

An activity is a task, or item of work to be done, that consumes time, effort, money or other resources. It lies between two events, called the 'Preceding' and 'Succeeding' ones. An activity is represented on the network by an arrow with its head indicating the sequence in which the events are to occur.

Head event Tail event

Preceding event Succeeding event

Predecessor Activity:

An activity which must be completed before one or more other activities start is known as Predecessor activity.

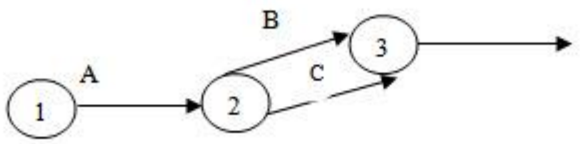
Successor Activity:

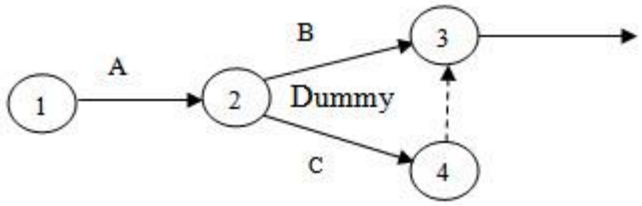
An activity which started immediately after one or more of other activities are completed is known as Successor activity.

Dummy Activity:

Certain activities which neither consume time nor resources but are used simply to represent a connection between events are known as dummies. A dummy activity is depicted by dotted line in the network diagram.

A dummy activity in the network is added only to represent the given precedence relationships among activities of the project and is needed when (a) two or more parallel activities in a project have same head and tail events, or (b) two or more activities have some (but not all) of their immediate predecessor activities in common.



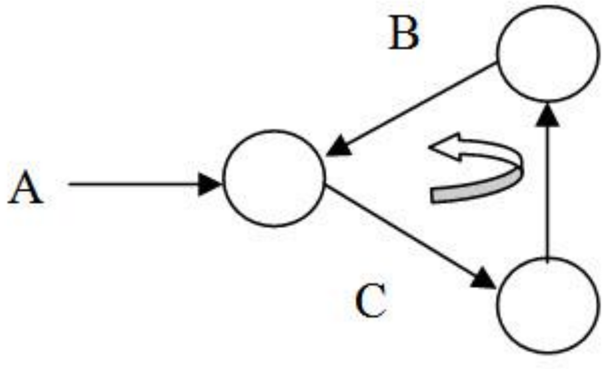


Common Errors

There are three types of errors which are most common in network drawing, viz.,

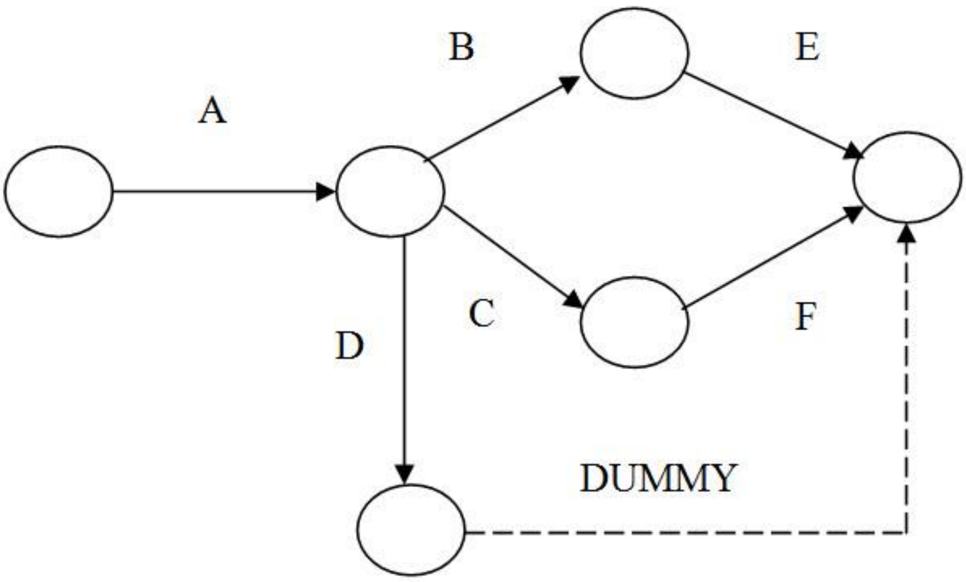
(a) Formation of a loop, (b) Dangling, and (c) Redundancy.

(a) **Formation of a loop:** If an activity were represented as going back in time, a closed loop would occur. This is show in fig which is simply the structure of Fig (b) with activity B reversed in direction. Cycling in a network can result through a simple error or when while developing the activity plans, one tries to show the repetition of an activity before beginning the next activity.



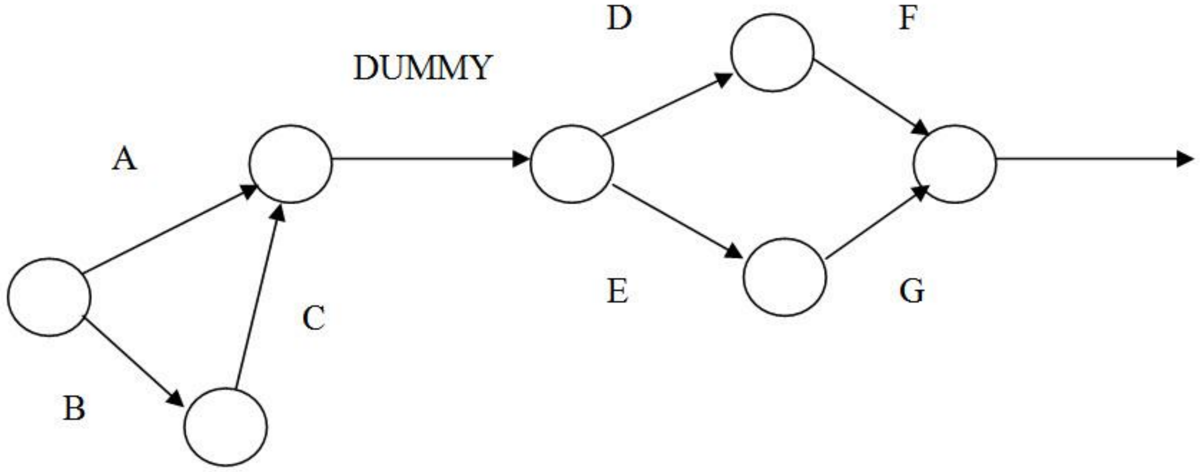
A closed loop would produce an endless cycle in computer programmes without a built-in routine for detection or i4-entification of the cycle; Thus one property of a correctly constructed network diagram is that it is "non-cyclic".

(a) **Dangling:** No activity should end without being joined to the end event. If it is not so, a dummy activity is introduced in order to maintain the continuity of the system. Such end-events other than the end of the project as a whole are called dangling events.



In the above network, activity D leads to dangling. A dummy activity is therefore introduced to avoid this dangling.

(c) **Redundancy**: If a dummy activity is the only activity emanating from an event, it can be eliminated. For example, in the network show in Fig the dummy activity is redundant and can be eliminated, and the network redrawn.



<u>Rules of Network Construction</u>

1. Each activity is represented by one & only one arrow so that no single activity can be represented twice in the network.
2. Time follows from left to right. Arrows pointing in opposite directions must be avoided.
3. Arrows should be kept straight and not curved or bent.
4. Use dummies freely.
5. Every node must have at least one activity preceding it and at least one activity following it, except that the beginning node has no activities before it and the ending node has no activities following it.
6. Only one activity may connect any two nodes. This rule is necessary so that an activity can be specified by giving the numbers of its beginning and ending nodes.

Numbering the Events

After the network is drawn in a logical sequence, every event is assigned a number. The number sequence must be such so as to reflect the flow of the network. In event numbering, the following rules should be observed:

1. Events numbers should be unique
2. Event numbering should be carried out an a sequential basis from left to right.
3. The initial event which has all outgoing arrows with no incoming arrow is numbered 0 or 1.
4. The head of an arrow should always bear a number higher than the one assigned at the tail of the arrow.
5. Gaps should be left in the sequence of event numbering to accommodate subsequent inclusion of activities, if necessary.

Example:

A television is manufactured in six steps, labeled A through F. Because of its size and Complexity, the television is produced one at a time. The production control manager thinks that network scheduling techniques might be useful in planning future production. He recorded the following information:

A is the first step and precedes B and C

C precedes D and E

B follows D and precedes E

F follows E

D is successor of F

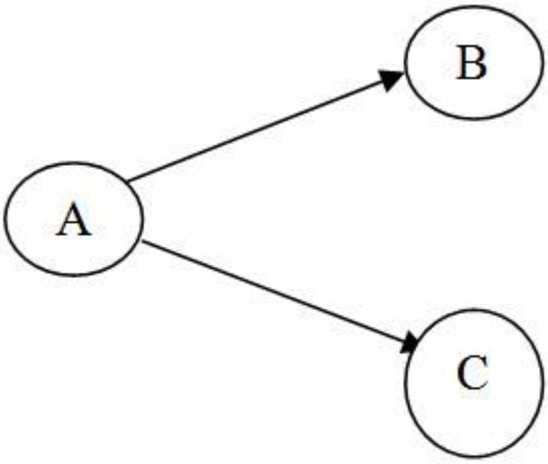
(i) Draw an activity-on-node diagram for the production manager.

(ii) On checking with the records, the production manager corrects his last note to read, “D is a predecessor of F”. Draw a revised diagram of this network incorporating this new change.

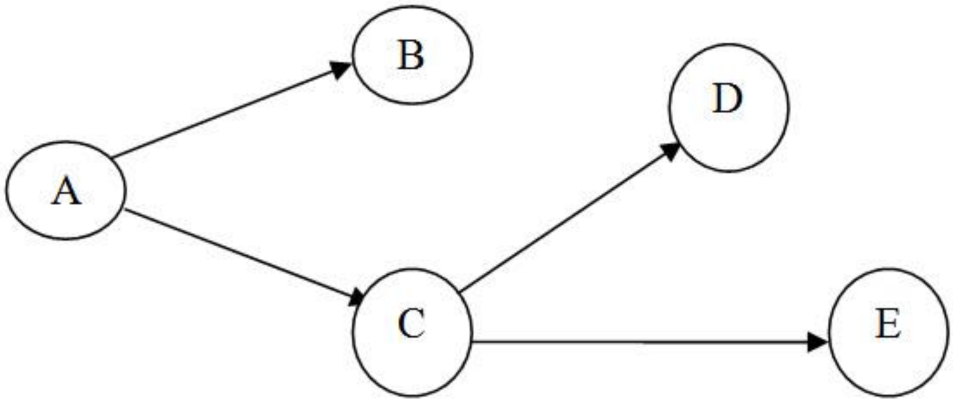
Solution:

(a)

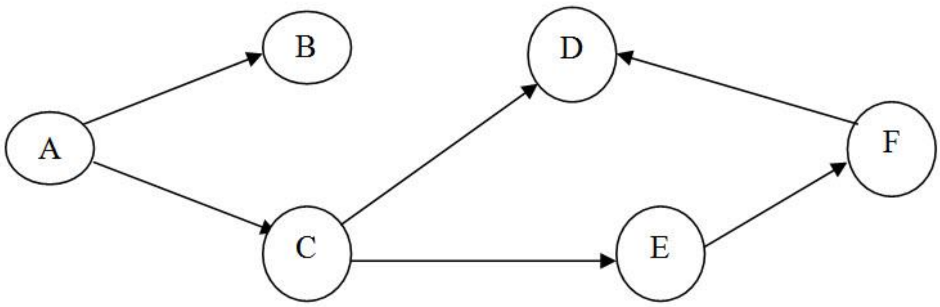
A is the first step which follows B and C



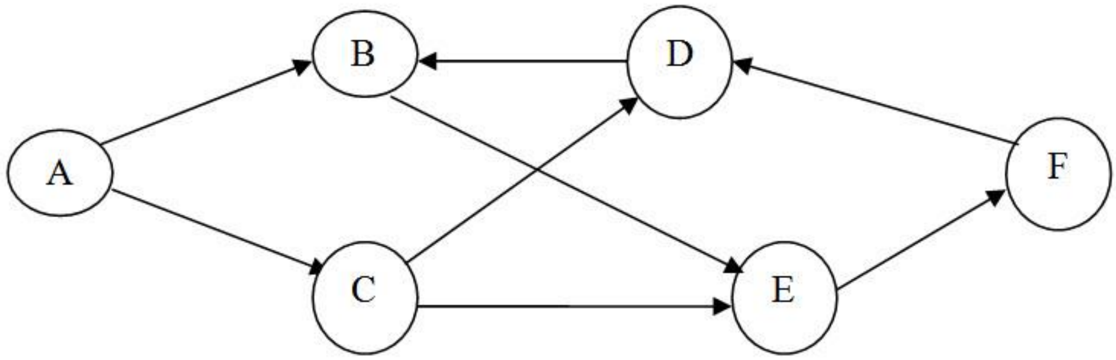
C precedes D and E



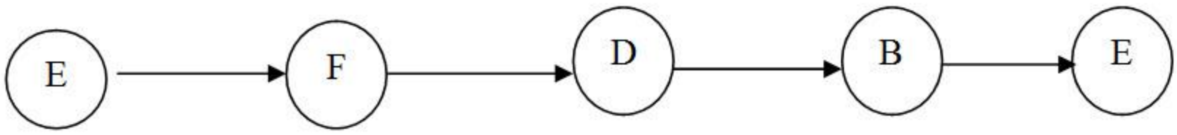
F follows E and D and is the successor of F



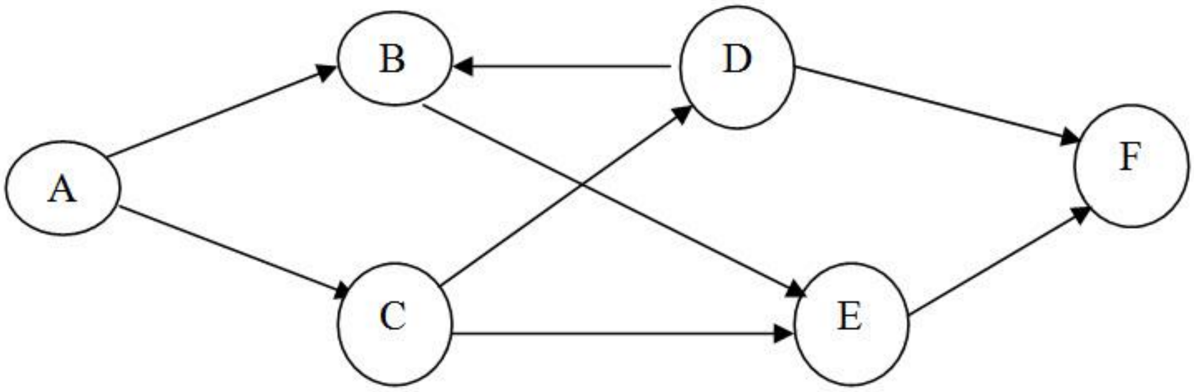
Now since B follows D and precedes E, the complete network diagram is shown below.



Evidently, this network contains a cycle as shown below



(b) Revised network when D is a predecessor of F is as follows:



P.E.R.T. & C.P.M. AND REPLACEMENT MODEL

Time Calculations in Networks

For each activity an estimate must be made of time that will be spent in the actual accomplishment of that activity. Estimates may be expressed in hours, days, weeks or any other convenient unit of time. The time estimate is usually written in the network immediately above the arrow. The next step after making the time estimates is the calculation of earliest times and latest times for each mode. These calculations are done in the following way.

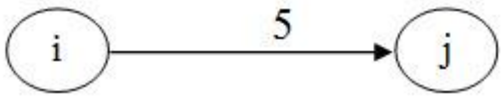
a) Let zero be the starting time for the project. Then for each activity there is an earliest starting time (ES) relative to the project starting time. The earliest finishing time is denoted by

Thus the formula is

EF_i (or) $ES_j = \max \{ES_i + t_{ij}\}$

Where Es_j denotes the earliest start time of all the activities emanating from node i and t_{ij} is the estimated duration of the activity i-j.

Example:



In the above example the activity is from i-j, the duration of time is 5 hours. Here start time is $ES_j = \max \{ES_i + t_{ij}\}$

Initial start time $ES_i=0$.

$ES_j = \max \{0+ 5\} =5$.

Initially the starting time will be 0 The finishing time for the i^{th} event is 5.Staring time for the j^{th} event is 5.

b) Let us suppose that we have a target time for completing the project. Then this time is called the latest finish time (LF) for the final activity. The latest start time (LS) is the latest time at which an activity can start if the target is to be maintained. It means that for the final activity, its LS is simply LF – activity time.

$LF_i = \min \{LF_j - t_{ij}\}$, for all defined (i, j) activities.

Critical Path:

Certain activities in a network diagram of a project are called critical activities because delay in their execution will cause further delay in the project completion time. Thus, **all activities having zero total float value are identified as critical activities.**

The critical path is the continuous chain of critical activities in a network diagram. It is the longest path starting from first to the last event and is shown by a thick line or double lines in a network.

The length of the critical path is the sum of the individual times of all the critical activities lying on it and defines the minimum time required to complete the project.

The critical path on a network diagram can be identified as:

(a) $ES_i = LF_i$

(b) $ES_j = LF_j$

(c) $ES_j - ES_i = LF_j - LF_i = t_{ij}$.

Critical Path Method (CPM)

The iterative procedure of determining the critical path is as follows:

Step 1: List all the jobs and then draw a network diagram. Each job is indicated by an arrow with the direction of the arrow showing the sequence of jobs. The length of the arrows has no significance. Place the jobs on the diagram one by one keeping in mind what precedes and follows each job as well as what job can be done simultaneously.

Step 2: Consider the job's times to be deterministic. Indicate them above the arrow representing the task.

Step 3: Calculate the earliest start time (EST) and earliest finish time (EFT) for each event and write them in the box marked $\boxed{\surd}\boxed{}$ Calculate the latest start time (LST) and latest finish time (LFT) and write them in the box marked $\boxed{}\boxed{\surd}$.

Step 4: Tabulate various times, i.e., activity normal times, earliest times and latest times, and mark EST and LFT on the arrow diagram.

Step 5: Determine the total float for each activity by taking differences between EST and LFT.

Step 6: Identify the critical activities and connect them with the beginning node and the ending node in the network diagram by double line arrows. This gives the critical path.

Step 7: Calculate the total project duration.

Slack / Float of an Activity and Event

The float (Slack) or free time is the length of time to which a non-critical activity and the time between its ES & LF is longer than its actual duration or an event can be delayed or extended without delaying the total project completion time.(ie) the difference between the latest finish and earliest start time.

There are four types of floats namely

a) Total float

b) Free float

c) Independent float

d) Interference float

a) Total float

Difference between the latest finish and earliest finish time for the activity

Total float = $TF_{ij} = LF_j - EF_j$

b) Free float

It is defined by assuming that all the activities start as early as possible. The free float for the activity (i, j) is the excess available time over its duration.

$$LF_{ij} = ES_j – ES_i - t_{ij}$$

c) Interference float

The difference between total float and free float.

d) Independent float

The time by which an activity can be rescheduled without affecting the preceding or the succeeding activities is known as independent float.

Independent float = Free float – Tail event Slack

Advantages of Critical Path Method(CPM)

1. CPM was developed for conventional projects like construction project which consists of well know routine tasks whose resource requirement and duration were known with certainty.
2. CPM is suited to establish a trade off for optimum balancing between schedule time and cost of the project.
3. CPM is used for projects involving well know activities of repetitive in nature.

However the distinction between PERT and CPM is mostly historical.

Problem

CPM

The following table gives the activities of a construction project and duration.

Activity	1-2	1-3	2-3	2-4	3-4	4-5
Duration (days)	20	25	10	12	6	10

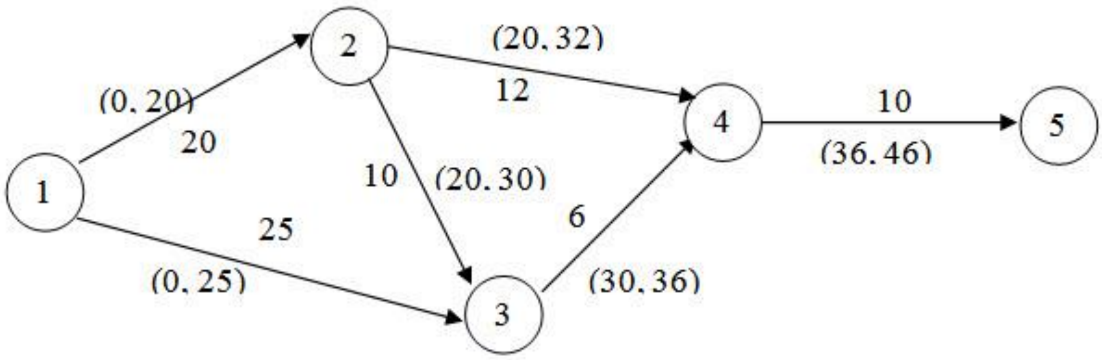
(i) Draw the network for the project.

(ii) Find the critical path.

(iii) Find the total, free and independent floats each activity.

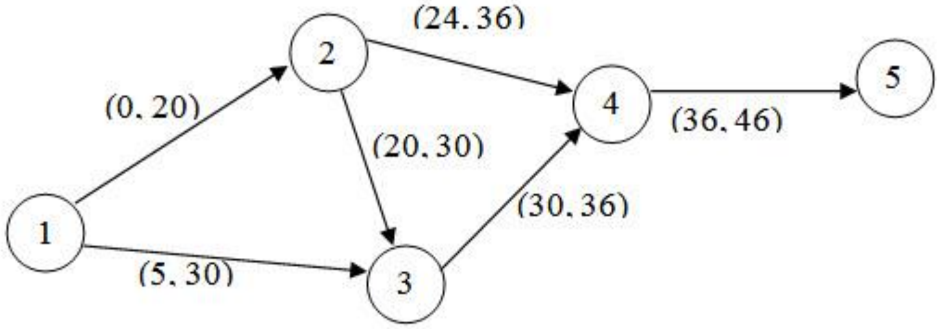
Solution:

The first step is to draw the network and fix early start and early finish schedule and then late start-late finish schedule as in figure 1 and figure 2.



Early – Start Early – Finish Schedule

Fig. 1



Late – Start Late – Finish Schedule

Fig.-2

Activity	Total Slack	Free Slack	Independent Slack
1-2	0	0	0
1-3	5	5	5
2-3	0	0	0
2-4	4	4	4
3-4	0	0	0
4-5	0	0	0

To find the critical path, connect activities with) total slack and we get 1-2-3-4-5 as the critical path.

Check with alternate paths.

1-2-4-5 = 42

1-2-3-4-5 = 46*

1-3-4-5 = 41

PROGRAMME EVALUATION REVIEW TECHNIQUE: (PERT)

This technique, unlike CPM, takes into account the uncertainty of project durations into account.

Deterministic network methods assume that the expected time is the actual time taken. Probabilistic methods, on the other hand, assume the reverse, more realistic situation, where activity times are represented by a probability distribution. This probability distribution of activity time is based upon three different time estimate made for each activity. There are as follows .

Optimistic (least) time estimate : (t_0 or a) is the duration of any activity when everything goes on very well during the project. i.e., laborers are available and come in time, machines are working properly, money is available whenever needed, there is no scarcity of raw materials needed etc.

Pessimistic (greatest) time estimate: (t_p or b) is the duration of any activity when almost every thing goes against our will and a lot of difficulties is faced while doing a project.

Most likely time estimate: (t_m or m) is the duration of any activity when sometimes things go on very well, sometimes things go on very bad while doing the project.

4.2 PERT Procedure

1. Draw the project net work
2. Compute the expected duration of each activity $t_e = {1 \over 6}\left[{4{t_m} + {t_0} + {t_p}} \right]$
3. Compute the expected variance $\sigma^2={\left({{{{t_p} - {t_0}} \over 6}} \right)^2}$ of each activity.
4. Compute the earliest start, earliest finish, latest start, latest finish time for each activity.
5. Determine the critical path and identify critical activities.
6. Compute the expected variance of the Project length (also called the variance of the critical path) σ^2 which is the sum of the variances of all the critical activities.
7. Compute the expected standard deviation of the project length and calculate the standard normal derivate ${{\rm{Z}}_{\rm{0}}}={{{\rm{duedate}} - {\rm{expected\ date\ of\ completion}}} \over {\sqrt {{\rm{project\ variance}}} }}$
8. Using (7) one can estimate the probability of completing the project within a specified time, using the normal curve (Area) tables.

4.3 Basic difference between PERT and CPM

PERT

1. PERT was developed in a brand new R and D Project it had to consider and deal with the uncertainties associated with such projects. Thus the project duration is regarded as random variable and therefore probabilities are calculated so as to characterize it.
2. Emphasis is given to important stages of completion of task rather than the activities required to be performed to reach a particular event or task in the analysis of network. i.e., PERT network is essentially an event- oriented network.
3. PERT is usually used for projects in which time estimates are uncertain. Example : R&D activities which are usually non-repetitive.
4. PERT helps in identifying critical areas in a project so that suitable necessary adjustments may be made to meet the scheduled completion date of the project.

Critical Path Method(CPM)

1. CPM was developed for conventional projects like construction project which consists of well know routine tasks whose resource requirement and duration were known with certainty.
2. CPM is suited to establish a trade off for optimum balancing between schedule time and cost of the project.
3. CPM is used for projects involving well know activities of repetitive in nature.

However the distinction between PERT and CPM is mostly historical.

Example:

PERT

The following table lists the jobs of a network with their time estimates.

Job I-j	Duration (days)		
	Optimistic	Most Likely	Pessimistic
1 2	3	6	15
1 6	2	5	14
2 3	6	12	30
2 4	2	5	8
3 5	5	11	17
4 5	3	6	15
6 7	3	9	27
5 8	1	4	7
7 8	4	19	28

(a) Draw the project network.

(b) Calculate the length and variance of the critical path.

(c) What the is approximate probability that the jobs on the critical path will be completed by the due date of 42 days?

(d) What due date has about 90% chance of being met?

Solution:

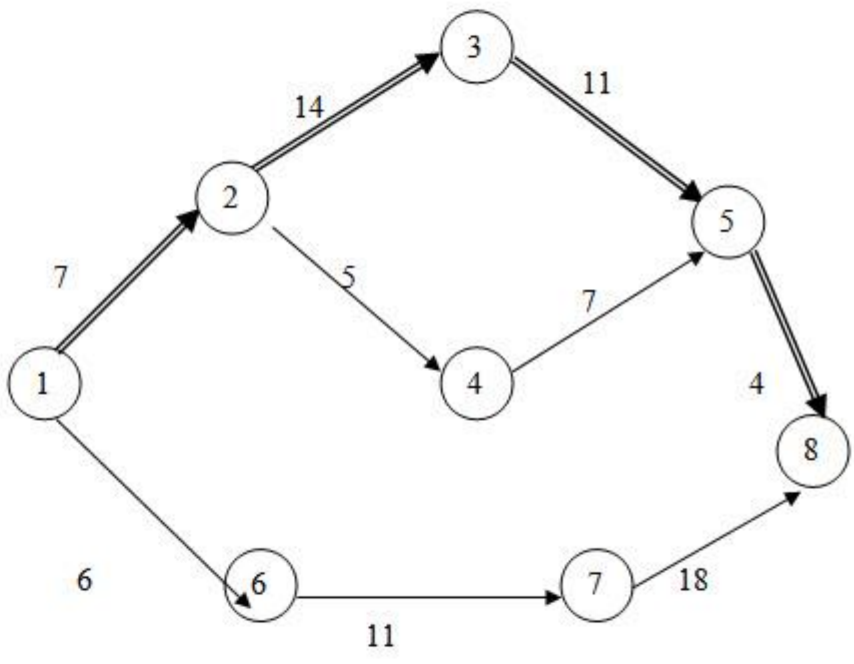
Before proceeding to draw the project network, let us calculate the expected time of activity t_e, standard deviation and variance of the expected time of activity using

\[{t_e}\,\,\,\,\,\,=\,{{\left({{t_0}\,\, + \,\,4t{}_m\, + \,{t_p}} \right)} \over 6}\]

\[S.D\,\,=\,\left({{t_p}\, - \,{t_0}} \right)/6;\,\,\,\,\,\,\,\,\,\,\,\,Variance\,\,\,\,\,=\,\,{(S.D)^2}\]

Activity	Total Slack	Free Slack	Independent Slack
1-2	7	2	4
1-6	6	2	4
2-3	14	4	16
2-4	5	1	1
3-5	11	2	4
4-5	7	2	4
6-7	11	4	16
5-8	4	1	1
7-8	18	4	16

(a) Project Network:



(b) There are three paths:

1-2-3-5-8 = 36 days

1-2-4-5-8 = 23 days

1-6-7-8 = 35 days

1-2-3-5-8 is the longest path and hence the critical path.

Expected length of the critical path is 36 days. The variance for 1-2, 2-3, 3-5 and 5-8 are 4, 16,4 and 1 respectively and variance of the projection duration is 25 and hence

Standard deviation of the project duration = 25 = 5 Days.

(c) Due date = 42 days (T)

Expected duration = 36 days (T_e) and S.D = 5 days (0)

$$Z = (T-T_e) / S.D. = (42-36)/5 = 1.2$$

The area under the normal curve for Z = 1.2 is 0.3849.

Therefore, the probability of completing the project in 42 days

= 0.5000 + 0.3849

= 0.8849

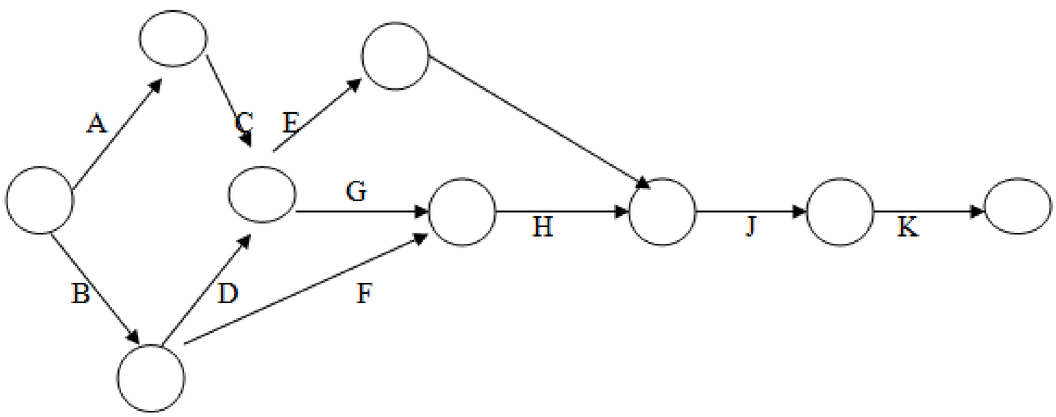
= 88.49%

SOLVED PROBLEMS

Problem 1. Draw a network for the simple project of erection of steel works for a shed. The various elements of project are as under:

Activity code	Description	Prerequisites
A	Erect site workshop	-
B	Fence site	-
C	Bend reinforcement	A
D	Dig foundation	B
E	Fabricate steel work	A, C
F	Install concrete plant	B
G	Place reinforcement	C, D
H	Concrete foundation	G, F
I	Paint steel work	E
J	Erect steel work	H,I
K	Give finishing touch	J

Solution:



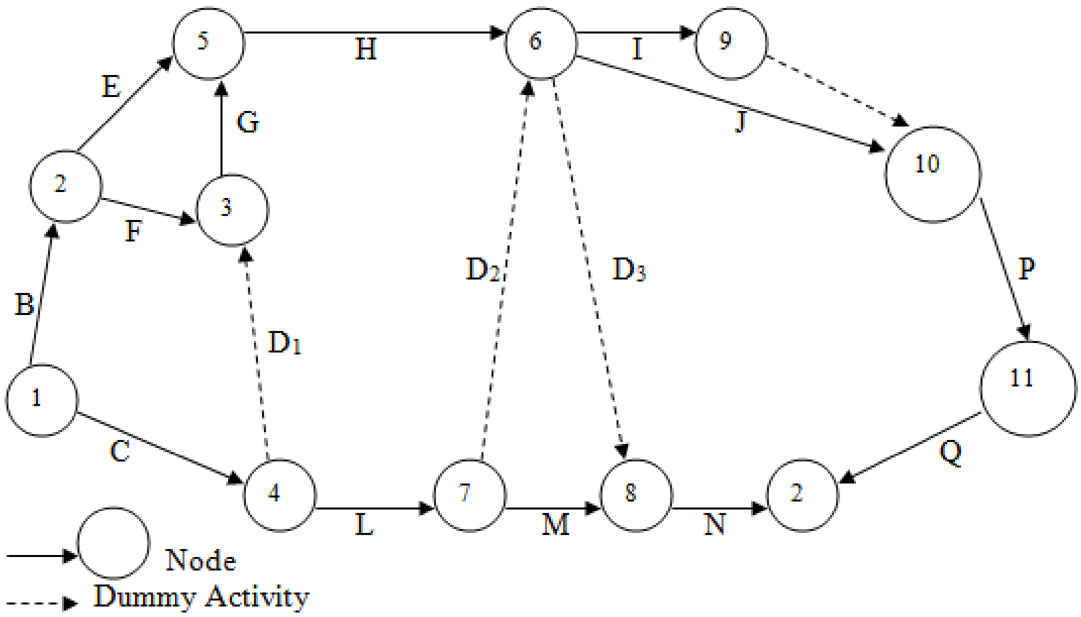
Problem 2. Construct the network diagram comprising activities *B*, *C*, ... *Q* and *N* such that the following constraints are satisfied:

B < E, F; C < G, L; E, G < H ; F, G < H/; L.H < I;

L < M; H, M < N; H < J: I, J < P; P < Q.

The notation X < Y means that the activity X must be finished before Y can begin.

Solution. The resulting network is shown in Fig. *27.6*. The dummy activities D_1, D_2 and D_3 are used to establish the correct precedence relationships. D_4 is used to identify the activities I and J with unique end nodes. The nodes of the project are numbered such that their ascending order indicates the direction of progress in the project:

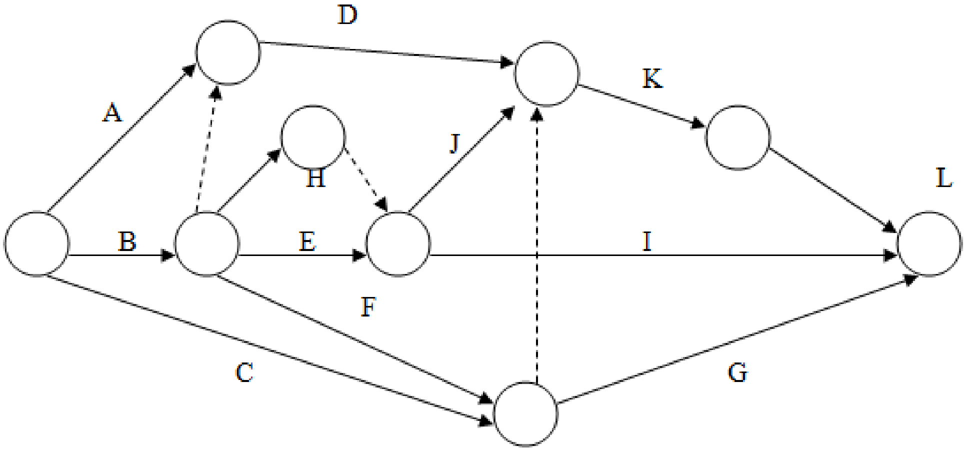


Problem 3. Construct the arrow diagram comprising activities *A, B,* … and *L* such that the following relationships are satisfied:

(i) *A*, *B*, and *C*, the first activities of the project, can start simultaneously,
(ii) *A* and *B* precede *D,* (iii) *B* precedes *E, F* and *H,* (iv) *F* and *C* precede *G,*
(v) *E* and *H* precede *I* and *J,* (vi) *C, D, F* and *J* precede *K,* (vii) *K* precedes *L,*

(viii) *I, G* and *L* are the terminal activities of the project.

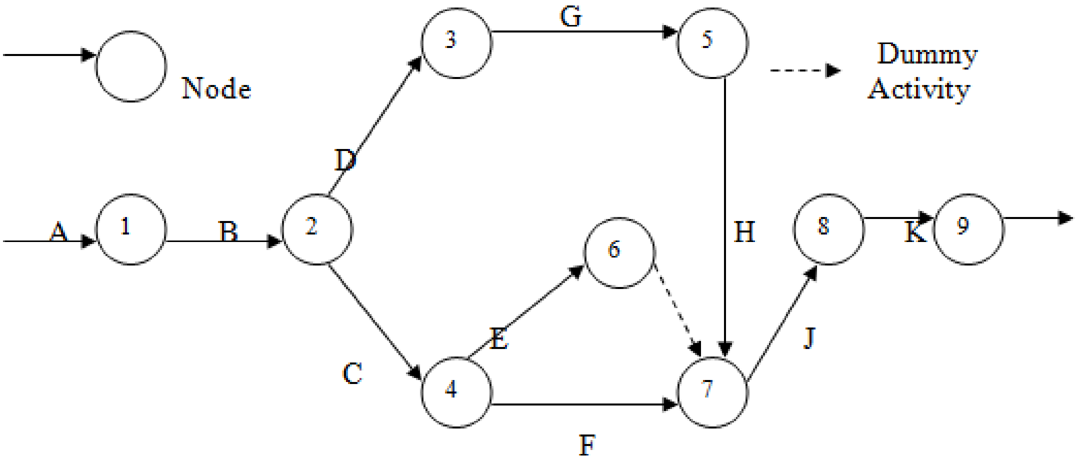
Solution. The resulting network is shown in Fig. *27.7* below. The dummy activities are used to identify the activities *A* and *B; E* and *H; C* and *F* with a unique end node:



Problem 4. Draw a network for the following project and number the events according to Fulkerson's rules:

(1) A is the start event and K is the end event.
(2) J is the successor event to F.
(3) C and D are successor events to B.
(4) D is the preceding event to G.
(5) E and F occur after event C.
(6) E precedes J
(7) C restrains the occurrence of G and G precedes H.
(8) H precedes J.
(9) F restrains the occurrence of H.
(10) K succeeds event J.

Solution. The resulting network is shown in the figure given below. The dummy activity is used to identify the activities E and F with unique end node:



The nodes of the network are numbered to indicate the direction of progress in the network

Problem 5. In a boiler overhauling project following activities are to be performed:

A. Inspection of boiler by boiler engineer and preparation of list of parts to be replaced/repaired.
B. Collecting quotations for the parts to be purchased.
C, Placing the orders and purchasing.
D. Dismantling of the detective parts from the boiler.
E. Preparation of necessary instructions for repairs.

F. Repair of parts in the workshop.
G. Cleaning of the various mountings and fittings.

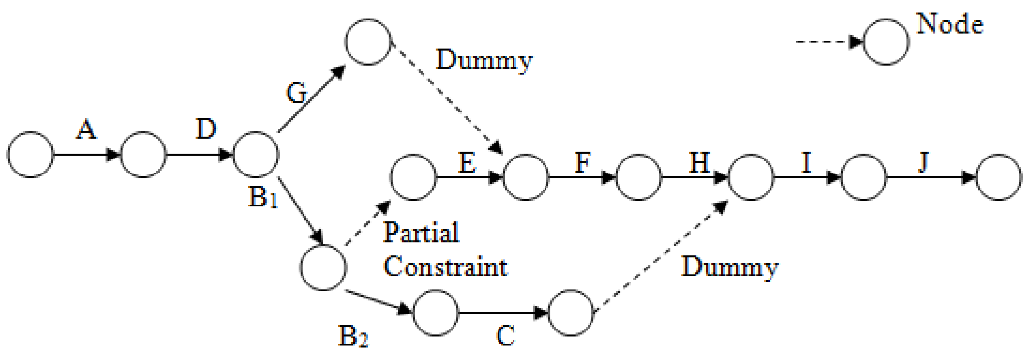
H. Installation of the repaired parts
I. Installation of the purchased parts.

J. Inspection.
K. Trial run.

Assuming that the work is assigned to the boiler engineer who has one boiler mechanic and one boiler attendant at his disposal, draw a network showing the precedence relationships.

Solution. Looking at the list of activities, we note that activity *A* (inspection of boiler) is to be followed by dismantling of defective parts (*D*) and only after that it can be decided which parts can be repaired and which will have to be replaced. Now the repairing and purchasing can go side by side. But the instructions for repairs may be prepared after sending the letters for quotations. Note that it becomes a partial constraint, also started after activity *D*. Now we assume that repairing will take less time than purchasing. But the installation of repaired parts can be started only when the cleaning is completed. This results in the use of a dummy activity. After the installation of repaired parts, installation of purchased parts can be taken up. This will be followed by inspection and trial run.

The network showing the precedence relationships is given below:



The dummy activities are used to identify the activities *C*, *H* and *E*, *G* with unique end nodes.

Problem 6. A television is manufactured in six steps, labelled *A* through *F*. The television is produced one at a time. The manager thinks that network scheduling will improve the position. It is noted that A is the first step and $A < B$. $A < C$, $C < D$, $C < E$, $B > D$, $B < E$, $F > E$, *D* is a successor of *F*.

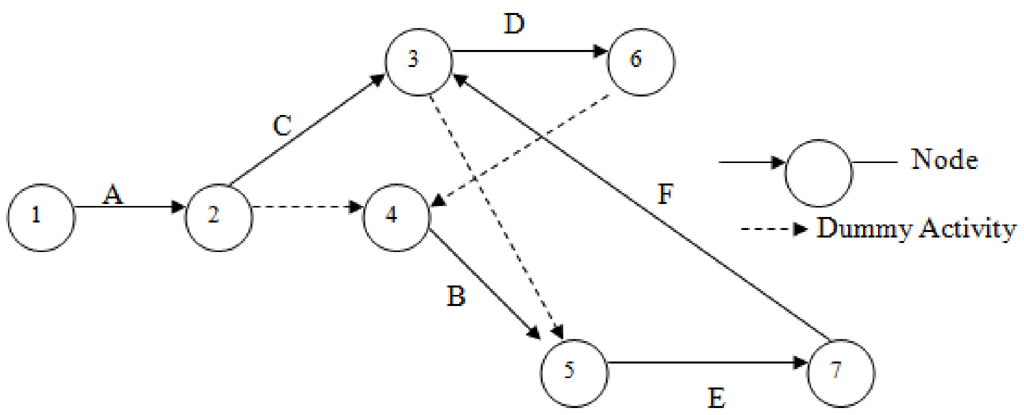
(a) Draw an activity node diagram and identify the cycle.

(b) If *D* precedes *F*, show that the cycle gets eliminated.

(c) If the activities take occupy time as indicated:

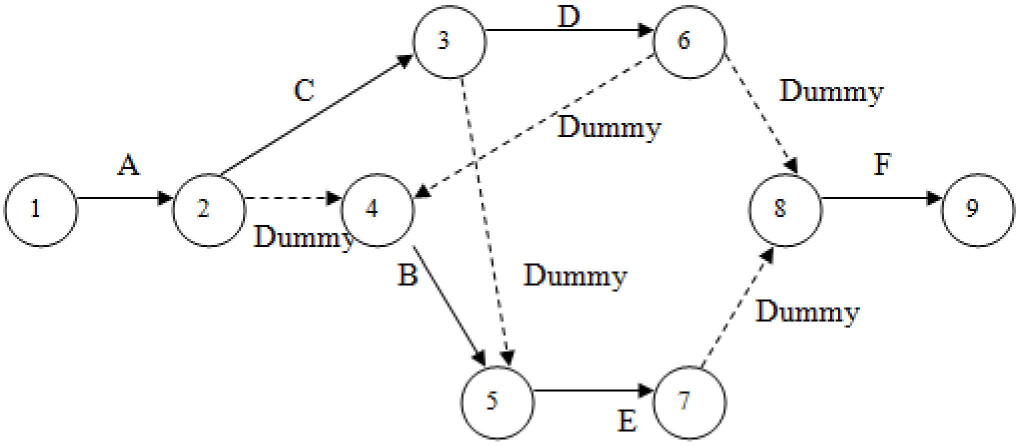
A — 1 hours, *B* — 3 hours, *C* — 2 hours, *D* — 1 hour, *E* — 3 hours and *F* — 2 hours, determine the minimum time required to complete the television set.

Solution, *(a)* Using given constraints, the activity node diagram (network) is shown in Fig. *27.70*. The dummy activities are introduced to establish the correct precedence relationships:



The nodes are numbered in such a way that their ascending order indicates the direction of progress in the manufacturing process. Among the various activities involved, obviously we have the following cycle: $B\rightarrow E\rightarrow F\rightarrow D\rightarrow B$.

(b) If *D* precedes *F*, the revised network is as shown below



c) To determine the minimum time of completion of the project, we compute *ESi* and *LFj* for each activity of the project (TV set). The critical path calculations as applied to Fig. 27.11 are:

$ES_1, = 0$ $\quad ES_2 = ES_1 + t_{12} = 0 + 1 = 1$ $\quad ES_3 = ES_2 + t_{23} = 1+2=3$

$ES_6 = ES_3 + t_{26} = 3 + 1 = 4$ $\quad ES_4 = \underset{i\,=\,2,\,6}{\text{Max}} \{ES_i + t_{i4}\} = \text{Max.} \{1 + 0, 4 + 0\} = 4$

$ES_5 = \underset{i\,=\,3,4}{\text{Max.}} Es_i + t_{i5}\} = \text{Max.} \{3 + 0, 4 + 3\} = 7$ $\quad ES_7 = ES_5 + t_{57} = 7 + 3 = 10$

$ES_8 = \underset{i\,=\,6,\,7}{\text{Max.}} \{ES_i + t_{i8}\} = \text{Max.} \{4 + 0, 10 + 0\} = 10$ $\quad ES_9 = ES_8 + t_{89} =10 + 2 = 12.$

The values of LF_j are similarly computed as follows:

$LF_9 = ES_9 = 12$ $\quad LF_8 = LF_9 – t_{89} = 12 - 2 = 10$
$LF_7 = LF_8 - t_{78} = 10 - 0 = 10$ $\quad LF_5 = LF_7 – t_{57} = 10 - 3 = 7$
$LF_4 = LF_5 - t_{45} = 7 - 3 = 4$ $\quad LF_6 = \text{Min.} \{10 - 0, 4 - 0\} = 4$
$LF_3 = \text{Min.} \{4 - 1, 7 - 0\} = 3$ $\quad LF_2 = \text{Min.} \{3 - 2, 4 - 0\} = 1$
$LF_1 = LF_2 – t_{12} = 1 - 1 = 0.$

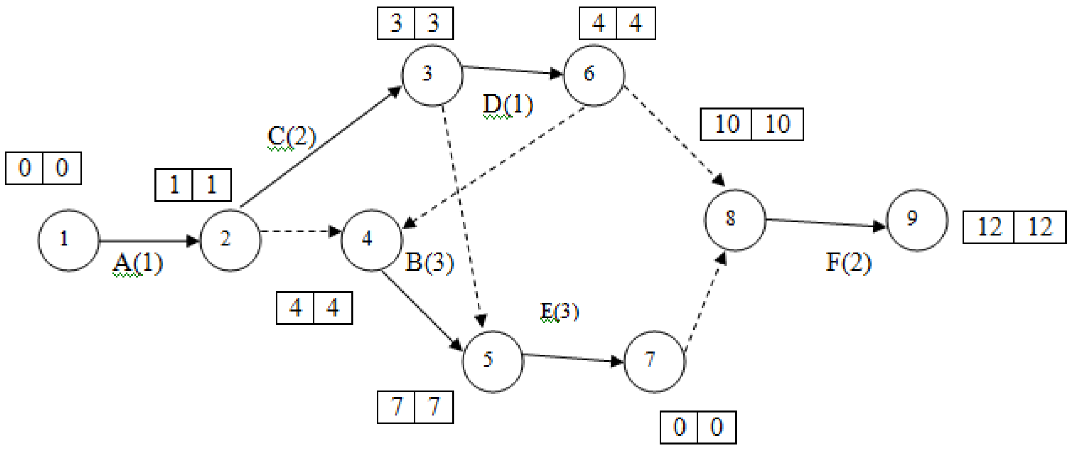
For determining the critical nodes, calculations are displayed in the following table:

Activity	Normal time	Earliest	time	Latest	time	Total float

	(hours)	Start	Finish	Start	Finish	
(1. 2)	1	0	1	0	1	0
(2, 4)	0	1	1	4	4	3
(2,3)	2	I	3	I	3	0
(3,6)	1	3	4	3	4	0
(6, 4)	0	4	4	4	4	0
(4, 5)	3	4	7	4	7	0
(3, 5)	0	3	3	7	7	4
(5,7)	3	7	10	7	10	0
(7, 8)	0	10	10	10	10	0
(6, 8)	0	4	4	10	10	6
(8, 9)	2	10	12	10	12	0

It is apparent from the table that the critical nodes are for the activities (1, 2), (2, 3), (3, 6), (6, 4), (4, 5), {5, 7), (7, 8) and (8, 9).

The critical path, therefore, comprises the activities *A, C, D, B, E* and *F*, which is shown below



The Critical path is 1-2-3-6-4-5-7-8-9

The minimum time required to complete the television set is 1+2+1+3+3+2 = 12 hours.

PROBABILITY OF COMPLETING THE PROJECT WITHIN GIVEN TIME

The Three time estimate for different activities of a project are given below: -

Activity	T_0	T_m	T_p
1-2	2	5	8
2-3	4	9	20
2-4	4	7	16
2-5	8	11	20
3-6	3	7	17
4-6	7	10	13
4-5	0	0	0
5-7	3	5	13
6-7	2	3	10

7-8 2 4 6

Z -2 -1 0 1 2

P % -2.28 15.87 50 84.13 97.72

What is the probability of completing the project in 35 Days?

Step 1 :-

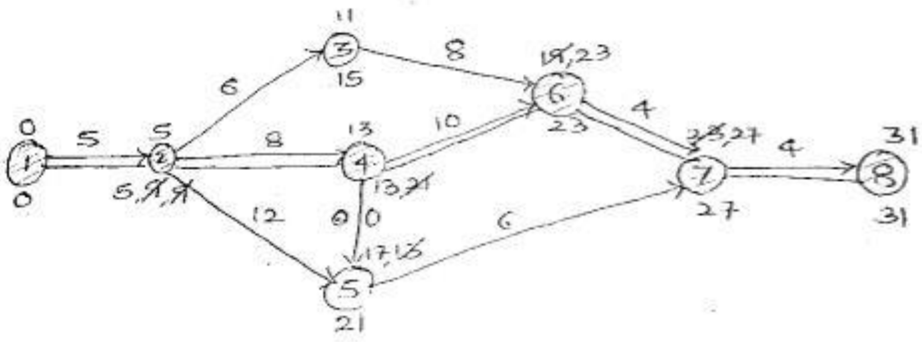
For critical activity only

$\sigma = \frac{t_p - t_o}{6}$

Activity	T_0	T_m	T_p	$t_e = \frac{t_0 + 4t_m + t_p}{6}$	σ	σ^2
1-2	2	5	8	5	1	1
2-3	4	9	20	6		
2-4	4	7	16	8	2	4
2-5	8	11	20	12		
3-6	3	7	17	8		
4-6	7	10	13	10	1	1
4-5	0	0	0	0		
5-7	3	5	13	6		
6-7	2	3	10	4	1.33	1.768
7-8	2	4	6	4	0.66	0.435

$\sum \sigma^2 = 8.203$

Step 2:-



Project duration = 31 days

Critical path =1-2-4-6-7-8

Step 3:-

The probability of completing the project in 35 days is

$z = \frac{\text{scheduled time} - \text{project duration}}{\epsilon}$

S.T.= 35 Days

P.D.= 31 Days

$\epsilon = \sqrt{\sigma^2} = \sqrt{8.203} = 2.86$

$z = \frac{35-31}{2.86} = 1.398$

From table given in question, interpolation

1	84.13
1.398	P%=?
2	97.72

$\therefore P\% = 89.54\%$

PROJECT CRASHING

Project crashing is when you shorten the duration of a project by reducing the time of one or more tasks. Crashing is done by increasing the resources to the project, which helps make tasks take less time than what they were planned for. Of course, this also adds to the cost of the overall project. Therefore, the primary objective of project crashing is to shorten the project while also keeping costs at a minimum.

Just as the triple constraint says, if you reduce the duration of the project, or its time, then costs will in turn have to increase. It's a trade off. Crashing project management accounts for the triple constraint, in that to achieve it, you include additional resources (as noted above) or reduce the project requirements or scope. However, such drastic measures cannot be implemented without the sponsor or primary stakeholders agreeing to the changes.

A result of project crashing can be a change to the critical path and the emergence of a new, different critical path. Project crash management requires that you return to your project schedule to make sure you're aware of changes that have occurred there because of the project crashing.

Different Interpretations of Project Crashing

Project crashing as a term is not etched in stone, and can mean a few different things. It could refer to spending more money to get things done faster. It can also refer to pinpointing the critical path, providing greater resources there, without necessarily thinking about being efficient. Or, you can review the critical path and see if there are any activities that can be shortened by an influx of resources.

A related method for truncating your schedule is called fast tracking. This is when you overlap tasks that were originally scheduled to run separately. But, this course of action should not be taken without first analyzing its feasibility and risk. Whichever route you take, it's always wise to give it thought and analysis.

What Prompts Crashing in Project Management?

When would a project manager want to increase investment to complete the project earlier? After all, a lot of time and effort went into the project planning and schedule. Obviously, since project crashing requires higher costs, it wouldn't be used unless there's an emergency.

One reason for using project crashing would be if the project was scheduled unrealistically, and this wasn't clear until the project has already been executed. This can even happen at the planning stage if the sponsor, customer or stakeholder insist on a due date that isn't feasible.

Another reason is that, during the process of a change control analysis (which shows impact on the time, cost, scope or other project factors), an issue comes up that must be addressed immediately. As issues arise in the course of managing a project that take it off track, the project manager must figure out a way to lock back to the schedule baseline.

As noted above, other than project crashing, there is the fast tracking method. Though we're discussing project crashing, it's important to touch on when fast tracking is preferable. Sometimes you can use either, but if the project is already over-budget and you don't have funds, then fast tracking is the likely option.

Best Practices When Crashing Your Project

Project crashing is usually a last resort, and it's not without substantial risks. There are some things you need to consider before taking your project down this road. For one, are the tasks you're looking to crash in the critical path? These tasks are going to impact the delivery of your project. If the tasks aren't in the critical path, you can probably ignore them.

Another thing to consider is the length of the tasks. A short task will be hard to speed up, especially if it doesn't repeat throughout the project. Long tasks are going to usually have some fat to trim. But regardless of the task, you need to have resources available. If you don't have access to the right resources, then it makes no sense for project crashing. Having to get new materials or team members is likely going to be too costly to be effective.

Related: How to Make a Resource Breakdown Structure (RBS)

Another consideration is if it would take too long to ramp up the project crashing; for instance, if the project involves very specific skills and on boarding new team members would be costly and time consuming. While it might seem logical to crash at the end of the project when it is becoming clear you're not going to hit your target, most experts suggest avoiding that scenario.

Project crashing is most effective earlier in the timeline—usually when a project is less than halfway done.

Project Crashing Management Stages

Once you've made the decision to use project crashing, there are some steps you'll want to follow to get the results you want.

1. Critical Path

The first thing to do is analyze the critical path of your project. This will help you determine which tasks can be shortened to bring the project to a close sooner. Therefore, if you haven't already, calculate your critical path, see which tasks are essential and which are secondary to the project's success.

2. Identify Tasks

Get a list of all the tasks you have, then meet with those who have been assigned to complete them. Ask if they believe any of the tasks they're responsible for are in the critical path and can be cut down. Then, start looking for ways to tighten up those tasks.

3. What's the Trade Off?

Once you've narrowed down the tasks in the critical path that you believe can be shortened, start calculating how much adding more resources will cost. Find the tasks that can be allocated additional resources, and come in sooner with the least amount of strain on your budget.

4. Make Your Choice

When you know what you will have to spend (compared to how much time you'll save) for each of the tasks in your critical path, you must now make a decision and choose the least expensive way forward. Project crashing is not just adding resources to get done faster, but it's getting the most in return for that extra expense.

5. Create a Budget

Like any project, once you've decided on your plan, you have to pay for it. Making a project crashing budget is the next step in executing your project crashing plan. You'll have to update your baseline, schedule and resource plan to align with your new initiative.

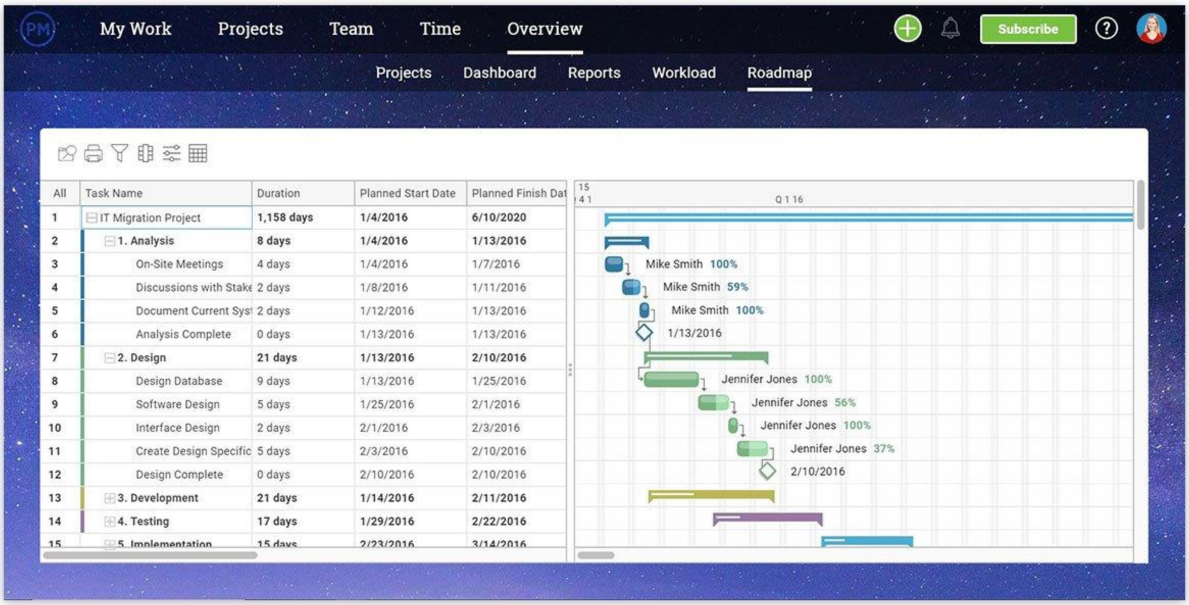
How ProjectManager.com Makes Project Crashing Easier

Project crashing involves knowing your resources and then reallocating them. If time is working against you, then this process needs to be as efficient as possible. ProjectManager.com is an award-winning software that tracks your resources, teams and projects to boost productivity.

Our resource management software is cloud-based, meaning that the data you see is the most current picture of your project resources possible. You also get visibility into what your team is doing as they're doing it, including costs. This allows you to make better decisions.

Track Resource Costs

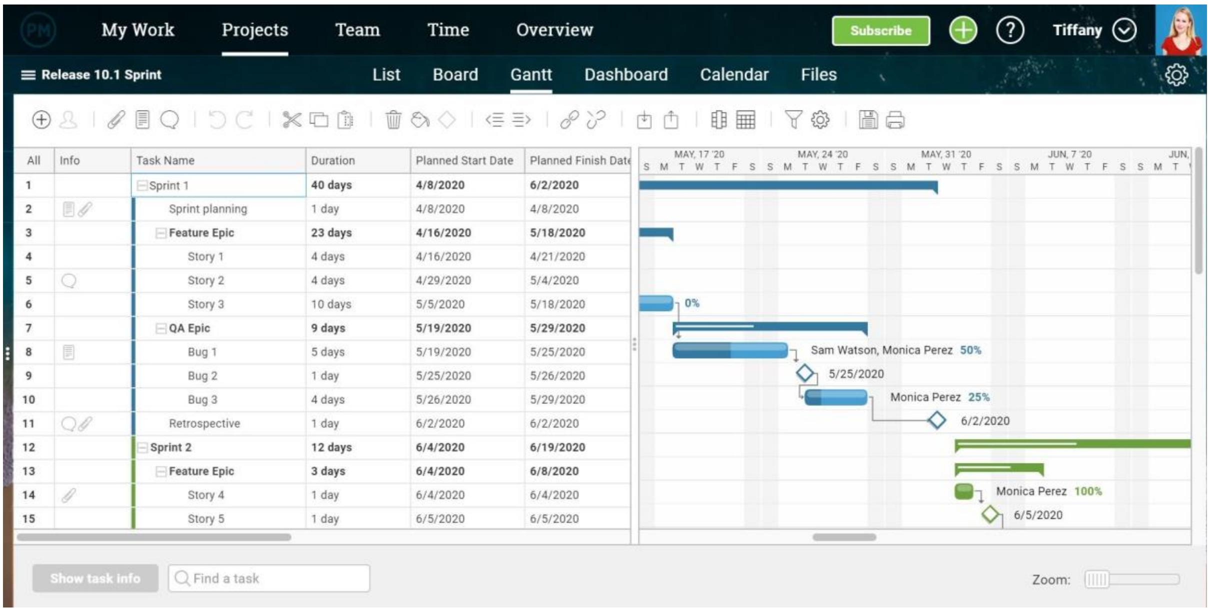
All your resources can be identified, from teams to supplies, equipment and more. When you add the hourly rates to the project, you can see them across all our software features. Once you team logs their hours, we calculate the actual costs for you. Then, they can be compared to the planned costs, so you know immediately if you're on target.



Our online roadmap software helps you identify and track your resource costs.

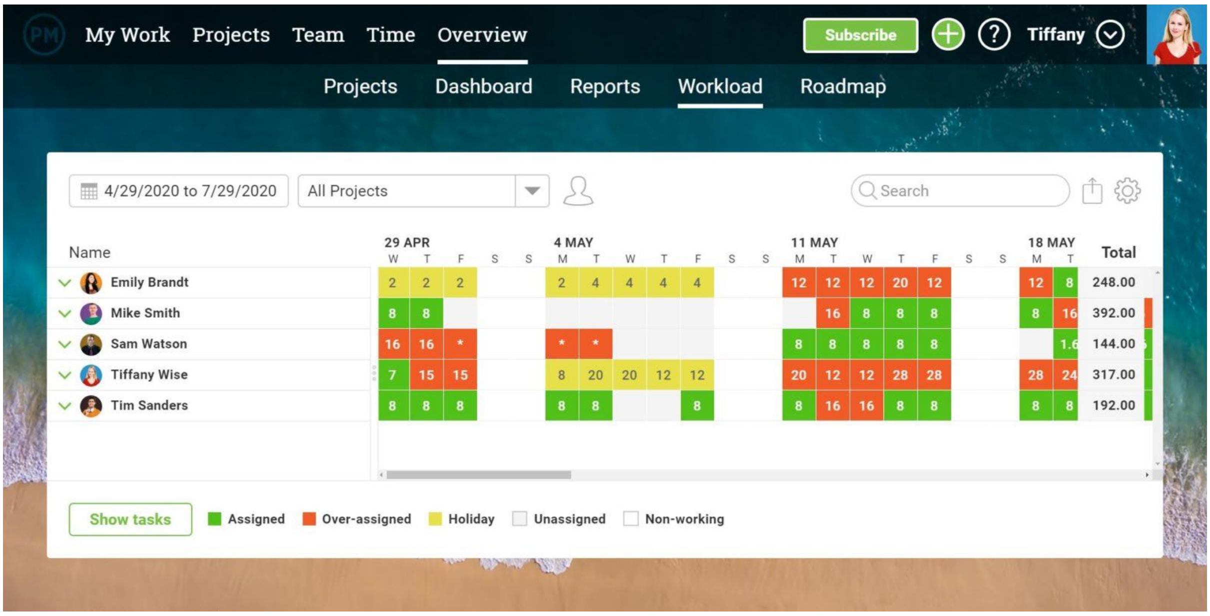
Use the Gantt to Schedule Resources & Tasks

Use our online Gantt chart tool to schedule your resources. Assignments can be made while in this view. A popup window will also tell you how many hours your team member is working, and if they have too many or too few tasks assigned to them. The Gantt will also filter the critical path, estimate planned costs for your resources and set a baseline. You can even track progress on the Gantt, or by using our real-time dashboard.



Balance Your Team’s Workload for Efficient Project Crashing

To make sure your project crashing management plan isn’t overloading some of your team while others are under-allocated, use the workload page. On the color-coded chart, you can easily see who has too many tasks, balance the workload or assign them to the project crashing.



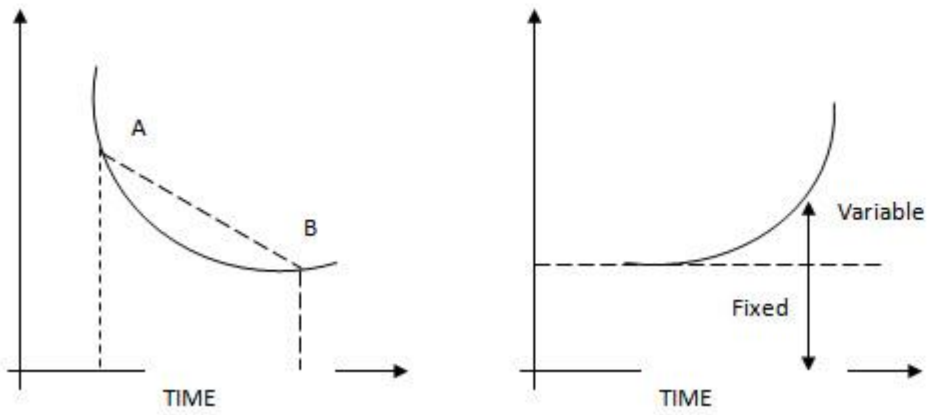
Color-coded workload chart makes it easy to see who has too many tasks and who has too few.

PROJECT COST

In order to include the cost aspects in project scheduling, we must first define the cost-duration relationships for various activities in the project. The total project cost comprises direct and indirect costs. The direct costs are associated with the individual activities such as manpower loading, equipment utilized, materials consumed directly, etc., in respect of various activities. The indirect costs are those expenditures which cannot be allocated to individual activities of the project. These may include administration or supervision costs. loss of revenue, fixed overheads, depreciation, insurance, and so on. While indirect cost allocated to a project goes up with the increase in project duration, direct costs go high as the time for individual activity is reduced. Such deliberate reduction of activity times by putting an extra effort is called crashing the activity.

It may be noted that for technical reasons, the duration of an activity cannot be reduced indefinitely. The crash time represents the fully expedited or the minimum activity duration time that is possible, and any attempts to further 'crash' would only raise the activity direct costs without reducing the time. The activity cost corresponding to the crash time is called the crash cost which equals the minimum direct cost required to achieve the crash performance time. These are in contrast to the normal time and the normal cost of the activity. The normal cost is equal to the absolute minimum of the direct costs required to perform an activity. The corresponding activity duration is known as the normal time.

The direct cost-time relationship and indirect cost-time relationships are



The point A in denotes the normal time for completion of an activity whereas point B denotes the crash time which indicates the least duration in which the particular activity can be completed. The cost curve is non-linear and asymptotic but for the sake of simplicity it can be approximated by a straight line with its slope given by

$$Cost\ slope = \frac{Crash\ \text{cost} - Normal\ \text{cost}}{Normal\ time - Crash\ time}$$

The cost slope represents the rate of increase in the cost of performing the activity per unit decrease in time and is called cost/time trade off. It varies from activity to activity. Having assessed the direct and indirect project cost the total costs can be found out. The total project cost is the sum total of the project direct and indirect costs.

Shows both the direct and the indirect project cost. As these two curves have been plotted against the same time scale, at each ordinate, the project direct and indirect costs can be added to obtain the various points on the graph, indicating the total project cost corresponding to the various project durations. If these total time-cost points are joined to form a curve, the curve so obtained will be total project cost curve. If this curve is examined carefully, it may be seen that for different project durations, there will be corresponding project costs. However, compared to all other points on this curve, there will be one point which is at the lowest, indicating the least cost and the minimum time for the project. It may again be observed that any point on the right hand side of this lowest point, the cost of the project increases and project duration also increases. On the other hand, if the point~ on the left hand side of the lowest point of this curve is considered, it may be seen that though there is reduction in project duration time, the cost of the project, however, increases. So much so, any point either on the right hand side or left hand side of the lowest point of this total projects. cost curve, it is not advantageous to the project manager to choose any of the points for the implementation of the project Obviously, it is the lowest point which indicates the lowest overall cost against the minimum time required for the completion of the project. This point may therefore be called as the optimum time-cost point of the curve which the project manager can choose for the implementation of the project

RESOURCE LEVELLING

In PERT and CPM techniques there is an implied assumption that required resources are always available. When resources are limited, two alternative courses of action are available. In the first alternative, the activities are critically sequenced and the minimum period of project is redetermined. This process is called Resource Levelling. The problem here is to manipulate the activity slacks, schedules and resource requirements throughout the duration of the project. In resource levelling two types of problems are involved:

1. Levelling resource demands with constraint on the total project duration time.

2. Minimization of the project duration time with a constraint on the total availability of certain key resources.

The first problem arises when resources are adequate but they are desired to be used at a relatively constant rate during the life of the project. The second problem occurs when the resources cannot be increased and the object is to minimize project duration with available resources.

Thus, resource levelling or load levelling is required when the demands on specified resources are required not to exclude the specified level and the duration of the project is not invariant.

Remark. In order to stabilize the use of existing level of resources the total float of non-critical activities is used. By shifting a non-critical activity between its earliest start time and latest allowable time, project manager may be able to lower the maximum resource requirements. The following two general rules are normally used in scheduling non-critical activities.

1. If the total float of a non-critical activity is equal to its free float, then it can he schedule any where between its earliest start and latest completion times.
2. If the total float is a non-critical activity is more than its free float, then its starting time can be delayed relative to its earliest start time by no more than the amount of its free float without affecting the scheduling of its immediately succeeding activities.

OPTIMUM COST AND OPTIMUM DURATION.

ALGORITHM

The process of shortening a project is called crashing and is usually achieved by adding extra resources to an activity. Project crashing involves following steps:

Step 1: Critical path. Find the normal critical path and identify the critical activities.

Step 2: Cost slope. Calculate the cost slope for the different activities by using the formula:

$$Cost\ slope = \frac{Crash\ \text{cost} - Normal\ \text{cost}}{Normal\ time - Crash\ time}$$

Step 3: Ranking. Rank the activities in the ascending order of cost slope.

Step 4: Crashing. Crash the activities in the critical path as per the ranking, i.e., activity having lower cost slope would be crashed first to the maximum extent possible. Calculate the new direct cost by cumulative adding the cost of crashing to the normal cost.

Step 5: Parallel crashing. As the critical path duration is reduced by the crashing in Step 3, other paths also become critical, i.e., we get parallel critical paths. This means that project duration can be reduced duly by simultaneous crashing of activities in the parallel critical paths.

Step 6: Optimal duration. Crashing as per step 3 and step 4, optimal project duration is determined. It would be the time duration corresponding to which the total cost (i.e., direct cost plus indirect cost) is a minimum.

SAMPLE PROBLEMS

The following table gives the activities in a construction project and other, relevant information:

Activity i -- j	Normal time (days)	Crash time (days)	Normal cost (Rs.)	Crash cost (Rs.)
1-2	20	17	600	720
1-3	25	25	200	200
2-3	10	8	300	440
2-4	12	6	400	700
3-4	5	2	300	420
4-5	10	5	300	600
4-6	5	3	600	900
5-7	10	5	500	800
6-7	8	3	400	700

(a) Draw the activity network of the project

(b) Find the total float and free float for each activity

(c) Using the above information crash the activity step by step until all paths are critical.

Solution: (a) using the normal time duration the network is given below:

(b) Considering the normal time of the project, the earliest times and latest times as well as the total floats and free floats in respect of the node points is obtained in the following table:

Activity (i-j)	Normal duration (days)	Earliest time		Latest time		Float	
		Start	Finish	Start	Finish	Total	Free
1-2	20	0	20	0	20	0	0
1-3	25	0	25	5	30	5	5
2-3	10	20	30	20	30	0	0
2-4	12	20	32	23	35	3	3
3-4	5	30	35	30	35	0	0
4-5	10	35	45	35	45	0	0
4-6	5	35	40	42	47	7	7
5-7	10	45	55	45	55	0	0
6-7	8	40	48	47	55	7	7

The critical path of project is

$1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 7$

and duration of the project is 55 days with total cost as Rs. 3600

(c)The cost slopes of the activities of the above network are computed as follows:

Activity	Cost slope
1-2	(720 – 600)/(20 – 17) = 40
1-3	(200 – 200)/(25 – 25) =0
2-3	(440 – 300)/(10 – 8) = 70
2-4	(700 – 400)/(12 - 6) = 50
3-4	(420 – 300)/(5 – 2) = 40
4-5	(600 – 300)/(10 – 5) = 60
4-6	(900 – 600)/(5 – 3) = 150
5-7	(800 – 500)/(10 – 5) = 60
6-7	(700 – 400)/(8 – 3) =60

Crashing of Activities

Since the activities lying on the critical path control the project duration, we crash the activities lying on the critical path.

First crashing. First of all we crash that activity of critical path which involves the minimum cost slope. Since the activities (1, 2) and (3, 4) give the minimum cost slope, we compress the duration of activity (3, 4) from 5 to 2 days with an additional cost Rs. 3 x 40, i.e., Rs. 120.

Thus the revised network is

Duration of project is now 52 days and total cost

= Rs. 3,600 + Rs. 40 x 3 = Rs. 3,720.

Second crashing. Now, Since there are two parallel critical paths, we choose the minimum cost slope of the activity which lies on any of the two critical paths. As the minimum cost slope is for the activity (1, 2) we compress this activity from 20 to 17 days with an additional cost of Rs. 40 x 3 i.e. Rs. 120.

Thus we have the following network:

Duration of project is now 49 days and total cost

= Rs. 3,720 + Rs. 40 x 3 = Rs. 3,840.

Third crashing. Minimum cost slope now is for the activities (4, 5), (5, 7), and (6, 7). Therefore crashing the activities (4, 5) and (5, 7) by 5 days each and activity (6, 7) by 3 days at an extra cost of Rs. 60 per day, we have

Duration of project is now 39 days and total cost

= Rs. 3,840 + Rs. 60 x (5 + 5+ 3) = Rs. 4,620.

Fourth crashing. Finally, crashing the activities (2, 3) and (2, 4) by two days each, we find that all the activities are lying on the critical path. Thus, we have

Duration of project is now 39 days and total cost

= Rs. 4,620 + Rs. 70 x 2 + Rs. 50 x 2 = Rs. 4,860.

Since all the activities are now on critical paths, the process of crashing is completed.

REPLACEMENT MODELS COMPRISING SINGLE REPLACEMENT AND GROUP REPLACEMENT.

When a PR action is performed, it incurs a **replacement** cost and a downtime cost. Hence, when the **replacement** cost is high, it might be worthwhile **replacing** both machines at the same time (called **group replacement** policy; GRP) instead of **replacing** them separately (called **individual replacement** policy; IRP).

TYPES OF REPLACEMENT PROBLEM

Replacement study can be classified into two categories: (a) Replacement of assets that deteriorate with time (Replacement due to gradual failure, or wear and tear of the components of the machines). This can be further classified into the following types: (i) Determination of economic life of an asset. (ii) Replacement of an existing asset with a new asset. (b) Simple probabilistic model for assets which fail completely (replacement due to sudden failure).

Types of Replacement Problem

Replacement study can be classified into two categories:

(a) Replacement of assets that deteriorate with time (Replacement due to gradual failure, or wear and tear of the components of the machines).

This can be further classified into the following types:

(i) Determination of economic life of an asset.

(ii) Replacement of an existing asset with a new asset.

(b) Simple probabilistic model for assets which fail completely (replacement due to sudden failure).

Determination of Economic Life of an Asset

Any asset will have the following cost components:

ü Capital recovery cost (average first cost), computed from the first cost (purchase price) of the machine.

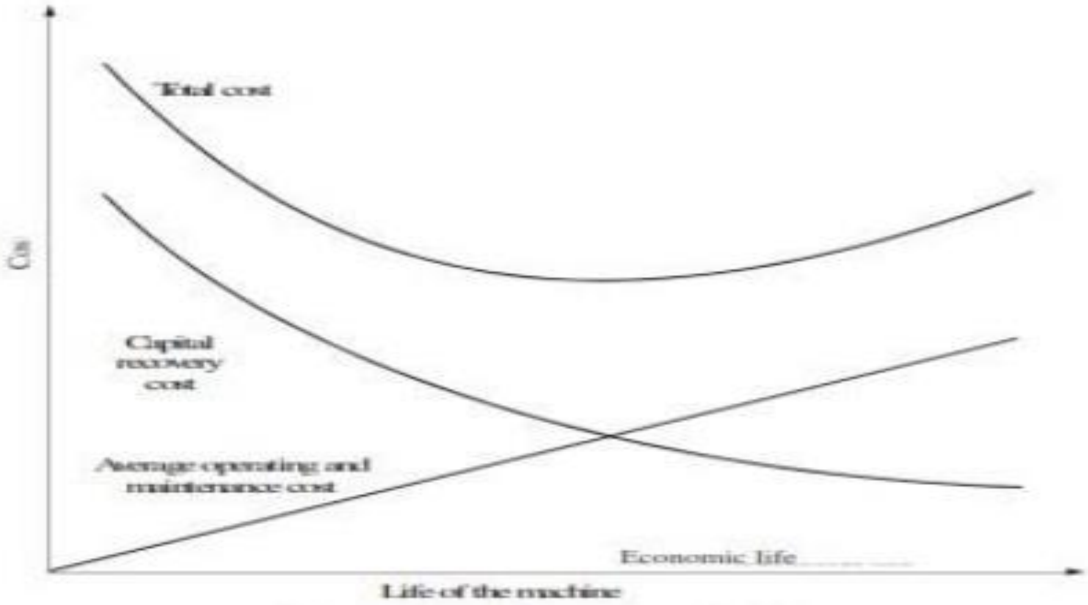


Fig. Chart showing economic life.

ü Average operating and maintenance cost (O & M cost)

Total cost which is the sum of capital recovery cost (average first cost) and average maintenance cost.

EXAMPLE

A firm is considering replacement of an equipment, whose first cost is Rs. 4,000 and the scrap value is negligible at the end of any year. Based on experience, it was found that the maintenance cost is zero during the first year and it increases by Rs. 200 every year thereafter.

(a) When should the equipment be replaced if i = 0%?

(b) When should the equipment be replaced if i = 12%?

(a) When i = 0%. In this problem

(i) First cost = Rs. 4,000

(ii) Maintenance cost is Rs. 0 during the first year and it increases by Rs. 200 every year thereafter.

This is summarized in column B of Table

End of year (n)	Maintenance cost at end of year	Summation of maintenance costs	Average cost of maintenance through year given	Average first cost if replaced at year end given	Average total cost through year given
		B	C/A	4,000/A	D + E
A	B (Rs.)	C (Rs.)	D (Rs.)	E (Rs.)	F (Rs.)
1	0	0	0	4,000.00	4,000.00
2	200	200	100	2,000.00	2,100.00
3	400	600	200	1,333.33	1,533.33
4	600	1,200	300	1,000.00	1,300.00
5	800	2,000	400	800.00	1,200.00
6	1,000	3,000	500	666.67	1,166.67*
7	1,200	4,200	600	571.43	1,171.43

*Economic life of the machine = 6 years

Column C summarizes the summation of maintenance costs for each replacement period. The value corresponding to any end of year in this column represents the total maintenance cost of using the equipment till the end of that particular year.

Average total cost = [First cost (FC) + Summation of maintenance cost] / Replacement period

$$\text{Average total cost} = \frac{\text{First cost (FC)} + \text{Summation of maintenance cost}}{\text{Replacement period}}$$

$$= \frac{\text{Average first cost for he given period}}{n} + \frac{\text{Average maintenance cost for the given period}}{n}$$

Column F = Column E + Column D

Column F = Column E + Column D

The value corresponding to any end of year (n) in Column F represents the average total cost of using the equipment till the end of that particular year.

For this problem, the average total cost decreases till the end of year 6 and then it increases. Therefore, the optimal replacement period is six years, i.e. economic life of the equipment is six years.

(b) When interest rate, i = 12%. When the interest rate is more than 0%, the steps to be taken for getting the economic life are summarized with reference to Table

End of year (*n*)	Maintenance cost at end of year	P/F, 12%, *n*	Present worth as of beginning of year 1 of maintenance costs	Summation of present worth of maintenance costs through year given	Present worth of cumulative maintenance cost & first cost	A/P, 12%, *n*	Annual equivalent total cost through year given
			(B C)	D	E + Rs. 4,000		F G
A	B (Rs)	C	D (Rs.)	E (Rs.)	F (Rs.)	G	H (Rs.)
1	0	0.8929	0.00	0.00	4,000.00	1.1200	4,480.00
2	200	0.7972	159.44	159.44	4,159.44	0.5917	2,461.14
3	400	0.7118	284.72	444.16	4,444.16	0.4163	1,850.10
4	600	0.6355	381.30	825.46	4,825.46	0.3292	1,588.54
5	800	0.5674	453.92	1,279.38	5,279.38	0.2774	1,464.50
6	1,000	0.5066	506.60	1,785.98	5,785.98	0.2432	1,407.15
7	1,200	0.4524	542.88	2,328.86	6,328.86	0.2191	1,386.65*
8	1,400	0.4039	565.46	2,894.32	6,894.32	0.2013	1,387.83
9	1,600	0.3606	576.96	3,471.28	7,471.28	0.1877	1,402.36
10	1,800	0.3220	579.60	4,050.88	8,050.88	0.1770	1,425.00

*Economic life of the machine = 7 years

Table Calculations to Determine Economic Life (First cost = Rs. 4,000, Interest = 12%)

The steps are summarized now:

1. Discount the maintenance costs to the beginning of year 1.

Column D = Column B
1/ (1 + i)n

= Column B (P/F, i, n) = Column B Column C.

2. Find the summation of present worth of maintenance costs through the year given (Column E = Column D).

3. Find Column F by adding the first cost of Rs. 4,000 to Column E.

4. Find the annual equivalent total cost through the years given.

Column H = Column F i(1 + i)n / (1 + i)n − 1

= Column F (A/P, 12%, n) = Column F Column G

5. Identify the end of year for which the annual equivalent total cost is minimum.

For this problem, the annual equivalent total cost is minimum at the end of year 7. Therefore, the economic life of the equipment is seven years.

Replacement Of Exist Ing Asset With A New Asset

In this section, the c oncept of comparison of replacement of an existing asset with a new asset is presented. In this analysis, the annual equivalent cost of each alternative should be co mputed first.

Then the alternative which has the least cost should be selected as the best alternative. Before discu ssing details, some preliminary concepts whic h are essential for this type of replacem ent analysis are presented.

Capital Recovery with Return

Consider the follow ing data of a machine.

Let

P = purchase priice of the machine,

F = salvage valu e of the machine at the end of machine life,

n = life of the m achine in years, and

i = interest rate,, compounded annually

The corresponding cash flow diagram is shown in Fig

Fig. Cash flow diagram of machine.

The equation for the annual equivalent amount for the above cash flow diagram is

$$AE(i) = (P - F) \quad (A/P, i, n) + F \quad i$$

This equation represents the *capital recovery with return*.

Concept of Challenger and Defender

- If an existing equipment is considered for replacement with a new equipment, then the existing equipment is known as the *defender* and the new equipment is known as *challenger*.

- Assume that an equipment has been purchased about three years back for Rs. 5,00,000 and it is considered for replacement with a new equipment. The supplier of the new equipment will take the old one for some money, say, Rs. 3,00,000.

- This should be treated as the present value of the existing equipment and it should be considered for all further economic analysis.

- The purchase value of the existing equipment before three years is now known as *sunk cost*, and it should not be considered for further analysis.

EXAMPLE

Two years ago, a machine was purchased at a cost of Rs. 2,00,000 to be useful for eight years. Its salvage value at the end of its life is Rs. 25,000. The annual maintenance cost is Rs. 25,000.

The market value of the present machine is Rs. 1,20,000. Now, a new machine to cater to the need of the present machine is available at Rs. 1,50,000 to be useful for six years. Its annual maintenance cost is Rs. 14,000. The salvage value of the new machine is Rs. 20,000.

Using an interest rate of 12%, find whether it is worth replacing the present machine with the new machine.

Solution

Alternative 1—

Present machine

Purchase price = Rs. 2,00,000

Present value (P) = Rs. 1,20,000

Salvage value (F) = Rs. 25,000

Annual maintenance cost (A) = Rs. 25,000

Remaining life = 6 years

Interest rate = 12%

The cash flow diagram of the present machine is illustrated in Fig.

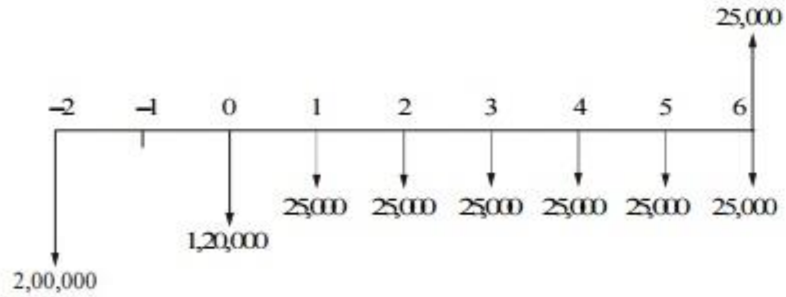


Fig. Cash flow diagram for alternative 1.

Fig. Cash flow diagram for alternative 1.

annual maintenance cost for the preceding periods are not shown in this figure. The annual equivalent cost is computed as

$$AE(12\%) = (P – F)(A/P, 12\%, 6) + F \quad i + A$$

$$= (1,20,000 – 25,000)(0.2432) + 25,000 \quad 0.12 + 25,000$$

$$= \text{Rs. } 51,104$$

Alternative 2 —

New machine

Purchase price (P) = Rs. 1,50,000

Salvage value (F) = Rs. 20,000

Annual maintenance cost (A) = Rs. 14,000

Life = 6 years

Interest rate = 12%

The cash flow diagram of the new machine is depicted in Fig.

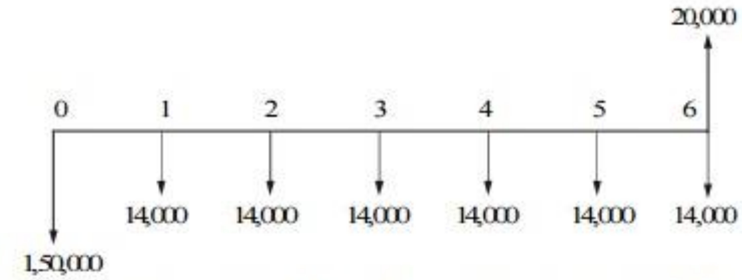


Fig. Cash flow diagram for alternative 2.

Fig. Cash flow diagram for alternative 2.

The formula for the annual equivalent cost is

$$AE(12\%) = (P - F)(A/P, 12\%, 6) + F \quad i + A$$

$$= (1{,}50{,}000 - 20{,}000)(0.2432) + 20{,}000 \quad 0.12 + 14{,}000$$

$$= \text{Rs. } 48{,}016$$

Since the annual equivalent cost of the new machine is less than that of the present machine, it is suggested that the present machine be replaced with the new machine.

Simple Probabilistic Model For Items Which Fail Completely

Electronic items like transistors, resistors, tubelights, bulbs, etc. could fail all of a sudden, instead of gradual deterioration. The failure of the item may result in complete breakdown of the system. The system may contain a collection of such items or just one item, say a tubelight.

Therefore, we use some replacement policy for such items which would avoid the possibility of a complete breakdown.

The following are the replacement policies which are applicable for this situation.

(i) Individual replacement policy. Under this policy, an item is replaced immediately after its failure.

(ii) Group replacement policy. Under this policy, the following decision is made:

At what equal intervals are all the items to be replaced simultaneously with a provision to replace the items individually which fail during a fixed group replacement period?

There is a trade-off between the individual replacement policy and the group replacement policy. Hence, for a given problem, each of the replacement policies is evaluated and the most economical policy is selected for implementation. This is explained with two numerical problems.

EXAMPLE

The failure rates of transistors in a computer are summarized in Table .

Table Failure Rates of Transistors in Computers

End of week	1	2	3	4	5	6	7
Probability of failure to date	0.07	0.18	0.30	0.48	0.69	0.89	1.00

The cost of replacing an individual failed transistor is Rs. 9. If all the transistors are replaced simultaneously, it would cost Rs. 3.00 per transistor. Any one of the following two options can be followed to replace the transistors:

(a) Replace the transistors individually when they fail (individual replacement policy).

(b) Replace all the transistors simultaneously at fixed intervals and replace the individual transistors as they fail in service during the fixed interval (group replacement policy).

Find out the optimal replacement policy, i.e. individual replacement policy or group replacement policy. If group replacement policy is optimal, then find at what equal intervals should all the transistors be replaced.

Solution

Assume that there are 100 transistors in use.

Let,

pi be the probability that a transistor which was new when placed in position for use, fails during the *i*th week of its life. Hence,

$$p_1 = 0.07, \quad p_2 = 0.11, \quad p_3 = 0.12, \quad p_4 = 0.18,$$
$$p_5 = 0.21, \quad p_6 = 0.20, \quad p_7 = 0.11$$

Since the sum of *pis* is equal to 1 at the end of the 7th week, the transistors are sure to fail during the seventh week.

Assume that

(a) transistors that fail during a week are replaced just before the end of the week, and

(b) the actual percentage of failures during a week for a sub-group of transistors with the same age is same as the expected percentage of failures during the week for that sub-group of transistors.

Let

Ni = the number of transistors replaced at the end of the *i*th week

*N*0 = number of transistors replaced at the end of the week 0 (or at the beginning of the first week).

= 100

N1 = number of transistors replaced at the end of the 1st week

= N0 p1 = 100 0.07 = 7
N2 = number of transistors replaced at the end of the 2nd week
= N0 p2 + N1 p1
= 100 0.11 + 7 0.07 = 12
N3 = N0 p3 + N1 p2 + N2 p1
= 100 0.12 + 7 0.11 + 12 0.07
= 14
N4 = N0 p4 + N1 p3 + N2 p2 + N3 p1
= 100 0.18 + 7 0.12 + 12 0.11 + 14 0.07
= 21
N5 = N0 p5 + N1 p4 + N2 p3 + N3 p2 + N4 p1
= 100 0.21 + 7 0.18 + 12 0.12 + 14 0.11 + 21 0.07
= 27

N6 = N0 p6 + N1 p5 + N2 p4 + N3 p3 + N4 p2 + N5 p1
= 100 0.2 + 7 0.21 + 12 0.18 + 14 0.12 + 21 0.11 + 27 0.07
= 30
N7 = N0 p7 + N1 p6 + N2 p5 + N3 p4 + N4 p3 + N5 p2
+ N6 p1
= 100 0.11 + 7 0.2 + 12 0.21 + 14 0.18 + 21 0.12
+ 27 0.11 + 30 0.07
= 25

Calculation of individual replacement cost

Expected life of each transistor = $= \sum_{i=1}^{7} i \quad p_i$

= 1 0.07 + 2 0.11 + 3 0.12 + 4 0.18

+ 5 0.21 + 6 0.2 + 7 0.11

= 4.39 weeks

Average No. of failures/week = 100/4.39 = 23

(approx.) Therefore,

Cost of individual replacement

= (No. of failures/week Individual replacement cost/transistor) = 23 9 = Rs. 207.

Determination of group replacement cost

Cost of transistor when replaced simultaneously = Rs. 3

Cost of transistor when replaced individually = Rs. 9

The costs of group replacement policy for several replacement periods are summarized in Table.

Table Calculations of Cost for Preventive Maintenance

Table Calculations of Cost for Preventive Maintenance

End of week A	*Cost of replacing 100 transistors at a time* B (Rs.)	*Cost of replacing transistors individually during given replacement period* C (Rs.)	*Total cost* (B + C) D (Rs.)	*Average cost/week* (D/A) E (Rs.)
1	300	7 9 = 63	363	363.00
2	300	(7 + 12) 9 = 171	471	235.50
3	300	(7 + 12 + 14) 9 = 297	597	199.00
4	300	(7 + 12 + 14 + 21) 9 = 486	786	196.50*
5	300	(7 + 12 + 14 + 21 + 27) 9 = 729	1,029	205.80
6	300	(7 + 12 + 14 + 21 + 27 + 30) 9 = 999	1,299	216.50
7	300	(7 + 12 + 14 + 21 + 27 + 30 + 25) 9 = 1,224	1,524	217.71

*Indicates the minimum average cost/week.

From Table it is clear that the average cost/week is minimum for the fourth week. Hence, the group replacement period is four weeks.

Individual replacement cost/week = Rs. 207

Minimum group replacement cost/week = Rs. 196.50

Since the minimum group replacement cost/week is less than the individual replacement cost/week, the group replacement policy is the best, and hence all the transistors should be replaced once in four weeks and the transistors which fail during this four-week period are to be replaced individually.