SHIVAJI UNIVERSITY, KOLHAPUR

CENTRE FOR DISTANCE AND ONLINE EDUCATION

Operations Research

(Mathematics)

For

M. Sc. Part-I : Semester-II

(In accordance with National Education Policy 2020)
(Academic Year 2022-23 onwards)

First Edition 2017
Revised Edition 2023

Prescribed for **M. Sc. Part-I**

Copies : 500

Published by:
Dr. V. N. Shinde
Ag. Registrar,
Shivaji University,
Kolhapur-416 004

Printed by :

Shri. B. P. Patil
Superintendent,
Shivaji University Press,
Kolhapur-416 004

ISBN- 978-81-8486-690-2

★ Further information about the Centre for Distance and Online Education & Shivaji University may be obtained from the University Office at Vidyanagar, Kolhapur-416 004, India.

Centre for Distance and Online Education
Shivaji University, Kolhapur

ADVISORY COMMITTEE

Centre for Distance and Online Education
Shivaji University, Kolhapur

MEMBERS OF B.O.S. IN MATHEMATICS

Preface

Large number of students appears for M. Sc. examinations externally every year. In view of this, Shivaji University has introduced the distance education mode for external students from the year 2008-09 and entrust the task to us to prepare the Self Instructional Material (SIM) for aspirants. An objective of the SIM is to provide students the material on the subject from which they can prepare for examination on their own without the help of a tutor. Today we are extremely happy to present the book on "Operations Research" for M. Sc. Part I students as a SIM prepared by well devoted experts. We hope that the exposition of the material in the book will meet the needs of all the students.

In the context of tremendous pace at which engineering and technology is advancing, the scientist and engineer who has to interpret science to practical end has the obligation to keep himself alert to understand the implications and complexities of science and engineering before he can utilize them to the benefits of his fellowmen. The subject matter covered in this book brings better awareness in planning, scheduling, cost and job control to the efficient and economical conduct of industrial projects for which the optimum use of men, money, machines and materials and their management at all levels is necessary.

The main aim of this book is to make clear the fundamentals of operations research and its techniques used in different fields of interests. This book covers topics like Convex Sets, Theory of Linear Programming problems, Duality Theorem, Information Theory.

An attempt has been made to make the presentation of the various units comprehensive, rigorous and yet simple. Numerous examples have been solved for the use of students. Although the book is aimed to M. Sc. Distance Education Students, even SET/NET aspirant students and students of management and engineering would find it useful.

I owe a deep sense of gratitude to the Vice Chancellor Prof. (Dr.) D. T. Shirke who has given impetus to go ahead with ambitious project like the present one. I thank Mr. Dayanand Gawade, Assistant Professor, Centre for Distance & Online Education for his continuous help to complete this book. I also thank Dr. D. K. More, Director, Center for Distance and Online Education, Shivaji University, Kolhapur for his help and keen interests in the completion of SIM.

Prof. S. H. Thakar
Department of Mathematics,
Shivaji University, Kolhapur-416004.

Writing Team	Unit No.
Dr. Mrs. S. H. Thakar Dept. of Mathematics, Shivaji University, Kolhapur.	**1 to 4**

■ **Editor** ■

Dr. S. H. Thakar
Professor & I/c. Dean of Science
Head of Department of Mathematics,
Shivaji University, Kolhapur.
Maharashtra.

M. Sc. (Mathematics)

Operation Research

Contents

Each Unit begins with the section Objectives -

Objectives are directive and indicative of :

1. What has been presented in the Unit and
2. What is expected from you
3. What you are expected to know pertaining to the specific Unit once you have completed working on the Unit.

The self check exercises with possible answers will help you to understand the Unit in the right perspective. Go through the possible answer only after you write your answers. These exercises are not to be submitted to us for evaluation. They have been provided to you as Study Tools to help keep you in the right track as you study the Unit.

UNIT 01

CONVEX SETS AND FUNCTIONS

1.0 INTRODUCTION

The roots of operations research can be found when early attempts were made to use a scientific approach in technical problems and in the management of organisations at the time of world war II. Britian had very limited military resources and therefore there was an urgent need to allocate resources to the various military operations and to the activities of each operation in an effective manner. Therefore the british military executives and managers called upon a team of scientists to apply a scientific method to study the technical problems related to air and land defence of the country. As the team was dealing with (military) operations the work of this team of scientists was named as OR in Britian.

Their efforts were instrumental in winning the air battle of Britian, and of the North Attantic etc.

The success of this team of scientists in Britian encouraged United States, Canada and France to start with such efforts. The work of this team was given various names in United States such as Operational analysis, operations evaluation operations research etc.

The apparent success of OR in the military attracted the attention of industrial management in this new field. In this way OR began to creep into industry and many governmental organisations.

After the war, many scientists were motivated to pursue research relevant in this new branch. The first technique in this field called the simplex method for solving linear programming problem was developed by American mathematician, George Dantzing in 1947. Since then many techniques such as quadratic programming, dynamical programming, inventory theory, queing theory etc. are developed. Thus the impact of OR can be experienced in almost all walks of life.

Definition of OR

We give few definitions of OR.

1) OR is the application of the theories of probability, linear programming, queuing theory etc. to the problems of war, industry, agriculture and many organisation.
2) OR is the art of winning war without actually fighting.
3) OR is the art of giving bad answers to the problems where otherwise the worse answers are given. (T. L. Saathy 58)

Use of OR

In general we can say that whenever there is a problem there is OR for help. In addition to the military operations research is widely used in many organisations. Now we discuss the scope of OR in various fields.

1) **Defence :** There is a necessity to formulate optimum strategies that may give maximum benefit. OR helps the military executives to select the best course of action to win the battle.

2) **Industry :** The company executives require the use of OR for the following :

   1) Production department to minimize the cost of production.

   2) Marketing department to maximize the amount sold and to minimize the cost of sales.

   3) Finance department to minimize the capital required to maintain any level of business.

   The various departments come in conflict with each other as the policy of one department is against the policy of the other. This difficulty is solved by the application of OR techniques. Thus OR has great scope in industry. Now a days almost all big industries in India make use of OR techniques.

3) **L. I. C. :** OR techniques are applicable to enable L. I. C. officers to decide the premium rates of various policies in the best interest of the corporation.

4) **Agriculture :** With the increase of population and resulting shortage of food there is a need to increase agriculture output for a country. But there are many problems faced by the agriculture department of a country. e. g. (i) climate conditions (ii) Problem of optimal distribution of water from the resources etc.

   Thus there is a need of the policy under the given restrictions for which OR techniques are useful to determine the best policies.

5) **Planning :** Careful planning plays an important role in the economic development of many organisations for which OR techniques are fruitful for such planning.

CONVEX SETS AND THEIR PROPERTIES

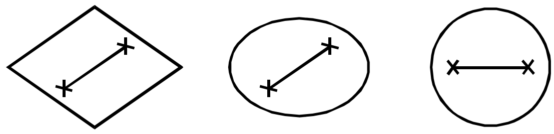
1.1 Definition I (Convex Set) Let $R^n = \left\{ \bar{x} = (x_1, x_2, \ldots, x_n) \middle| x_i \in R,\ i = 1, 2, \ldots, n \right\}$

A subset $S \subset R^n$, is said to be convex, if for any two points $\bar{x}_1, \bar{x}_2$ in S the line segment joining the points $\bar{x}_1$ and $\bar{x}_2$ is also contained in S.

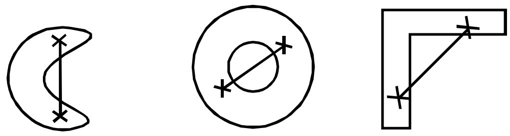
In other words, a subset $S \subset R^n$ is convex, if and only if

$$\bar{x}_1, \bar{x}_2 \in S \Rightarrow \lambda\, \bar{x}_1 + (1-\lambda)\, \bar{x}_2 \in S \ ; \ 0 < \lambda \le 1$$

Some convex and non - convex sets in R^2 are given below.



Convex Sets



Non - convex Sets

Example 1.1

Show that the set $S=\{(x_1,x_2):3x_1^2+2x_2^2\le 6\}$ is convex.

Solution :

Let $\bar{x},\bar{y}\in S$ where $\bar{x}=(x_1,x_2)$ and $\bar{y}=(y_1,y_2)$.

Since $\bar{x},\ \bar{y}\in S,\ 3x_1^2+2x_2^2\le 6$ and $3y_1^2+2y_2^2\le 6$.

The line segment joining $\bar{x}$ and $\bar{y}$ is the set

$\{\bar{u}:\bar{u}=\lambda\,\bar{x}+(1-\lambda)\,\bar{y},0\le\lambda\le 1\}$

For some $\lambda,0\le\lambda\le 1$, let $\bar{u}=(u_1,u_2)$ be a point of this set, so that

$u_1=\lambda\,x_1+(1-\lambda)\,y_1$, and $u_2=\lambda\,x_2+(1-\lambda)\,y_2$

Now,

$$3u_1^2+2u_2^2=3\left[\lambda\,x_1+(1-\lambda)\,y_1\right]^2+2\left[\lambda\,x_2+(1-\lambda)\,y_2\right]^2$$

$$=\lambda^2\left(3x_1^2+2x_2^2\right)+(1-\lambda)^2\left[3y_1^2+2y_2^2\right]+2\lambda(1-\lambda)(3x_1y_1+2x_2y_2)$$

$$\le 6\lambda^2+6(1-\lambda)^2+12\lambda(1-\lambda)=6$$

Since $(x_1-y_1)^2\ge 0$, $x_1y_1\le\frac{1}{2}\left(x_1^2+y_1^2\right)$ similarly $x_2y_2\le\frac{1}{2}\left(x_2^2+y_2^2\right)$ and $3x_1y_1+2x_2y_2\le 6$ and we have

$3u_1^2+2u_2^2\le 6$ and hence $\bar{u}=(u_1,u_2)\in S$.

Hence S is a convex set.

Example 1.2

In R^n consider,

$S_1 = \{\bar{x} \mid |\bar{x}| \le 1\}$ where $|\bar{x}| = \left(x_1^2 + x_2^2 + \ldots + x_n^2\right)^{1/2}$

Take $\bar{x}_1, \bar{x}_2 \in S$

Then $|\bar{x}_1| \le 1, |\bar{x}_2| \le 1$ and for $0 \le \lambda \le 1$,

$$|\lambda\,\bar{x}_1 + (1-\lambda)\,\bar{x}_2| \le |\lambda|\,|\bar{x}_1| + |(1-\lambda)\,\bar{x}_2|$$

$$= \lambda\,|\bar{x}_1| + (1-\lambda)\,|\bar{x}_2| \le 1$$

$\Rightarrow \lambda\,\bar{x}_1 + (1-\lambda)\,\bar{x}_2 \in S_1 \Rightarrow S_1$ is a convex set.

Example 1.3

Show that $C = \left\{(x_1, x_2) \,\middle|\, 2x_1 + 3x_2 = 7\right\} \subseteq R^2$ is convex set.

Solution :

Let $\bar{x} = (x_1, x_2)$ and $\bar{y} = (y_1, y_2) \in C$ and let $o \le \lambda \le 1$.

Let $\bar{w} = \lambda\,\bar{x} + (1-\lambda)\,\bar{y} = (w_1, w_2)$

$$\Rightarrow \bar{w} = \lambda(x_1, x_2) + (1-\lambda)(y_1, y_2)$$

$$\Rightarrow (w_1, w_2) = (\lambda x_1 + (1-\lambda) y_1, \lambda x_2 + (1-\lambda) y_2)$$

$$\Rightarrow w_1 = \lambda x_1 + (1-\lambda) y_1, w_2 = \lambda x_2 + (1-\lambda) y_2$$

We have $2w_1 + 3w_2 = 2(\lambda x_1 + (1-\lambda) y_1) + 3(\lambda x_2 + (1-\lambda) y_2)$

$$\Rightarrow 2w_1 + 3w_2 = \lambda(2x_1 + 3x_2) + (1-\lambda)(2y_1 + 3y_2)$$

Since $\bar{x}, \bar{y} \in C$, $2x_1 + 3x_2 = 7$, $2y_1 + 3y_2 = 7$

Hence $2w_1 + 3w_2 = \lambda.7 + (1-\lambda).7 = 7$

$\Rightarrow \bar{w} = (w_1, w_2) = \lambda\,\bar{x} + (1-\lambda)\,\bar{y} \in C, \forall\,\lambda,\ 0 \le \lambda \le 1.$

Hence C is a convex set.

Example 1.4

Show that $S=\left\{(x_1,x_2,x_3)\middle|2x_1-x_2+x_3\le 4\right\}\subseteq R^3$ is a convex set.

Solution :

Let $\bar{x}=(x_1,x_2,x_3)$ and $\bar{y}=(y_1,y_2,y_3)$ be any two points in S. Then by hypothesis,

$2x_1-x_2+x_3\le 4$, $2y_1-y_2+y_3\le 4$ .... (i)

Let $\bar{w}=(w_1,w_2,w_3)=\lambda\bar{x}+(1-\lambda)\bar{y}$ where $0\le\lambda\le 1$

$$\Rightarrow\bar{w}=\lambda(x_1,x_2,x_3)+(1-\lambda)(y_1,y_2,y_3)$$

$$\Rightarrow\bar{w}=(\lambda x_1,\lambda x_2,\lambda x_3)+((1-\lambda)y_1,(1-\lambda)y_2,(1-\lambda)y_3)$$

$$\Rightarrow\bar{w}=(\lambda x_1+(1-\lambda)y_1,\lambda x_2+(1-\lambda)y_2,\lambda x_3+(1-\lambda)y_3)$$

$$\Rightarrow w_1=\lambda x_1+(1-\lambda)y_1, w_2=\lambda x_2+(1-\lambda)y_2, w_3=\lambda x_3+(1-\lambda)y_3$$

We have,

$$2w_1-w_2+w_3=2(\lambda x_1+(1-\lambda)y_1)-(\lambda x_2+(1-\lambda)y_2)+(\lambda x_3+(1-\lambda)y_3)$$

$$=\lambda(2x_1-x_2+x_3)+(1-\lambda)(2y_1-y_2+y_3)$$

$$\le\lambda.4+(1-\lambda)4=4 \quad \text{.......... by (1)}$$

$$\Rightarrow\bar{w}=\lambda\bar{x}+(1-\lambda)\bar{y}\in S \text{ for all } \bar{x},\bar{y},\in S \text{ and for all } \lambda \text{ such that } 0\le\lambda\le 1$$

$\Rightarrow$ S is a convex set.

Example 1.5

Show that in R^3, $S=\left\{(x_1,x_2,x_3)\middle|\|x\|^2=x_1^2+x_2^2+x_3^2\le 1\right\}$ is a convex set.

Solution :

Let $\bar{x}=(x_1,x_2,x_3)$ and $\bar{y}=(y_1,y_2,y_3)\in S$.

Then $\|\bar{x}\|^2=x_1^2+x_2^2+x_3^2\le 1$ and $y_1^2+y_2^2+y_3^2=\|y\|^2\le 1$ .... (i)

Let $0\le\lambda\le 1$ and $\bar{z}=\lambda\bar{x}+(1-\lambda)\bar{y}$ where $\bar{z}=(z_1,z_2,z_3)$

Then $\bar{z}=\lambda(x_1,x_2,x_3)+(1-\lambda)(y_1,y_2,y_3)$

$$\Rightarrow\bar{z}=(\lambda x_1,\lambda x_2,\lambda x_3)+((1-\lambda)y_1,(1-\lambda)y_2,(1-\lambda)y_3)$$

$$\Rightarrow \bar{z} = (\lambda x_1 + (1-\lambda) y_1, \lambda x_2 + (1-\lambda) y_2, \lambda x_3 + (1-\lambda) y_3)$$

$$\Rightarrow \|\bar{z}\|^2 = [\lambda x_1 + (1-\lambda) y_1]^2 + (\lambda x_2 + (1-\lambda) y_2)^2 + (\lambda x_3 + (1-\lambda) y_3)^2$$

$$\Rightarrow \|\bar{z}\|^2 = \lambda^2 [x_1^2 + x_2^2 + x_3^2] + (1-\lambda)^2 [y_1^2 + y_2^2 + y_3^2] + 2\lambda(1-\lambda)(x_1 y_1 + x_2 y_2 + x_3 y_3) \text{ (ii)}$$

For i = 1, 2, 3, since $(x_i - y_i)^2 \geq 0$, $x_i y_i \leq \frac{1}{2}(x_i^2 + y_i^2)$ and therefore

$$x_1 y_1 + x_2 y_2 + x_3 y_3 \leq \frac{1}{2}[x_1^2 + x_2^2 + x_3^2 + y_1^2 + y_2^2 + y_3^2]$$

$$\leq \frac{1}{2}(1+1) = 1$$

Thus, $x_1 y_1 + x_2 y_2 + x_3 y_3 \leq 1$ .... (iii)

Hence from (i), (ii) and (iii) we have

$$\|\bar{z}\|^2 \leq \lambda^2 + (1-\lambda)^2 + 2\lambda(1-\lambda)1.1 = [\lambda + (1-\lambda)]^2 = 1$$

$\Rightarrow \lambda \bar{x} + (1-\lambda)\bar{y} = \bar{z} \in S$ for all $\bar{x}, \bar{y} \in S$ and for all λ such that $0 \leq \lambda \leq 1$.

$\Rightarrow S$ is a convex set.

Theorem 1.1

The intersection of any finite number of convex sets is a convex set.

Proof

Let $S_1, S_2, \ldots, S_n$ be a finite number of convex sets, and let $S = S_1 \cap S_2 \cap \ldots \cap S_n$.

Let $\bar{x}, \bar{y} \in S$ and $0 \leq \lambda \leq 1$

Then $\bar{x}, \bar{y} \in S_i$ for each i = 1, 2, ..., n where each S_i is a convex set. Then

$\lambda \bar{x} + (1-\lambda)\bar{y} \in S_i$ for each i = 1, 2,..., n

$\Rightarrow \lambda \bar{x} + (1-\lambda)\bar{y} \in S_, \cap S_2 \cap \ldots S_n = S$

$\Rightarrow S$ is a convex set.

Theorem 1.2

Let S and T be convex sets in R^n. Then $\alpha S + \beta T$ is also convex for any α, β in R.

Proof

Let $\bar{x}, \bar{y} \in \alpha S + \beta T$

Then $\bar{x}=\alpha\bar{u}_1+\beta\bar{v}_1$ and $\bar{y}=\alpha\bar{u}_2+\beta\bar{v}_2$, where $\bar{u}_1,\bar{u}_2\in S$ and $\bar{v}_1,\bar{v}_2\in T$

For any λ with $0\le\lambda\le 1$, we have

$$\lambda\bar{x}+(1-\lambda)\bar{y}=\lambda(\alpha\bar{u}_1+\beta\bar{v}_1)+(1-\lambda)(\alpha\bar{u}_2+\beta\bar{v}_2)$$

$$\Rightarrow\lambda\bar{x}+(1-\lambda)\bar{y}=\alpha(\lambda\bar{u}_1+(1-\lambda)\bar{u}_2)+\beta(\lambda\bar{v}_1+(1-\lambda)\bar{v}_2)$$

$\bar{u}_1,\bar{u}_2\in S$, S is convex.

$\therefore\ \lambda\bar{u}_1+(1-\lambda)\bar{u}_2\in S$

Similarly, $\lambda\bar{v}_1+(1-\lambda)\bar{v}_2\in T$

$$\lambda\bar{x}+(1-\lambda)\bar{y}\in\alpha S+\beta T,$$

Hence $\alpha S+\beta T$ is convex.

Definition 1.2

A convex combination of a finite number of points $\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n$ is a point

$$\bar{x}=\lambda_1\bar{x}_1+\lambda_2\bar{x}_2+\ldots+\lambda_n\bar{x}_n$$

where $\lambda_1,\lambda_2,\ldots,\lambda_n\ge 0$ and $\lambda_1+\lambda_2+\ldots+\lambda_n=1$

Remark

From this definition it follows that a subset $K\subseteq R^n$ is convex, if convex combination of any two points of K belongs to K.

Theorem 1.3

For a set K to be convex it is necessary and sufficient that every convex combination of points in K belongs to K.

Proof

Let every convex combination of points in K belong to K.

Then every convex combination of two points in K belongs to K.

Therefore K is convex. Hence the condition is sufficient.

Converly let K be convex.

To prove that the condition is necessary we shall follow the method of induction. We shall first prove that if the condition is true for r points it is also true for r + 1 points.

Let $\sum_{i=1}^{r}\lambda_i\bar{x}_i\in K$ where K is convex and $\bar{x}_i\in K,\sum_{i=1}^{r}\lambda_i=1,\lambda_i\ge 0,i=1,2,\ldots,r$

Consider $\sum_{i=1}^{r+1} \mu_i \overline{x}_i, \overline{x}_i \in K, \sum_{i=1}^{r+1} \mu_i = 1, \mu_i \geq 0$, $i = 1, 2, \ldots, r+1$

Here two cases arise.

i) $\mu_{r+1} = 0$

ii) $\mu_{r+1} \neq 0$

Case (I)

$$\mu_{r+1} = 0 \Rightarrow \sum_{i=1}^{r+1} \mu_i \overline{x}_i = \sum_{i=1}^{r} \mu_i \overline{x}_i \in K$$

Since by hypothesis $\mu_i \geq 0$ and $\sum_{i=1}^{r} \mu_i = 1$.

Case (II)

$$\mu_{r+1} \neq 0 \Rightarrow \sum_{i=1}^{r+1} \mu_i \overline{x}_1 = \left(1 - \mu_{r+1}\right) \frac{\sum_{i=1}^{r} \mu_i \overline{x}_i}{\left(1 - \mu_{r+1}\right)} + \mu_{r+1} \overline{x}_{r+1}$$

$$= \left(1 - \mu_{r+1}\right) \overline{y} + \mu_{r+1} \overline{x}_{r+1}$$

where $\overline{y} = \dfrac{\sum_{i=1}^{r} \mu_i \overline{x}_i}{\left(1 - \mu_{r+1}\right)} = \sum_{i=1}^{r} \dfrac{\mu_i}{1 - \mu_{r+1}} \overline{x}_i = \sum_{i=1}^{r} \lambda_i \overline{x}_i$ and $\sum_{i=1}^{r+1} \mu_i = 1$

and $\sum_{i=1}^{r} \lambda_i = \dfrac{\sum_{i=1}^{r} \mu_i}{1 - \mu_{r+1}} = \dfrac{1 - \mu_{r+1}}{1 - \mu_{r+1}} = 1$

Thus $\lambda_i \geq 0$, $\sum_{i=1}^{r} \lambda_i = 1$ and therefore $\overline{y} \in K$.

Hence $\sum_{i=1}^{r+1} \mu_i \overline{x}_i = \left(\sum_{i=1}^{r} \mu_i \right) \overline{y} + \mu_{r+1} \overline{x}_{r+1} = \left(1 - \mu_{r+1}\right) \overline{y} + \mu_{r+1} \overline{x}_r \in K$ becasue the right hand side is the convex linear combination of two points $\overline{y}$ and $\overline{x}_{r+1}$ in K which by hypothesis is convex.

This proves the theorem for r + 1 points. It is true for r = 2 by definition. Hence theorem is proved.

Definition 1.3

The convex hull of a set S is the intersection of all convex sets containing S. We shall denote by [S] the convex hull of S.

Remark

Every set has a convex hull , because R^n is a convex set and so there is always at least one convex set R^n of which every set is a subset. Also a convex set is its own convex hull.

Theorem 1.4

The convex hull of S is the set of all finite convex combinations of points in S.

Proof

Let K be the set of all finite convex combination of the points in S.

Then by theorem 1.3, K is a convex set containing S.

Hence $S \subseteq K$. Let K_1 be any convex set which contains S. Then K_1 contains all convex combinations of points in K_1. Hence it contains all convex combinations of points in S.

Hence $K \subseteq K_1$.

Thus K is a subset of all convex sets containing S which shows that K is the intersection of all convex sets containing S. Hence K = [S].

i.e. K is the convex hull of S.

Theorem 1.5

The set of all convex combinations of a finite number of points $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m$ is a convex set.

Proof

$$\text{Let } S = \left\{ \bar{x} \,\middle|\, \bar{x} = \sum_{i=1}^{m} \lambda_i \bar{x}_i, \lambda_i \geq 0, \sum_{i=1}^{m} \lambda_i = 1 \right\}$$

To show that S is a convex set take $\bar{x}'$ and $\bar{x}''$ in S, so that $\bar{x}' = \sum_{i=1}^{m} \lambda'_i \bar{x}_i$ where $\lambda'_i \geq 0$ and $\sum_{i=1}^{m} \lambda'_i = 1$ and $\bar{x}'' = \sum_{i=1}^{m} \lambda''_i \bar{x}_i$ where $\lambda''_i \geq 0$ and $\sum_{i=1}^{m} \lambda''_i = 1$.

Consider the vector $\bar{x} = \lambda \bar{x}' + (1-\lambda) \bar{x}'', 0 \leq \lambda \leq 1$

$$\Rightarrow \qquad \bar{x} = \lambda \sum_{i=1}^{m} \lambda_i \bar{x}_i + (1-\lambda) \sum_{i=1}^{m} \lambda''_i \bar{x}_i$$

$$\Rightarrow \qquad \bar{x} = \sum_{i=1}^{m} \left[\lambda \lambda'_i + (1-\lambda) \lambda''_i\right] \bar{x}_i$$

We can write $\bar{x} = \sum_{i=1}^{m} \mu_i \bar{x}_i$

where $\mu_i = \lambda \lambda'_i + (1-\lambda) \lambda''_i$

Since $0 \le \lambda \le 1, \lambda'_i \ge 0, \lambda''_i \ge 0$ it follows that $\mu_i \ge 0 \; \forall \, i = 1, 2, \ldots, m$. Also

$$\sum_{i=1}^{m} \mu_i = \sum_{i=1}^{m} \{\lambda \lambda'_i + (1-\lambda) \lambda''_i\}$$

$$= \lambda \sum_{i=1}^{m} \lambda'_i + (1-\lambda) \sum_{i=1}^{m} \lambda''_i = \lambda.1 + (1-\lambda)1 = 1$$

Hence $\bar{x}$ is a convex combination of $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m \Rightarrow \bar{x} \in S$.

Thus for each pair of points $\bar{x}'$ and $\bar{x}''$ in S the line segment joining them is contained in S. Hence S is a convex set.

Theorem 1.6

Every point of [S] can be expressed as a convex combination of at most (n + 1) points of $S \subseteq R^n$.

Proof

By definition of convex hull and theorem 1.1, [S] is a convex set.

Let $\bar{x}_i \in S, i = 1, 2, \ldots, m$.

$$\bar{x} = \sum_{i=1}^{m} \lambda_i \bar{x}_i, \sum_{i=1}^{m} \lambda_i = 1, \lambda_i \ge 0 \,, \quad \bar{x} \in [S]$$

Now $\bar{x} \in [S]$ can be expressed as a convex combination of points in S follows from the above theorem (1.3). What we have to prove now is that for any given $\bar{x}$ we can always find $m \le n+1$.

Let us suppose if possible that there is an $\bar{x} \in [S]$ for which m > n + 1. Since the space R^n is n - dimensional, not more than n vectors in R^n can be linearly independent. Consider the vectors, $\bar{x}_1 - \bar{x}_m, \bar{x}_2 - \bar{x}_m, \ldots, \bar{x}_{m-1} - \bar{x}_m$.

Since m - 1 > n these (m - 1) vectors cannot be linearly independent.

Hence it is possible to find $\alpha_i, i = 1, 2, \ldots, m-1$ not all zero such that

$$\sum_{i=1}^{m-1} \alpha_i (\bar{x}_i - \bar{x}_m) = \bar{0}$$

or $$\sum_{i=1}^{m-1} \alpha_i \bar{x}_i - \left(\sum_{i=1}^{m-1} \alpha_i \right) \bar{x}_m = \bar{0}$$

or $\sum_{i=1}^{m} \alpha_i \bar{x}_i = \bar{0}$ where $\alpha_m = -\sum_{i=1}^{m-1} \alpha_i$

or $$\sum_{i=1}^{m} \alpha_i = 0$$

Let $\mu_i = \lambda_i - \beta \alpha_i, i = 1, 2, \ldots, m$. Since $\lambda_i \geq 0$ we can choose β such that $\mu_i \geq 0$ with $\mu_i = 0$ for at least one i. This will happen if $\beta = \overset{\text{mim}}{i} \left\{ \frac{\lambda_i}{\alpha_i} \right\}$ over those values of i for which $\alpha_i > 0$ or $\beta_i = \max \left\{ \frac{\lambda_i}{\alpha_i} \right\}$ over i for which $\alpha_i < 0$.

Also $$\sum_{i=1}^{m} \mu_i = \sum_{i=1}^{m} \lambda_i - \beta \sum_{i=1}^{m} \alpha_i = 1 \qquad \left[\sum_{1}^{m} \lambda_i = 1 \ \& \ \sum_{1}^{m} \alpha_i = 0 \right]$$

Now $$\sum_{i=1}^{m} \mu_i \bar{x}_i = \sum_{i=1}^{m} \lambda_i \bar{x}_i - \sum_{i=1}^{m} \beta \alpha_i \bar{x}_i$$

$$= \sum_{i=1}^{m} \lambda_i \bar{x}_i = \bar{x}$$

(Since $\sum_{i=1}^{m} \alpha_i \bar{x}_i = 0$)

Since at least one $\mu_i = 0$ it follows that $\bar{x}$ is a convex linear conbination of at most (m - 1) points. If m - 1 > n + 1 we can again apply the above argument and express $\bar{x}$ as a convex combination of m - 2 points, and so on till m - k = n + 1, k > 0. This proves the theorem.

Definition 1.4

A point $\bar{x}$ of a convex set K is an extreme point or vertex of K if it is not possible to find two points $\bar{x}_1, \bar{x}_2$ in K such that

$$\bar{x} = (1-\lambda)\bar{x}_1 + \lambda\,\bar{x}_2, 0 < \lambda < 1$$

A point of K which is not a vertex of K is called an internal point of K.

Theorem 1.7

The set of all internal points of a convex set K is again a convex set.

Proof

Let V be the set of vertices of K. Then K - V is the set of internal points.

Let $\bar{x}_1, \bar{x}_2 \in K - V$. Then $\bar{x}_1, \bar{x}_2 \in K$ and $\bar{x}_1, \bar{x}_2 \notin V$

Hence $\bar{x} = (1-\lambda)\bar{x}_1 + \lambda\,\bar{x}_2 \in K, 0 < \lambda < 1$, is by definition not a vertex of K, but $\bar{x} \in K$.

i. e. $\bar{x} \in K - V$.

Hence K - V is a convex set.

Definition 1.5

The set of all convex combinations of a finite number of points $\bar{x}_i$, i = 1, 2, ..., m is the convex polynedron spanned by these points.

Theorem 1.8

The convex polyhedron is a convex set.

Proof

Let $\bar{y}_1$ and $\bar{y}_2$ be any two points in the polyhedron spanned by $\bar{x}_i, i = 1, 2, \ldots, m$

Then by definition

$$\bar{y}_1 = \sum_{i=1}^{m} \lambda_i\,\bar{x}_i, \sum_{i=1}^{m} \lambda_i = 1, \lambda_i \geq 0$$

$$\bar{y}_2 = \sum_{i=1}^{m} \mu_i\,\bar{x}_i, \sum_{i=1}^{m} \mu_i = 1, \mu_i \geq 0$$

Now let, $\bar{y} = (1-\alpha)\bar{y}_1 + \alpha\,\bar{y}_2, 0 \leq \alpha \leq 1$

$$\Rightarrow \bar{y} = (1-\alpha)\sum_{i=1}^{m} \lambda_i\,\bar{x}_i + \alpha \sum_{i=1}^{m} \mu_i\,\bar{x}_i$$

$$\Rightarrow \quad \bar{y} = \sum_{i=1}^{m} \left[(1-\alpha)\lambda_i + \alpha\,\mu_i\right] \bar{x}_i = \sum_{i=1}^{m} \beta_i\, \bar{x}_i \,,$$

where $\beta_i = (1-\alpha)\lambda_i + \alpha\,\mu_i$

Since $\sum_{i=1}^{m} \beta_i = (1-\alpha)\sum_{i=1}^{m} \lambda_i + \alpha \sum_{i=1}^{m} \mu_i = 1$, $\bar{y}$ is also in the polyhedron. Hence polyhedron is a convex set.

Theorem 1.9

The set of vertices of a convex polyhedron is a subset of its spanning points.

Proof

Let W be the set of points spanning the convex polyhedron, and V be the set of its vertices. If possible let $\bar{y} \in V$ but $\bar{y} \notin W$. Since $\bar{y}$ is in the poly hedron by definition it is a convex linear combination of points of W all of which are other than $\bar{y}$ (by assumption). Hence by definition $\bar{y}$ is not a vertex which is a contradiction. Therefore $\bar{y} \in W$ or $V \subseteq W$.

Remark

It is obvious that there can be spanning points which are not vertices. For example consider the points A, B, C, D in R^2 such that D is in the triangle formed by the vertices A, B, C. The four points span the triangle ABC but D is not a vertex.

HYPERPLANES AND HALF SPACES

Definition 1.5

Let $\bar{x} \in R^n, \bar{C}(\neq 0)$ a constant row n - vector and $\alpha \in R$. Then we define,

i) A hyperplane as $\{\bar{x} \mid \bar{c}\,\bar{x} = \alpha\}$

ii) A closed half - space as $\{\bar{x} \mid \bar{c}\,\bar{x} \le \alpha\}$ or $\{\bar{x} \mid \bar{c}\,\bar{x} \ge \alpha\}$

iii) An open half space as $\{\bar{x} \mid \bar{c}\,\bar{x} < \alpha\}$ or $\{\bar{x} \mid \bar{c}\,\bar{x} > \alpha\}$

Definition 1.6

A set $X \subseteq R^n$ is said to be an $\in -nbd$ of a point $\bar{x}_0 \in R^n$ if,

$\{\bar{x} \mid |\bar{x} - \bar{x}_0| < \in\} \subseteq X$ where $|\bar{x}| = |(x_1, x_2, \ldots, x_n)| = (x_1^2 + x_2^2 + \ldots + x_n^2)^{1/2}$

Definition 1.7

The $\delta.nbd$ of $\bar{x}$ in R^n is defined as the set of all points $\bar{y}$ in R^n such that $|\bar{y}-\bar{x}|<\delta$ (Where $\delta>0, \delta \in R$)

Definition 1.8

If R^n the point $\bar{x}$ is a boundry point of the set S if every δ - neighbourhood of $\bar{x}$ contains some points which are in S and some points which are not in S.

For example in

$S_1=\{\bar{x}\,|\,|\bar{x}|\le 1\}$, $S_2=\{\bar{x}\,|\,|\bar{x}|<1\}$, $\bar{x}\in R^2$ the points on the circumference of the circle $x_1^2+x_2^2=1$ are the boundry points of S_1 and S_2. S_1 contains all its boundry points while S_2 contains none of them.

Definition 1.9

A set is said to be closed if it contains all its boundry points and is said to be open if its complement is closed.

Definition 1.10

A set S is said to be bounded from below if there exists $\bar{y}$ in R^n with each component finite such that for every $\bar{x}\in S$, $\bar{y}\le\bar{x}$. $[\text{Note}: \bar{y}\le\bar{x}\Leftrightarrow y_j\le x_j, j=1,2,...,n]$.

Definition 1.11

A set S is bounded if there exists a finite real number $M\ge 0$ such that for all $\bar{x}$ in S, $|\bar{x}|\le M$.

Corollary 1.10

A hyperplane is a closed set.

Proof

Let $\{\bar{x}\,|\,\bar{c}\,\bar{x}=\alpha_0\}$ be a hyperplane.

Let $\bar{x}_1$ be the boundry point of the hyperplane. Suppose it is not a point of the hyper plane.

Then either $\bar{c}\,\bar{x}_1>\alpha_0$ or $\bar{c}\,\bar{x}_1<\alpha_0$.

Suppose $\bar{c}\,\bar{x}_1<\alpha_0$ and let $\bar{c}\,\bar{x}_1=\alpha_1<\alpha_0$

Now $\bar{c}\,\bar{x}=\bar{c}[\bar{x}_1+\bar{x}-\bar{x}_1]$

$$\Rightarrow \bar{c}\,\bar{x}=\bar{c}\,\bar{x}_1+\bar{c}(\bar{x}-\bar{x}_1)\Rightarrow \bar{c}\,\bar{x}\le|\bar{c}\,\bar{x}|=|\bar{c}\,\bar{x}_1+\bar{c}(\bar{x}-\bar{x}_1)|$$

$$\Rightarrow \bar{c}\,\bar{x} \le |\bar{c}\,\bar{x}_1| + |\bar{c}(\bar{x}-\bar{x}_1)|$$

$$\Rightarrow \bar{c}\,\bar{x} \le \alpha_1 + |\bar{c}(\bar{x}-\bar{x}_1)| \qquad \left[|\bar{c}\,\bar{x}_1| = |\alpha_1| = \alpha_1\right]$$

$$\Rightarrow \bar{c}\,\bar{x} \le \alpha_1 + |\bar{c}||\bar{x}-\bar{x}_1|$$

Consider the $\in$ nbd of $\bar{x}_1$, $\{\bar{x}\,|\,|\bar{x}-\bar{x}_1| < \in\}$ where $\in$ is an orbitary positive number.

Let $\in = \dfrac{\alpha_0 - \alpha_1}{2|\bar{c}|}$

Hence if $\bar{x}$ is in the $\in$-nbd of $\bar{x}_1$ we get $\bar{c}\,\bar{x} < \alpha_1 + \dfrac{(\alpha_0 - \alpha_1)}{2} = \dfrac{\alpha_0 + \alpha_1}{2} < \alpha_0$

This shows that $\bar{x}$ is in the half space $\bar{c}\,\bar{x} < \alpha_0$. Hence there exits a nbd. of $\bar{x}_1$ which contains no points of the hyperplane $\bar{c}\,\bar{x} = \alpha_0$. Hence $\bar{x}_1$ is not a boundry point of the hyperplane. This is a contradiction. Thus there is no boundry point of the hyper plane which is not in the hyperplane. Hence the hyperplane is a closed set.

Definition 1.12

In R^n, every hyper plane $\{\bar{x} / \bar{c}\,\bar{x} = \alpha\}$ determines two open half spaces and two closed half spaces. The open half spaces are :

$X_1 = \{\bar{x}\,|\,\bar{c}\,\bar{x} > \alpha\}$ and $X_2 = \{\bar{x}\,|\,\bar{c}\,\bar{x} < \alpha\}$

The closed half - spaces are

$X_3 = \{\bar{x}\,|\,\bar{c}\,\bar{x} \ge \alpha\}$ and $X_4 = \{\bar{x}\,|\,\bar{c}\,\bar{x} \le \alpha\}$.

Corollery 1.11

A hyperplane is a convex set.

Proof

Let $X = \{\bar{x}\,|\,\bar{c}\,\bar{x} = \alpha\}$ be a hyperplane and let $\bar{x}_1, \bar{x}_2$ be any two points of this hyperplane. Then $\bar{c}\,\bar{x}_1 = \alpha$ and $\bar{c}\,\bar{x}_2 = \alpha$. Now if $0 \le \lambda \le 1$, we have

$$\bar{c}[\lambda\,\bar{x}_1 + (1-\lambda)\bar{x}_2] = \bar{c}(\lambda\,\bar{x}_1) + \bar{c}(1-\lambda)\bar{x}_2$$

$$= \lambda(\bar{c}\,\bar{x}_1) + (1-\lambda)\bar{c}\,\bar{x}_2$$

$$= \lambda\,\alpha + (1-\lambda)\alpha = \alpha$$

Therefore the point $\lambda\,\bar{x}_1 + (1-\lambda)\bar{x}_2$ for $0 \le \lambda \le 1$ is in the hyperplane. Hence the hyperplane is a convex set.

Corollary 1.12

The closed half spaces $H_1 = \{\bar{x} | \bar{c}\,\bar{x} \geq \alpha\}$ and $H_2 = \{\bar{x} | \bar{c}\,\bar{x} \leq \alpha\}$ are convex sets.

Proof

Let $\bar{x}_1, \bar{x}_2$ be any two points of H_1. Then $\bar{c}\,\bar{x}_1 \geq \alpha$ and $\bar{c}\,\bar{x}_2 \geq \alpha$. If $0 \leq \lambda \leq 1$.

$$\bar{c}\left[\lambda\,\bar{x}_1 + (1-\lambda)\,\bar{x}_2\right] = \lambda\,(\bar{c}\,\bar{x}_1) + (1-\lambda)\,\bar{c}\,\bar{x}_2$$

$$\geq \lambda\,\alpha + 1(1-\lambda)\,\alpha = \alpha$$

$\Rightarrow \lambda\,\bar{x}_1 + (1-\lambda)\,\bar{x}_2 \in H_1$. Hence H_1 is a convex set. Similarly H_2 is a convex set.

Corollary 1.13

The open half spaces $H_1 = \{\bar{x} | \bar{c}\,\bar{x} > \alpha\}$ and $H_2 = \{\bar{x} | \bar{c}\,\bar{x} < \alpha\}$ are convex sets.

Proof

Let $\bar{x}_1, \bar{x}_2$ be any two points of H_1.

Then $\bar{c}\,\bar{x}_1 > \alpha, \bar{c}\,\bar{x}_2 > \alpha$

If $0 \leq \lambda \leq 1$, we have

$$\bar{c}\left[\lambda\,\bar{x}_1 + (1-\lambda)\,\bar{x}_2\right] = \lambda\left(\bar{c}\,\bar{x}_1\right) + (1-\lambda)\,\bar{c}\,\bar{x}_2$$

$$> \lambda\,\alpha + (1-\lambda)\,\alpha = \alpha$$

$\Rightarrow \lambda\,\bar{x}_1 + (1-\lambda)\,\bar{x}_2 \in H_1, \forall\, \bar{x}_1, \bar{x}_2 \in H_1$

$\Rightarrow H_1$ is a convex set.

Similarly H_2 is a convex set.

SUPPORTING AND SEPARATING HYPERPLANES

Definition 1.13 (Supporting hyperplane)

Let $S \subset R^n$ be any closed convex set and $\bar{w} \in S$ be a boundary point. Then a hyperplane $\bar{c}\,\bar{x} = z$ is called a supporting hyperplane of S at $\bar{w}$, if

i) $\bar{c} \cdot \bar{w} = z$ and

ii) $S \subset H_+$ or $S \subset H_-$

where $H_+ = \{\bar{x} : \bar{c}\,\bar{x} \geq z\}$ and $H_- = \{\bar{x} : \bar{c}\,\bar{x} \leq z\}$

Remarks

1) The supporting hyperplane need not be unique.

2) S may intersect the supporting hyperplane in more than one boundary points.

Theorem 1.14

Let S be a closed convex set. Then S has extreme points in every supporting hyperplane.

Proof

Let $\overline{w}$ be a boundary point of a closed convex set S.

Let $\overline{c}\,\overline{x}=z$ be a supporting hyperplane at $\overline{w}\in S$. Let $B=S\cap\left\{\overline{x}\,\middle|\,\overline{c}\,\overline{x}=z\right\}$.

Then B is a closed convex set and $B\neq\varphi$ for $\overline{w}\in B$.

We claim that every extreme point of B is also an extreme point of S.

Let us assume to the contrary that an extreme point $\overline{b}$ of B, is not an extreme point of S. Then there exist $\overline{x}_1,\overline{x}_2\in S$, such that

$$\overline{b}=\lambda\,\overline{x}_1+(1-\lambda)\,\overline{x}_2, \qquad 0<\lambda<1$$

Therefore $\overline{c}\,\overline{b}=\lambda\,\overline{c}\,\overline{x}_1+(1-\lambda)\,\overline{c}\,\overline{x}_2$. .... (i)

Since $\overline{c}\,\overline{x}=z$ is a supporting hyperplane at $\overline{w}$ and $\overline{x}_1,\overline{x}_2\in S$

$$\overline{c}\,\overline{x}_1\leq z \text{ and } \overline{c}\,\overline{x}_2\leq z$$

or $\overline{c}\,\overline{x}_1\geq z$ and $\overline{c}\,\overline{x}_2\geq z$ .... (ii)

From (i) and (ii)

$$\overline{c}\,\overline{b}\leq\lambda z+(1-\lambda)z=z \quad \text{or} \quad \overline{c}\,\overline{b}\geq\lambda z+(1-\lambda)z=z$$

Therefore $\overline{b}$ is not a point of B.

This is a contradiction.

Therefore every extreme point of B is also an extreme point of S.

Definition 1.14 (Separating hyperplane)

Let S and T be two non-empty subsets of R^n. The hyperplane H is said to separate S and T if H is contained in one of the closed half spaces generated by H and T is contained in the other closed half space. The hyperplane H is called separating hyperplane.

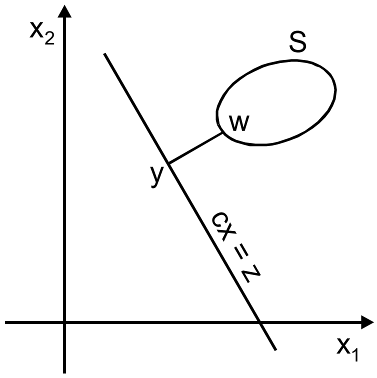
Remark :

A hyperplane H strictly separates S and T if S is contained in one of the open half spaces generated by H and T is contained in the other open half space.

Theorem 1.15 (Separating Hyperplane)

Let $S \subset R^n$ be a closed convex set. Then for any point $\bar{y}$ not is S, there is a hyperplane containing $\bar{y}$ such that S is contained in one of the open half spaces determined by the hyperplane.

Proof



We are given that $\bar{y} \notin S$.

Since S is a closed set, there exist $w \subset S$, such that,

$|\bar{w}-\bar{y}| = \min_{\bar{x} \in S} |\bar{x}-\bar{y}|$ i.e. $|\bar{w}-\bar{y}| \le |\bar{x}-\bar{y}|, \bar{w} \in S, \bar{x} \in S$ .... (i)

Observe that $|\bar{w}-\bar{y}| > 0$ (S is closed and $\bar{y} \notin S$)

Let $\bar{u}$ be any point of S. Since S is a convex set

$\left[\lambda \bar{u} + (1-\lambda)\bar{w}\right] \in S$ for $0 \le \lambda \le 1$ .... (ii)

From (i) and (ii)

$$|\lambda \bar{u} + (1-\lambda)\bar{w} - \bar{y}| \ge |\bar{w}-\bar{y}|$$

$$\Rightarrow \quad |(\bar{w}-\bar{y}) + \lambda(\bar{u}-\bar{w})|^2 \ge |\bar{w}-\bar{y}|^2$$

$$\Rightarrow \quad \lambda^2 |\bar{u}-\bar{w}|^2 + |\bar{w}-\bar{y}|^2 + 2\lambda(\bar{w}-\bar{y})(\bar{u}-\bar{w}) \ge |\bar{w}-\bar{y}|^2$$

$$\Rightarrow \quad \lambda^2 |\bar{u}-\bar{w}|^2 + 2\lambda(\bar{w}-\bar{y})(\bar{u}-\bar{w}) \ge 0$$

$$\Rightarrow \quad \lambda |\bar{u}-\bar{w}|^2 + 2(\bar{w}-\bar{y}) \cdot (\bar{u}-\bar{w}) \ge 0.$$

Letting $\lambda \to 0$, and $\bar{c} = (\bar{w} - \bar{y})$; we get

$$(\bar{w} - \bar{y})(\bar{u} - \bar{w}) \geq 0 \text{ or } \bar{c}(\bar{u} - \bar{w}) \geq 0 \text{ i.e. } \bar{c} \cdot \bar{u} \geq \bar{c} \cdot \bar{w}$$

or $\quad \bar{c} \cdot \bar{u} - \bar{c} \cdot \bar{y} \geq \bar{c} \cdot \bar{w} - \bar{c} \cdot \bar{y}$

or $\quad \bar{c}(\bar{u} - \bar{y}) \geq \bar{c}(\bar{w} - \bar{y}) = |\bar{c}|^2$

Hence $\bar{c}\,\bar{u} > \bar{c}\,\bar{y}$.

Putting $\bar{c}\,\bar{y} = z$, we get $\bar{c}\,\bar{u} > z$.

Thus $\bar{y}$ lies on the hyperplane $\bar{c}\,\bar{x} = z$ and for all $\bar{u} \in S$, $\bar{c}\,\bar{u} \geq z$.

This completes the proof.

CONVEX FUNCTIONS

Definition 1.14 (Convex Functions)

Let S be a non - empty convex subset of R^n. A function $f(\bar{x})$ on S is said to be convex if for any two vectors $\bar{x}_1$ and $\bar{x}_2$ in S.

$$f\left[\lambda \bar{x}_1 + (1-\lambda)\bar{x}_2\right] \leq \lambda f(\bar{x}_1) + (1-\lambda) f(\bar{x}_2) \qquad 0 \leq \lambda \leq 1$$

Definition 1.15 (Strictly convex function)

Let S be a non empty convex subset of R^n. A function f (x) on S is said to be strictly convex if for any two different vectors x_1 and x_2 is S.

$$f\left[\lambda \bar{x}_1 + (1-\lambda)\bar{x}_2\right] < \lambda f(\bar{x}_1) + (1-\lambda) f(\bar{x}_2) \qquad 0 < \lambda < 1$$

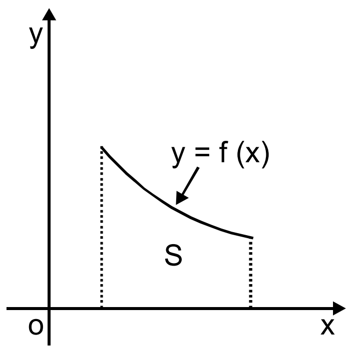


Fig A : Strictly Convex Function

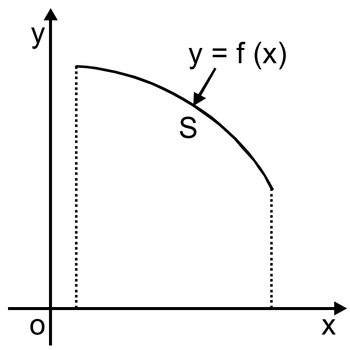


Fig B : Strictly Concave Function

It follows from the above two definitions that every strictly convex function is also convex. The graph of a strictly convex function has been illustrated in Fig. A.

Definition 1.16 [Concave (strictly concave) function]

A function $f(\bar{x})$ on a non - empty subset S of R^n is said to be concave (strictly concave) if $-f(\bar{x})$ is convex (strictly convex).

Clearly, every strictly concave function is also concave. The graph of a strictly concave function has been illustrated in Fig. B.

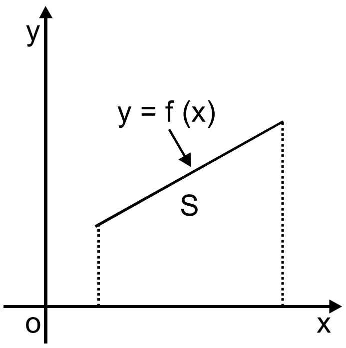


Fig C : Both Convex and Concave Functions

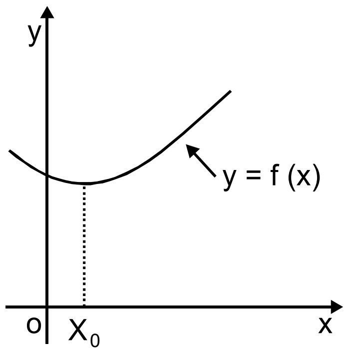


Fig D

It is possible for a function to be both convex and concave. For example, $f(\bar{x})=\bar{x}$ is such a function (Fig. C). The function in Fig. D is strictly convex for $\bar{x}\geq\bar{x}_0$ but not strictly convex for $\bar{x}<\bar{x}_0$.

The following results are the immediate consequences of the above definitions :

i) A linear function $z=c\bar{x}, \bar{x}\in R^n$ is a convex (concave) function but not strictly convex (concave).

ii) The sum of convex (concave) functions is convex (concave) and if at least one of the functions is strictly convex (concave) then so is their sum.

Note : In what follows we shall deal with convex functions only. However, all the results remain valid if we deal with concave functions.

LOCAL AND GLOBAL EXTREMA

In the problems of constrained optimization, we are interested in determining a vector $\bar{x}$ that minimises the function f ($\bar{x}$) [or maximises - f($\bar{x}$)] subject to the 'constraints' $g_i(\bar{x})\leq 0 \left(i=1,2,\ldots,m\right)$. The set of the vectors $\bar{x}$ satisfying these constraints is usually called the 'feasible region'.

Definition 1.17 (Global minima)

A global minimum of the function $f(\bar{x})$ is said to be attained at $\bar{x}_0$ if $f(\bar{x}_0) \le f(\bar{x})$ for all $\bar{x}$ in the feasible region.

Example : Function $f(\bar{x}) = x_1^2$, subject to the constraint $x_1 \ge 0$, has a minimum at $x_1 = 0$.

Definition 1.18 (Local minima)

A local minimum $f(\bar{x}_0)$ of function $f(\bar{x})$ is said to be attained at $\bar{x}_0$ if there exists a positive ε such thai $f(\bar{x}_0) \le f(\bar{x})$ for all $\bar{x}$ in the feasible region which also satisfy the condition $|\bar{x}_0 - \bar{x}| \le \varepsilon$.

Example :

The function $f(\bar{x}) = x_1^2 - x_1^3$ subject to the constraint $x_1 \ge 0$, has a local minimum at $x_1 = 0$. Note that $f(\bar{x})$ has no global minimum at all.

Note : The word extremum is used to indicate either maximum or minimum.

Theorem 1.16

Let $f(\bar{x})$ be a convex function on a convex set S. If $f(\bar{x})$ has a local minimum on S, then this local minimum is also a global minimum on S.

Proof :

Let $f(\bar{x})$ have a local minimum $f(\bar{x}_0)$ at $\bar{x}_0$ which is not a global minimum on S. Then, there exists at least one x_1 in $S(\bar{x}_1 \ne \bar{x}_0)$ such that $f(\bar{x}_1) < f(\bar{x}_0)$. Since $f(\bar{x})$ is a convex function on S, we have

$$f[\lambda \bar{x}_1 + (1-\lambda)\bar{x}_0] \le \lambda f(\bar{x}_1) + (1-\lambda) f(\bar{x}_0)$$

Also $\lambda f(\bar{x}_1) + (1-\lambda) f(\bar{x}_0) < \lambda f(\bar{x}_0) + (1-\lambda) f(\bar{x}_0) = f(\bar{x}_0)$

Thus $f[\lambda \bar{x}_1 + (1-\lambda)\bar{x}_0] \le f(\bar{x}_0)$

Now, for any $\varepsilon > 0$, we observe that

$$\left|[\lambda \bar{x}_1 + (1-\lambda)\bar{x}_0] - x_0\right| = \lambda |\bar{x}_1 - \bar{x}_0| < \varepsilon, \qquad \left(\text{if } \lambda < \frac{\varepsilon}{|\bar{x}_1 - \bar{x}_0|}\right)$$

Thus $\lambda \bar{x}_1 + (1-\lambda)\bar{x}_0$ will give a smaller value for $f(\bar{x})$ in the ε - neighbourhood of $\bar{x}_0$, whenever $\lambda < \min\left\{1, \frac{\varepsilon}{|\bar{x}_1 - \bar{x}_0|}\right\}$. This contradicts the fact that $f(\bar{x})$ takes on a local minimum at $\bar{x}_0$. Hence $\bar{x}_0$ is a global minimal point.

Corollary 1.17

If a function f(x) has a local minimum on a convex set S on which it is strictly convex, then this local minimum is also a global minimum on that set. This global minimum is attained at a single point.

Theorem 1.18

Let $f(\bar{x})$ be a convex function on a convex set S. Then the set of points in S at which $f(\bar{x})$ takes on its global minimum, is a convex set.

Proof :

The result is obvious if the global - minimum is attained at just a single point. Let us assume that the global minimum is attained at two different points $\bar{x}_1$ and $\bar{x}_2$ of S. Then $f(\bar{x}_1)=f(\bar{x}_2)$.

Now, since $f(\bar{x})$ is convex,

$$f\left[\lambda\bar{x}_2+(1-\lambda)\bar{x}_1\right]\le\lambda f(\bar{x}_2)+(1-\lambda)f(\bar{x}_1)=f(\bar{x}_2) \qquad 0\le\lambda\le1$$

$$\Rightarrow \quad f\left[\lambda\bar{x}_2+(1-\lambda)\bar{x}_1\right]\le f(\bar{x}_2)=f(\bar{x}_1)$$

$$\Rightarrow \quad f\left[\lambda\bar{x}_2+(1-\lambda)\bar{x}_1\right]\le f(\bar{x}_1)$$

Thus every point $\bar{x}=\lambda\bar{x}_2+(1-\lambda)\bar{x}_1$ corresponds to a global minima. The set of all such $\bar{x}$ is, obviously, a convex set.

Corollary 1.19

If the global minimum is attainable at two different points of S, then it is attainable at an infinite number of points of S.

Theorem 1.20

Let $f(\bar{x})$ be differentiable on its domain. If $f(\bar{x})$ is defined on an open convex set S, then $f(\bar{x})$ is convex if

$$f(\bar{x}_2)-f(\bar{x}_1)\ge(\bar{x}_2-\bar{x}_1)^T\nabla f(\bar{x}_1)$$

for all $\bar{x}_1,\bar{x}_2\in S$.

Proof :

We shall prove that if

$f(\bar{x}_2)-f(\bar{x}_1)\ge(\bar{x}_2-\bar{x}_1)^T\nabla f(\bar{x}_1)$ then $f(\bar{x})$ is convex.

Since $\overline{x}_1, \overline{x}_2 \in S, \overline{x}_0 = \lambda\,\overline{x}_2 + (1-\lambda)\,\overline{x}_1$ for $0 \le \lambda \le 1$ implies that $x_0 \in S$.

Using the above condition for $\overline{x}_1$ and $\overline{x}_0$, we have

$$f(\overline{x}_1) - f(\overline{x}_0) \ge (\overline{x}_1 - \overline{x}_0)^T \nabla f(\overline{x}_0) \qquad \text{.......... (i)}$$

Similarly, for $\overline{x}_2$ and $\overline{x}_0$,

$$f(\overline{x}_2) - f(\overline{x}_0) \ge (\overline{x}_2 - \overline{x}_0)^T \nabla f(\overline{x}_0) \qquad \text{.......... (ii)}$$

Multiplying (ii) by λ and (i) by $(1-\lambda)$ and then adding, we get

$$\lambda f(\overline{x}_2) + (1-\lambda) f(\overline{x}_1) \ge f(\overline{x}_0) + \left[\lambda\,\overline{x}_2^T + (1-\lambda)\,\overline{x}_1^T\right] \nabla f(\overline{x}_0) - \overline{x}_0^T \nabla f(\overline{x}_0)$$

$$= f(\overline{x}_0) + \overline{x}_0^T \nabla f(\overline{x}_0) - \overline{x}_0^T \nabla f(\overline{x}_0) = f(\overline{x}_0)$$

Using the definition of $\overline{x}_0$, this yields $\lambda f(\overline{x}_2) + (1-\lambda) f(\overline{x}_1) \ge f\left[\lambda\,\overline{x}_2 + (1-\lambda)\,\overline{x}_1\right]$,

which implies that $f(\overline{x})$ is convex.

◆ ◆ ◆ ◆ EXERCISES ◆ ◆ ◆ ◆

1) Define : Convex set, hyperplane, extreme point, convex combination of points.

2) a) Prove that a hyperplane is a convex set.

   b) Show that $c = \left\{\{x_1, x_2\} \middle| 2x_1 + 3x_2 = 7\right\} \subseteq R^2$ is a convex set.

   c) For any point $\overline{x}, \overline{y} \in R^n$ show that the line segment joining $\overline{x}, \overline{y}$ i. e. [x : y] is a convex set.

3) a) Show that $S = \left\{(x_1, x_2, x_3) \middle| 2x_1 - x_2 + x_3 \le 4\right\} \subseteq R^3$ is convex set.

   b) Show that in R^3 the closed ball $x_1^2 + x_2^2 + x_3^2 \le 1$ is a convex set.

   c) Show that a hyperplane in R^3 is a convex set.

4) a) Show that the closed half spaces $H_1 = \{\overline{x} / \overline{c}\,\overline{x} \ge z\}$ as $H_2 = \{\overline{x} / \overline{c}\,\overline{x} \le z\}$ are convex sets.

   b) The open half spaces $\{\overline{x} | \overline{c}\,\overline{x} > z\}$ and $\{\overline{x} | \overline{c}\,\overline{x} > z\}$ are convex sets.

   c) The intersection of any finite number of convex sets is a convex set.

5) a) Show that $S=\{(x_1,x_2,x_3) | 2x_1-x_2+x_3\le 4, x_1+2x_2-x_3\le 1\}$ is a convex set.

b) Let A be an $m\times n$ matrix and $\overline{b}$ be on n - vector then show that $\{\overline{x}\in R^n | A\overline{x}\le\overline{b}\}$ is a convex set.

c) Let S and T be convex sets in R^n. Then for any scalars α,β prove that $\alpha S+\beta T$ is a convex set.

d) Prove that the set of all convex combinations of a finite number of points $\overline{x}_1,\overline{x}_2,\ldots,\overline{x}_m$ is a convex set.

6) a) If V is any finite subste of vectors in R^n, then prove that the convex hull of V is the set of all convex combinations of vertors in V.

b) If $A=\{\overline{x},\overline{y}\}\subseteq R^n$ these prove that $<A>=[\overline{x}\,\overline{y}]$.

c) Prove that : A linear function $z=c_1x_1+c_2x_2+\ldots+c_nx_n$ defined over a convex polyhedron C takes its maximum (or minimum) value at an extreme point of C.

7) a) Let $S\subseteq R^n$ be a convex set with a nonempty interior. If $\overline{x}_1\in C/S$ and $\overline{x}_2\in \text{int}\,S$ then prove that for each $0<\lambda<1$ the point $\lambda\overline{x}_1+(1-\lambda)\overline{x}_2$ lies in int S.

b) If $S\subseteq R^n$ is a convex then prove that int S is also a convex set.

c) Let S be a convex set with a non empty interior. Then prove that cl S is also a convex set.

8) a) Let $S\subseteq R^n$ be a closed convex set and $\overline{y}\notin S$. Then prove that there exist unique $\overline{x}_0\in S$ such that $|\overline{y}-\overline{x}_0|=mn\{|\overline{y}-\overline{x}|=\overline{x}\in S\}$.

b) Let $X\subseteq R^n$ be a closed convex set. Then show that for any point $\overline{y}$ not in X. There exist a hyerplane containing $\overline{y}$ s. t. X is contained in one of the open half spaces determined by the hyperplane.

❑ ❑ ❑ ❑

UNIT 02

LINEAR PROGRAMMING PROBLEMS

2.0 INTRODUCTION

In 1947, George Dantzig and his associates, while working in the US department of Air Force, observed that a large number of military planning problems could be formulated as maximizing / minimizing a linear function (profit / cost) whose variables were restricted to values satisfying a system of linear constraints (e.g. 2x1 + 3x2 „T 5). The term programming refers to the process of determining a particular action plane. Since the objective function (profit / cost) and constraints are linear, problems are called linear programming problems.

The general Linear Programming Problem (L.P.P.)

The general linear programming problem is to find a vector $(x_1, x_2,..., x_n)$ which minimizes the linear form (i. e. objective function)

$$z = c_1x_1 + c_2x_2 + ... + c_nx_n \qquad .... (2.1)$$

subject to the linear constraints

$$x_J \geq 0 \quad (j = 1, 2, ..., n) \qquad .... (2.2)$$

and

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

$$a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \qquad .... (2.3)$$

$$a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m$$

Where the a_{ij}, b_i and c_j (i = 1, 2, ..., m, j = 1, 2, ..., m) are given constants and m < n. We shall assume that the equations (2.3) have been multiplied by (-1) where necessary to make all $b_i \geq 0$. The function (2.1) is called objective function and system (2.2) and (2.3) are called constraints.

The general L. P. P. is also denoted by : Minimize $z = \sum_{j=1}^{n} c_jx_j$

subject to $x_j \geq 0$, j = 1, 2, ..., n and

ai j xj bi

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \quad (i = 1, 2, ..., m)$$

Definition 2.1

A feasible solution to the L. P. P. is a vector $\bar{x} = (x_1, x_2, \ldots, x_n)$ which satisfies the conditions (2.2) and (2.3).

Definition 2.2

A basic solution (BS) to (2.3) (or L. P. problem) is a solution obtained by setting any n - m variables equal to zero and solving for the remaining m variables, provided that the determinant of the coefficients of these m variables is non zero. The m variables are called the basic variables.

Definition 2.3

A basic feasible solution (BSF) is a basic solution in which all the basic variables are non negative.

Definition 2.4

A non degenerate basic feasible solution is a basic feasible solution in which all the basic variables are positive.

Definition 2.5

A feasible solution which either maximizes or minimizes the objective function is called an optimal feasible solution.

Theorem 2.1

The collection of all felsible solutions to the L. P. P. is a convex set.

Proof

Let F be the set of all felsible solutions to the system $A\bar{x} = \bar{b}, \bar{x} \geq \bar{0}$

If the set F has only one point then obviously F is a convex set. Assume that F has more than one point.

Let $\bar{x}_1, \bar{x}_2 \in F$. Then we have

$A\bar{x}_1 = \bar{b}, \bar{x}_1 \geq \bar{0}$ and $A\bar{x}_2 = \bar{b}, \bar{x}_2 \geq \bar{0}$

Let $\bar{x}_0 = \lambda\bar{x}_1 + (1-\lambda)\bar{x}_2$ where $\bar{x}_1, \bar{x}_2 \in F, 0 \leq \lambda \leq 1$.

Then $A\bar{x}_0 = A\left[\lambda\bar{x}_1 + (1-\lambda)\bar{x}_2\right]$

$= \lambda A\bar{x}_1 + (1-\lambda)A\bar{x}_2$,

$= \lambda\bar{b} + (1-\lambda)\bar{b} = \bar{b}$

Also since $0 \leq \lambda \leq 1, \bar{x}_1 \geq \bar{0}, \bar{x}_2 \geq \bar{0}$ it follows that $\bar{x}_0 \geq \bar{0}$. This shows that $\bar{x}_0 \in F$ and consequently F is a convex set.

Remark

In general the convex set F is either (i) empty (ii) Unbounded or (iii) closed.

The empty set occurs when the constraints of the set can not be satisfied simultaneously. In this case the system yields no solution.

An unbounded set implies that the region of fisible solutions is not constrained in atleast one direction.

Finally closed set implies that the region of fessible solutions is a convex polyhedron since it is defined by the intersection of a finite number of linear constraints.

Note : We shall rewrite the definition of basic solution.

Basic Solution

Consider a system of simultaneous linear equations in n unknowns $A\bar{x}=\bar{b}\,(m<n)$, $r(A)=m$. If any n - m variables are equated to zero then the solution of the resulting system for m variables provided the determinant of the coefficient matrix of these variables is $\neq 0$ is called a basic solution, where r (A) = rank of A.

OR

If any $m \times m$ non singular matrix is chosen from A and if all the remaining n - m variables not associated with the columns in this matrix are set equal to 0 the solution to the resulting system of equations is called a basic solution. The m variables which can be different from zero are called basic variables.

Theorem 2.2

A necessary and sufficient condition for a point $\bar{x}\geq\bar{0}$ in F to be an extreme point is that $\bar{x}$ is a basic feasible solution to the system $A\bar{x}=\bar{b}, \bar{x}\geq\bar{0}$.

OR

Every basic feasible solution of $A\bar{x}=\bar{b}$ is an extreme point of the convex set of feasible solutions (of $A\bar{x}=\bar{b}$) and conversely every extreme point of the convex set of feasible solutions is a basic feasible solution to $A\bar{x}=\bar{b}$.

Proof

Let F denote the set of feasible solutions of $A\bar{x}=\bar{b}$.

Let $\bar{x}$ be a basic feasible solution of $A\bar{x}=\bar{b}$ which is a n - component vector $(x_1, x_2, ..., x_n)$. Thus both non basic (zero) and basic (some of which may be zero) variables are contains in $\bar{x}$. Suppose the components of $\bar{x}$ are so arranged that the first m components are the basic variables corresponding to basic vectors and are denoted by $\bar{x}_B$ Then,

$\bar{x}=(\bar{x}_B, \bar{o})$ where $\bar{o}$ is an (n - m) component null vector. Also assume that the vectors of the matrix A are so arranged that the first m column vectors correspond to $\bar{x}_B$ and we denote this sub matrix of A by B (called the basic matrix) and we denote the remaining (n - m) column vectors by R. Thus A = (B, R).

Accordingly the system $A\bar{x}=\bar{b}$ becomes

(B, R) $(\bar{x}_B, \bar{o})=\bar{b}$ or $B\bar{x}_B=\bar{b}$.

By the definition of a basic solution B must be non singular.

Hence $\bar{x}_B=B^{-1}\bar{b}$

To prove that every basic feasible solution is an extreme point of the convex set of feasible solutions.

If possible assume that the two distinct feasible solution $\bar{x}_1$ and $\bar{x}_2$ exist such that

$$\bar{x}=\lambda\bar{x}_1+(1-\lambda)\bar{x}_2,\ 0<\lambda<1 \qquad \text{.......... (1)}$$

But $\bar{x}_1$ and $\bar{x}_2$ can be expressed as,

$$\bar{x}_1=\left[\bar{x}_B^{(1)}, \bar{u}_1\right], \bar{x}_2=\left[\bar{x}_B^{(2)}, \bar{u}_2\right] \qquad \text{.......... (2)}$$

where $\bar{x}_B^{(1)}$ and $\bar{x}_B^{(2)}$ are the first m components of $\bar{x}_1$ and $\bar{x}_2$ respectively and $\bar{u}_1, \bar{u}_2$ denote the last (n - m) component vectors of $\bar{x}_1$ and $\bar{x}_2$ respectively.

From (1) and (2)

$$[\bar{x}_B, \bar{o}]=\lambda\left[\bar{x}_B^{(1)}, \bar{u}_1\right]+(1-\lambda)\left[\bar{x}_B^{(2)}, \bar{u}_2\right] \qquad \text{.......... (3)}$$

i. e.

$$[\bar{x}_B, \bar{o}]=\left[\lambda\bar{x}_B^{(1)}+(1-\lambda)\bar{x}_B^{(2)}, \lambda\bar{u}_1+(1-\lambda)\bar{u}_2\right]$$

Therefore $\lambda\bar{u}_1+(1-\lambda)\bar{u}_2=\bar{0}$ .... (4)

Since $\lambda>0$, $(1-\lambda)>0$ and $\bar{u}_1\geq\bar{0}, \bar{u}_2\geq\bar{0}$, therefore from (4)

$$\bar{u}_1=\bar{u}_2=\bar{0} \qquad \text{.......... (5)}$$

Since $\bar{x}_1, \bar{x}_2$ are in the set of feasible solutions,

$$A\bar{x}_1=\bar{b}, A\bar{x}_2=\bar{b}\Rightarrow B\bar{x}_B^{(1)}=\bar{b} \text{ and } B\bar{x}_B^{(2)}=\bar{b}$$

$$\Rightarrow \bar{x}_B^{(1)}=x_B^{(2)}=B^{-1}\bar{b}=\bar{x}_B$$

This shows that $\bar{x}=\bar{x}_1=\bar{x}_2$ which contradicts the fact that $\bar{x}_1\neq\bar{x}_2$. Conseqnently $\bar{x}$ cannot be expressed as a convex combination of any two distinct points in the set of feasible solutions and hence it must be an extreme point.

Conversely

Let $\bar{x}=(x_1, x_2, \ldots, x_n)$ be an extreme point of the convex set of feasible solutions.

We prove that $\bar{x}$ is a basic feasible solution of $A\bar{x}=\bar{b}$. By definition $\bar{x}$ will be a basic feasible solution of $A\bar{x}=\bar{b}$ if the column vectors associate with positive elements of $\bar{x}$ are linearly independent.

Assume that k - components of $\bar{x}$ are positives (remaining are zeros). Arrange the variables so that the first k components are positive. Then

$$\sum_{j=1}^{k} x_j \bar{a}_j = \bar{b}, x_j > 0, j = 1, 2, ..., k \qquad .... (6)$$

If possible assume that the vectors $\bar{a}_1, \bar{a}_2, ..., \bar{a}_k$ are not linearly independent. So they are linearly dependent and hence there exist scalars λ_j not all zero such that

$$\lambda_1 \bar{a}_1 + \lambda_2 \bar{a}_2 + ... + \lambda_k \bar{a}_k = \bar{0}$$

or $$\sum_{j=1}^{k} \lambda_j \bar{a}_j = \bar{0} \qquad .... (7)$$

From (6) and (7) it follows that for any $\delta > 0$,

$$\sum_{j=1}^{k} x_j \bar{a}_j \pm \delta \sum_{j=1}^{k} \lambda_j \bar{a}_j = \bar{b}$$

or $$\sum_{j=1}^{k} (x_j \pm \delta \lambda_j) \bar{a}_j = \bar{b}$$

Thus the two points

$$\bar{x}_1^* = (x_1 + \delta\lambda_1, x_2 + \delta\lambda_2, ..., x_k + \delta\lambda_k, \underbrace{0, 0, ..., 0}_{(n - k) \text{ components}}) \qquad .... (8)$$

and $$\bar{x}_2^* = (x_1 - \delta\lambda_1, x_2 - \delta\lambda_2, ..., x_k - \delta\lambda_k, \underbrace{0, 0, ..., 0}_{(n - k) \text{ components}}) \qquad .... (9)$$

satisfy the constraints $A\bar{x}=\bar{b}$

Since $x_j > 0$ select δ such that $0 < \delta < \min\left\{ \frac{x_j}{|\lambda_j|} \middle| \lambda_j \neq 0 \right\}$

Then the first k components of $\bar{x}_1^*, \bar{x}_2^*$ will always be positive.

Since the remaining components of $\bar{x}_1^*$ and $\bar{x}_2^*$ are zeros, it follows that $\bar{x}_1^*$ and $\bar{x}_2^*$ are feasible solutions different from $\bar{x}$. Adding (8) and (9) we obtain.

$$\bar{x}_1^* + \bar{x}_2^* = 2(x_1, x_2, ..., x_k, 0, 0, ..., 0)$$

$$\Rightarrow \frac{1}{2}\bar{x}_1^* + \frac{1}{2}\bar{x}_2^* = (x_1, x_2, ..., x_k, 0, 0, ..., 0) = \bar{x}$$

Thus $\bar{x}$ can be expressed as a convex combination of two distinct points $\bar{x}_1^*$ and $\bar{x}_2^*$ by selecting $\lambda = \frac{1}{2}$

i. e. $\bar{x} = \frac{1}{2}\bar{x}_1^* + \left(1 - \frac{1}{2}\right)\bar{x}_2^*$

This contradicts the assumption that $\bar{x}$ is an extreme point of the convex set of feasible solutions.

Hence $\bar{a}_1, \bar{a}_2, ..., \bar{a}_k$ are linearly independent and hence $\bar{x}$ is a basic feasible solution.

We have obviously $k \le m$. Because the number of linearly independent column vectors cannot be greater than m which is the row rank = column rank = rank of a matrix A. If k = m then the basic feasible solution is a non degenerate basic feasible solution.

Suppose k < m. Then the basic feasible solution is a degenerate basic feasible solution. Select other (m - k) additional column vectors with their corresponding variables equation 0. such that $\bar{a}_1, \bar{a}_2, ..., \bar{a}_m$ are linearly independent.

Thus the resulting set of k + (m - k) = m column vectors is linearly independent.

The sub matrix of A formed by these m columns is non singular.

Theorem 2.3

If the convex set of the feasible solutions of $A\bar{x} = \bar{b}$, is a convex polyhedron then at least one of the extreme points of the convex set of feasible solutions gives an optimal solution.

If the optimal solution occurs at more than one extreme point the value of the objective function will be the same for all convex combinations of these extreme points.

Proof

Let $\bar{x}_1, \bar{x}_2, ..., \bar{x}_k$ be the extreme points of the convex set F of the feasible solutions of the L. P. problem, $\max z = \bar{c} \cdot \bar{x}$ subject to $A\bar{x} = \bar{b}, \bar{x} \ge \bar{0}$.

Suppose $\bar{x}_m$ is the extreme point among $\bar{x}_1, \bar{x}_2, ..., \bar{x}_k$ at which the value of the objective function is maximum say z^*.

i. e. $$z^* = \max_{1 \le i \le k} \bar{c} \cdot \bar{x}_i = \bar{c} \cdot \bar{x}_m$$

Let $\bar{x}_0 \in F$ which is not an extreme point and let z_0 be the corresponding value of the objective function.

Then $$z_0 = \bar{c} \cdot \bar{x}_0 \qquad \text{........... (1)}$$

Since $\bar{x}_0$ is not an extreme point it can be expressed as convex combination of the extreme points $\bar{x}_1, \bar{x}_2, ..., \bar{x}_k$ of F (where F is assumed to be bounded).

Then $\bar{x}_0 = \lambda_1 \bar{x}_1 + \lambda_2 \bar{x}_2 + ... + \lambda_k \bar{x}_k$

where $\lambda_1, \lambda_2, ..., \lambda_k \geq 0$ and $\sum_{i=1}^{k} \lambda_i = 1$

So from (1) $$z_0 = \bar{c} \cdot (\lambda_1 \bar{x}_1 + \lambda_2 \bar{x}_2 + ... + \lambda_k \bar{x}_k)$$

$$\Rightarrow z_0 \leq \bar{c} \cdot \lambda_1 \bar{x}_m + \bar{c} \cdot \lambda_2 \bar{x}_m + ... + \bar{c} \cdot \lambda_k \bar{x}_m$$

$$\Rightarrow z_0 \leq \bar{c} \cdot (\lambda_1 + ... + \lambda_m) \bar{x}_m = \bar{c} \cdot \bar{x}_m$$

i. e. $z_0 \leq z^*$

This implies that the value of the objective function at any point in the set of fessible solutions is less than or equal to the maximal value z^* at extreme points.

Let $\bar{x}_1, \bar{x}_2, ..., \bar{x}_r$ $(r \leq k)$ be the extreme points of the set F at which the objective function assumes the same optimum value. This means.

$$z^* = \bar{c} \cdot \bar{x}_1 = \bar{c} \cdot \bar{x}_2 = ... = \bar{c} \cdot \bar{x}_r$$

Further let $\bar{x} = \lambda_1 \bar{x}_1 + \lambda_2 \bar{x}_2 + ... + \lambda_r \bar{x}_r, \lambda_j \geq 0$ and $\sum_{j=1}^{r} \lambda_j = 1$ be convex combination of there extreme points.

Then $$\bar{c} \cdot \bar{x} = \bar{c} \cdot [\lambda_1 \bar{x}_1 + \lambda_2 \bar{x}_2 + ... + \lambda_r \bar{x}_r]$$

$$= \lambda_1 (\bar{c} \cdot \bar{x}_1) + \lambda_2 (\bar{c} \cdot \bar{x}_2) + ... + \lambda_r (\bar{c} \cdot \bar{x}_r) = \lambda_1 z^* + ... + \lambda_r z^*$$

$$= (\lambda_1 + \lambda_2 + ... + \lambda_r) z^*$$

$= z^*$ Thus $\bar{c}\,\bar{x} = z^*$

This proves the result.

Note

Consider the general L. P. P.

Max. $z = \bar{c}\,\bar{x}$ subjects to $A\bar{x} = \bar{b}, \bar{x} \geq 0$ where

$$A = \begin{bmatrix} a_{11}.........a_{1n} \\ a_{21}.........a_{2n} \\ .... \\ a_{m1}.........a_{mn} \end{bmatrix}$$

$$\bar{c} = (c_1, c_2, ..., c_n)$$

$$\bar{x} = (x_1, ..., x_n),\ \bar{b} = (b_1, b_2, ..., b_m)$$

Where rank of A i. e. r (A) = m < n.

For convenience column vectors will also be represented by row vectors without using the transpose symbol (T). So there should be no confusion in understanding the scalar multiplication of two vectors $\bar{c}$ and $\bar{x}$.

We shall denote the j^{th} column of A by $\bar{a}_j, j = 1, 2, ..., n$

so that $$A = [\bar{a}_1, \bar{a}_2, ..., \bar{a}_n] \qquad .... (1)$$

Form an $m \times m$ non singular submatrix B of A called the basic matrix, whose columns are linearly independents vectors. Let these column vectors be renamed as

$\beta_1, \beta_2, ..., \beta_m$. Therefore

$$B = [\beta_1, \beta_2, ..., \beta_m] \qquad .... (2)$$

These columns of B form a basic of R^m.

Now any column $\bar{a}_j$ of A can be expressed as a linear combination of the columns of B.

Let $$\bar{a}_j = y_{1j}\,\beta_1 + y_{2j}\,\beta_2 + ... + y_{mj}\,\beta_m$$

$$\bar{a}_j = (\beta_1, \beta_2, ..., \beta_m) \cdot (y_{1j}, y_{2j}, ..., y_{mj})$$

i. e. $\bar{a}_j = B\,\bar{y}_j$ where $\bar{y}_j = (y_{ij}, y_{2j}, ..., y_{mj})$

i. e. $\bar{a}_j = B\,\bar{y}_j$ where $\bar{y}_j = (y_{1j}, y_{2j}, ..., y_{mj})$

i. e. $\bar{y}_j = B^{-1}\,\bar{a}_j$ where $y_{ij}\,(i = 1, ..., m)$ are scalars.

The vector $\bar{y}_j$ will change if the columns of A forming B change. Any basic matrix B will yield a basic solution to $A\,\bar{x} = \bar{b}$. The solution may be denoted by m component vector as $\bar{x}_B = (x_{B1}, x_{B2,}, ..., x_{Bm})$ where $\bar{x}_B$ is determined from $\bar{x}_B = B^{-1}\,\bar{b}$. .... (4)

Note that x_{Bi} corresponds to the column β_i of the matrix B. The variables $x_{B1}, x_{B2}, \ldots, x_{Bm}$ are called basic variables and the remaining (n - m) variables are non basic variables.

Correspondings to $\bar{x}_B$ we have $z = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n$

$$\text{Let } \bar{c}_B = \left(c_{B1}, c_{B_2}, \ldots, c_{Bm}\right)$$

where c_{Bi} is the coefficient of the basic variable x_{Bi} in the objective function.

So
$$z = c_{B1} x_{B1} + c_{B2} x_{B2} + \ldots + c_{Bm} x_{Bm} + \bar{0}$$

$$z = \left(c_{B1}, \ldots, c_{Bm}\right)\left(x_{B1}, \ldots, x_{Bm}\right)$$

$$z = \bar{c}_B \, \bar{x}_B \qquad \text{........... (5)}$$

Finally we form a new variable z_j defined as

$$z_j = y_{1j} c_{B1} + y_{2j} c_{B2} + \ldots + y_{mj} c_{Bm} = \sum_{i=1}^{m} c_{Bi} \, y_{ij}$$

$$z_j = \left(c_{B1}, \ldots, c_{Bm}\right)\left(y_{ij}, y_{2j}, \ldots, y_{mj}\right)$$

$$z_j = \bar{c}_B \, \bar{y}_j$$

There exists z_j for each $\bar{a}_j$.

Example 2.1

Illustrate the above definitions and notations for the following L. P. problem.

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

subject to $4x_1 + 2x_2 + x_3 + x_4 = 4$

$$x_1 + 2x_2 + 3x_3 - x_5 = 8$$

Solution :

Constraints equations in matrix form may be written as

$$\begin{array}{ccc} \bar{a}_1 \; \bar{a}_2 \; \bar{a}_3 \; \bar{a}_4 \; \bar{a}_5 & x & \bar{b} \\ \begin{bmatrix} 4 & 2 & 1 & 1 & 0 \\ 1 & 2 & 3 & 0 & -1 \end{bmatrix} & \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = & \begin{bmatrix} 4 \\ 8 \end{bmatrix} \end{array}$$

or $$A\bar{x}=\bar{b}$$

A basis matrix $B=(\beta_1,\beta_2)$ is formed using columns $\bar{a}_3$ and $\bar{a}_1$ where

$$\beta_1=\bar{a}_3=\begin{bmatrix}1\\3\end{bmatrix},\ \beta_2=\bar{a}_1=\begin{bmatrix}4\\1\end{bmatrix}$$

The rank of the matrix A is 2 and column vectors $\bar{a}_3,\bar{a}_1$ are linearly independent and thus form a basis for R^2. Thus basis matrix is

$$B=(\beta_1,\beta_2)=\begin{bmatrix}1 & 4\\3 & 1\end{bmatrix}$$
$$\quad\quad\quad\quad\quad\quad\bar{a}_3\ \ \bar{a}_1$$

Then the basic feasible solution is $\bar{x}_B=\bar{B}^{-1}\bar{b}$

$$\bar{x}_B=\left(\frac{+1}{|B|}\text{adj.}B\right)\bar{b}$$

$$\bar{x}_B=\frac{-1}{11}\begin{bmatrix}1 & -4\\-3 & 1\end{bmatrix}\begin{bmatrix}4\\8\end{bmatrix}=\frac{1}{11}\begin{bmatrix}28\\4\end{bmatrix}$$

$$\bar{x}_B=\begin{bmatrix}\frac{28}{11}\\\frac{4}{11}\end{bmatrix}=\begin{bmatrix}x_{B1}\\x_{B2}\end{bmatrix}$$

Hence the basic solution is $x_{B1}=\frac{28}{11}=x_3$, $x_{B2}=\frac{4}{11}=x_1$ and the remaining non basic variables are (always) zero i. e. $x_2=x_4=x_5=0$.

Also $\quad c_{B1}=\text{coeff. of } x_{B1}=\text{coeff. of } x_3=c_3=3$

$\quad\quad c_{B2}=\text{coeff. of } x_{B2}=\text{coeff. of } x_1=c_1=1$

Hence the value of the objective function is

$$z=\bar{c}_B\,\bar{x}_B=(3,1)\begin{pmatrix}28/11\\4/11\end{pmatrix}=\frac{88}{11}$$

Also any vector $\bar{a}_j=(j=1,2,3,4,5)$ can be expressed as a linear combination of vectors $\beta_j\,(j=1,2)$.

Let $$\bar{a}_j = y_{1j}\beta_1 + y_{2j}\beta_2 = y_{1j}\bar{a}_3 + y_{2j}\bar{a}_1$$

$$\bar{y}_2 = \bar{B}^{-1}\bar{a}_2 = -\frac{1}{11}\begin{bmatrix} 1 & -4 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 6/11 \\ 4/11 \end{bmatrix} = \begin{bmatrix} y_{12} \\ y_{22} \end{bmatrix}$$

Hence $$y_{12} = \frac{6}{11} \text{ and } 1\, y_{22} = \frac{4}{11}.$$

Now the variable z_2 corresponding to the column vector $\bar{a}_2$ can be obtained as

$$z_2 = \bar{c}_B\, \bar{y}_2 = (3,1)\begin{pmatrix} 6/11 \\ 4/11 \end{pmatrix}$$

$$= \left[3.\frac{6}{11} + 1.\frac{4}{11}\right] = \frac{22}{11} = 2$$

Similarly z_1, z_3, z_4 and z_5 can also be obtained.

Theorem 2.4

Consider a set of m simultaneous linear equations in n unknowns with $n > m, A\bar{x} = \bar{b}$ and $r(A) = m$. Then if there is a feasible solution $\bar{x} \geq \bar{0}$, there is a basic feasible solution.

Proof

To prove this assume that there exists a feasible solution to $A\bar{x} = \bar{b}$ with $p \leq n$ positive variables.

Number the variables, so that the first p variables are positive. Then the feasible solution can be written as

$$\sum_{j=1}^{n} x_j\, \bar{a}_j = \bar{b} \qquad \text{........... (1)}$$

and hence

$$x_j > 0, (j = 1, 2, \ldots., p), x_j = 0, (j = p+1, p+2, \ldots, n) \qquad \text{........... (2)}$$

Case (i)

Suppose the set $\bar{a}_j\,(j = 1, 2, \ldots, p)$ is linearly independent. Then $p \leq m$.

If p = m the given solution is automatically a nondegenerate basic feasible solution.

Supposep < m. We know that this set of p linearly independent column vectors can be extended to form a base $\{\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_m\}$ of the column space of A.

In this case $\{x_1, x_2, \ldots, x_p, x_{p+1}, \ldots, x_m\}$ where $x_j = 0, j = p+1, p+2, \ldots, m$ is a degenerate basic fessible solution.

Case (ii)

Suppose the vectors $\bar{a}_j\,(j=1,2,...,p)$ are linearly dependent. We shall show that under these circumstances it is possible to reduce the number of positive variables step by step until the columns associated with the positive variables are linearly independent.

When the $\bar{a}_j\,(j=1,2,...,p)$ are linearly dependent, there exist α_j not all zero such that

$$\sum_{j=1}^{p}\alpha_j\,\bar{a}_j=\bar{0} \qquad \text{.......... (3)}$$

and we proceed to reduce some x_r in

$$\sum_{j=1}^{p}x_j\,\bar{a}_j=\bar{b},\ x_j>0\,(j=1,2,...,p) \qquad \text{.......... (4)}$$

to zero.

Suppose some vector $\bar{a}_r$ of the p vectors in $\sum_{j=1}^{p}\alpha_j\,\bar{a}_j=\bar{0}$ is expressed in terms of the remaining (p - 1) vectors.

Thus $$\bar{a}_r=-\sum_{j\neq r}\frac{\alpha_j}{\alpha_r}\bar{a}_j \qquad \text{.......... (5)}$$

substituing (5) in (4) we obtain

$$\sum_{\substack{j=1\\ j\neq r}}^{p}\left(x_j-x_r\frac{\alpha_j}{\alpha_r}\right)\bar{a}_j=\bar{b} \qquad \text{.......... (6)}$$

Here we have not more than (p - 1) variables. However we are not sure that all these variables are non negative (In general if we choose $\bar{a}_r$ orbitrarily some variables may be negative)

We wish to obtain

$$x_j-x_r\frac{\alpha_j}{\alpha_r}\geq 0\ \ (j = 1, 2, ..., p),\ \ j\neq r \qquad \text{.......... (7)}$$

For any j for which $\alpha_j=0$ (7) will be satisfied automatically. When $\alpha_j\neq 0$ we have,

$$\frac{x_j}{\alpha_j}-\frac{x_r}{\alpha_r}\geq 0\ \text{ if }\ \alpha_j>0 \qquad \text{.......... (8)}$$

$$\frac{x_j}{\alpha_j} - \frac{x_r}{\alpha_r} \le 0 \text{ if } \alpha_j < 0 \qquad \text{.......... (9)}$$

We select $\bar{a}_r$ such that

$$\frac{x_r}{\alpha_r} = \min_j \left\{ \frac{x_j}{\alpha_j} \middle| \alpha_j > 0 \right\} \qquad \text{........ (10)}$$

(Note that $\sum \alpha_j \bar{a}_j = \bar{0} \Rightarrow$ at least one $\alpha_j \neq 0$ and hence $\alpha_j > 0$ for some j)

Thus a fessible solution $\sum_{\substack{j=1 \\ j \neq r}}^{p} \left(x_j - x_r \frac{\alpha_j}{\alpha_r} \right) \bar{a}_j = \bar{b}$

is obtained with not more than (p - 1) non zero variables.

These variables are also non negative. (since $\alpha_j > 0$)

If the columns associated with the positive variabls are linearly independent by case (i) we have a basic feasible solution. If the columns associates with the positive variables are linearly dependent we can repeat the same procedure and reduce one of the positive variables to 0. Utimately we shall arrive at a solution such that the columns corresponding to the positive variables are linearly independent. (Note that a single non zero vector is always linearly independent)

OR

Theorem 2.5

If a linear programming problem

max. $z = \bar{c}\,\bar{x}$ s. t. $A\,\bar{x} = \bar{b}, \bar{x} \ge \bar{0}$

has at least one feasible solution then it has at least one basic feasible solution.

Proof

Let

$$\bar{x}_0 = (x_1, x_2, \ldots, x_k, 0, 0, \ldots, 0)$$

be a feasible solution to the L. P. P. with positive components $x_1, x_2, \ldots, x_k$.

Let $\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_k$ be the first k columns of A (associated with the positive variables $x_1, x_2, \ldots x_x$ respectively)

Then by hypothesis

$$x_1\,\bar{a}_1 + x_2\,\bar{a}_2 + \ldots + x_k\,\bar{a}_k = \bar{b} \qquad \text{.......... (1)}$$

Case (i)

Suppose $\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_k$ are linearly indepedent. In this case $\bar{x}_0 = (x_1, x_2, x_k, 0, \ldots, 0)$ is a basic fessible solution.

Case (ii)

Suppose $\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_k$ are linearly dependent.

So there exist scalers $\lambda_1, \ldots, \lambda_k$ not all 0 such that

$\lambda_1 \bar{a}_1 + \ldots + \lambda_k \bar{a}_k = \bar{0}$ with atleast one $\lambda_j \neq 0$ and hence assume this $\lambda_j > 0$. .... (2)

Let $v = \max\limits_{1 \le j \le k} \left\{ \frac{\lambda_j}{x_j} \right\}, \lambda_j > 0$ (i.e. m +x is taken over those j fro which $\lambda_j > x$)

Obviously v > 0 for $x_j > 0$ (j = 1, 2, ..., k) and at least one $\lambda_j > 0$.

Multiply (2) by $\frac{1}{v}$ and then subtract from (1) to get

$$\sum_{j=1}^{k} x_j \bar{a}_j - \frac{1}{v} \sum_{j=1}^{k} \lambda_j \bar{a}_j = \bar{b}$$

$$\Rightarrow \sum_{j=1}^{k} \left(x_j - \frac{\lambda_j}{v} \right) \bar{a}_j = \bar{b} \qquad \text{.......... (3)}$$

$$\Rightarrow \hat{x} = \left(x_1 - \frac{\lambda_1}{v}, x_2 - \frac{\lambda_2}{v}, \ldots, x_k - \frac{\lambda_k}{v}, 0, 0, \ldots, 0 \right)$$

is a new solution of $A\bar{x} = \bar{b}$.

We have $v \ge \frac{\lambda_j}{x_j}$ or $x_j \ge \frac{\lambda_j}{v} (1 \le j \le k)$

The new solution $\hat{x}$ satisfies non negativity restriction.

Since $x_j - \frac{\lambda_j}{v} = 0$ for at least one j, $\hat{x}$ is a feasible solution with at the most k - 1 positive variables. All other variables are 0.

If the columns associated with the positive variables are still linearly. dependent, repeat the above procedure. Cuntinuing in this way we get the column vectors ossociated with positive variables which are linearly independent. Thus by case (i) we get a basic feasible solution.

Example 2.2

If $x_1 = 2, x_2 = 3, x_3 = 1$ is a feasible solution of a L. P. P. problem

max. $z = x_1 + 2x_2 + 4x_3$

subject to $2x_1 + x_2 + 4x_3 = 11$

$3x_1 + x_2 + 5x_3 = 14$

$x_1, x_2, x_3 \geq 0$

find a Basic Feasible Solution

Solution :

We have $A\bar{x} = \bar{b}$

where $A = \begin{bmatrix} 2 & 1 & 4 \\ 3 & 1 & 5 \end{bmatrix}, \bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \bar{b} = \begin{bmatrix} 11 \\ 14 \end{bmatrix}$

The given feasible solution is $x_1 = 2, x_2 = 3, x_3 = 1.$

Hence $2\bar{a}_1 + 3\bar{a}_2 + 1.\bar{a}_3 = \bar{b}$

Where $\bar{a}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \bar{a}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \bar{a}_3 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \bar{b} = \begin{bmatrix} 11 \\ 14 \end{bmatrix}$

Step (2)

The vectors $\bar{a}_1, \bar{a}_2, \bar{a}_3$ associated wsith the positive variables x_1, x_2, x_3 are linearly dependent so one of the vectors is a linear combination of the remaining two.

Let $\bar{a}_3 = \lambda_1 \bar{a}_1 + \lambda_2 \bar{a}_2$ Thus

$$\begin{bmatrix} 4 \\ 5 \end{bmatrix} = \lambda_1 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Maximum no. of lin. independnet columns is less than 3 since row rank of coefficient matrix A is 2.

$$\text{Now } \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 2\lambda_1 + \lambda_2 \\ 3\lambda_1 + \lambda_2 \end{bmatrix}$$

$$\Rightarrow 2\lambda_1 + \lambda_2 = 4, 3\lambda_1 + \lambda_2 = 5$$

$$\Rightarrow \lambda_1 = 1, \lambda_2 = 2$$

$$\Rightarrow \bar{a}_3 = \bar{a}_1 + 2\bar{a}_2$$

i. e. $\bar{a}_1 + 2\bar{a}_2 - \bar{a}_3 = \bar{0}$

Where $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = -1$

Step (3)

Now determine which of the variables x_1, x_2, x_3 should be 0. For this find

$$v = \max\left(\frac{\lambda_j}{x_j}\right), \lambda_j > 0$$

$$= \max\left(\frac{\lambda_1}{x_1}, \frac{\lambda_2}{x_2}\right) \qquad \text{(since } \lambda_1 = 1 > 0 \text{ , } \lambda_2 = 2 > 0 \text{)}$$

$$= \max\left\{\frac{1}{2}, \frac{2}{3}\right\} = \frac{2}{3}$$

$\hat{x} = \left(x_1 - \frac{\lambda_1}{v}, x_2 - \frac{\lambda_2}{v}, x_3 - \frac{\lambda_3}{v}\right)$ is a reduced solution where

$$x_1 - \frac{\lambda_1}{v} = 2 - \frac{1}{2/3} = \frac{1}{2}$$

$$x_2 - \frac{\lambda_2}{v} = 3 - \frac{2}{2/3} = 0$$

$$x_3 - \frac{\lambda_3}{v} = 1 - \left(-\frac{1}{2/9}\right) = \frac{5}{2}$$

$\therefore$ $$\hat{x} = \left(\frac{1}{2}, 0, \frac{5}{2}\right)$$

Step (4)

Now the solution $\hat{x} = \left(\frac{1}{2}, 0, \frac{5}{2}\right)$ is to be tested for basicness. The determinant of the matrix of the of column vectors corresponding to x_1, x_3 is

$$\left|\begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix}\right| \neq 0$$

Obviously $\bar{a}_1, \bar{a}_3$ are linearly independent.

Hence $\hat{x} = \left(\frac{1}{2}, 0, \frac{5}{2}\right)$ is a B. F. S.

Theorem 2.6

Let a L. P. P. have a B. F. S. If for any column $\bar{a}_j$ in A but not in $B = \{\bar{b}_1, \bar{b}_2, ..., \bar{b}_m\}$ (basic vectors for columns in A) we have $\bar{a}_j = \sum_{i=1}^{m} y_{ij}\, \bar{b}_i$ with at least one $y_{ij} > 0$ (i = 1, 2, ..., m) then we can find a new B. F. S. by replacing one of the columns in B by $\bar{a}_j$.

Proof

Consider a L. P. P. problem max $z = \bar{c}\,\bar{x}$ subject to $A\,\bar{x} = \bar{b}, \bar{x} \geq \bar{0}$ where A is $m \times n$ matrix m < n and r (A) = m, where r (A) = rank of A.

Let $\bar{x}_B$ be a BFS of the LPP, where $B = \{\bar{b}_1, \bar{b}_2, ..., \bar{b}_m\}$ forms a basis for the columns of A.

For any column $\bar{a}_j$ in A $(\bar{a}_j \notin B)$, we have

$$\bar{a}_j = \sum_{i=1}^{m} y_{ij}\, \bar{b}_i$$

Suppose some $y_{r_j} > 0$

Then
$$\bar{a}_j = \sum_{\substack{i=1 \\ i \neq r}}^{m} y_{ij}\, \bar{b}_i + y_{rj}\, \bar{b}_r$$

$$\Rightarrow \bar{b}_r = \frac{\bar{a}_j}{y_{rj}} - \frac{1}{y_{rj}} \sum_{\substack{i=1 \\ i \neq r}}^{m} y_{ij}\, \bar{b}_i$$

Hence $B\,\bar{x}_B = \bar{b}$ gives $\bar{b} = \sum_{i=1}^{m} x_{Bi}\, \bar{b}_i$

$$\Rightarrow \bar{b} = \sum_{\substack{i=1 \\ i \neq r}}^{m} x_{Bi}\ \bar{b}_i + x_{Br}\left[\frac{\bar{a}_j}{y_{rj}} - \frac{1}{y_{rj}}\sum_{\substack{i=1 \\ i \neq r}}^{m} y_{ij}\,\bar{b}_i\right]$$

$$\Rightarrow \bar{b} = \sum_{\substack{i=1 \\ i \neq r}}^{m}\left[x_{Bi} - x_{Br}\frac{y_{ij}}{y_{rj}}\right]\bar{b}_i + \frac{x_{Br}}{y_{rj}}\bar{a}_j$$

The new solution $\hat{x}_B$ is also a basic solution with the basic variables.

$$\hat{x}_{Bi} = \left(x_{Bi} - x_{Br}\frac{y_{ij}}{y_{rj}}\right), i = 1, 2, \ldots, m,\ i \neq r$$

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

Case (1)

Let $x_{Br} = 0$

In this case the new set of basic variables is obviously non negative, since we have assumed the existance of a BFS, $\bar{x}_B$.

Case (2)

$x_{Br} \neq 0$

We have $y_{rj} > 0$

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

If $y_{ij} = 0$ for some i, $\hat{x}_{Bi} = x_{Bi} \geq 0, \hat{x}_{Br} \geq 0$

If $y_{ij} < 0$ still $\hat{x}_{Bi} \geq 0$ and $\hat{x}_{Br} \geq 0$.

Suppose $y_{ij} > 0$

We require $\hat{x}_{Bi} = x_{Bi} - x_{Br}\dfrac{y_{ij}}{y_{rj}} \geq 0, i \neq r$

So we must have $\dfrac{x_{Bi}}{y_{ij}} \geq \dfrac{x_{Br}}{y_{rj}}$, where $y_{ij} > 0$.

We select r in such a way that $\frac{x_{Br}}{y_{rj}} = \min\left\{ \frac{x_{Bi}}{y_{ij}} \middle| y_{ij} > 0 \right\}$

Then we have a B. F. S.

Example 2.3

Given a basic feasible solution $x_3 = 4$ and $x_4 = 8$ to the L. P. P.

max. $z = x_1 + 2x_2$ subject to

$$x_1 + 2x_2 + x_3 = 4$$

$$x_1 + 4x_2 + x_4 = 8 ,$$

obtain a new B. F. S.

Solution :

We have $A\bar{x} = \bar{b}$

Where $A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}, \bar{x} = (x_1 x_2, x_3, x_4), \bar{b} = (4, 8)$

$$\bar{a}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \bar{a}_2 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \bar{a}_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \bar{a}_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

We have $B\bar{x}_B = \bar{b}$ where $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$$\bar{x}_B = (x_{B1}\, x_{B2}) = (4, 8), x_{B1} = x_3 = 4, x_{B2} = x_4 = 8$$

$$\beta_1 = \bar{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \beta_2 = \bar{b}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The y_j s for any column $\bar{a}_j$ in A but not in B are

$$\bar{y}_1 = B^{-1}\bar{a}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} y_{11} \\ y_{21} \end{bmatrix}$$

$$\bar{y}_2 = B^{-1}\bar{a}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} y_{12} \\ y_{22} \end{bmatrix}$$

Note that $\bar{a}_1 = \bar{B}^1\bar{a}_1 = y_{11}\bar{b}_1 + y_{12}\bar{b}_2$ and

$\bar{a}_2 = \bar{B}^1\bar{a}_2 = y_{21}\bar{b}_1 + y_{22}\bar{b}_2$.

Since $y_{11}=1, y_{21}=1>0$ we can insert $\bar{a}_1$ in B. We now select $\beta_r=\bar{b}_r$ for replacement by $\bar{a}_1$ which corresponds to the value of r determined by the minimum ratio rule :

$$\frac{x_{Br}}{y_{r1}} = \min_i \left\{ \frac{x_{Bi}}{y_{i1}}, y_{i1}>0 \right\}$$

$$= \min\left[\frac{x_{B1}}{y_{11}}, \frac{x_{B2}}{y_{21}}\right]$$

$$= \min\left[\frac{4}{1}, \frac{8}{1}\right] = 4 = \frac{x_{B1}}{y_{11}}$$

$$\Rightarrow r = 1$$

Hence we remove β_1 and enter $\bar{a}_1$ in place of $\beta_1=\bar{b}_1$.

The new basic matrix becomes

$$\hat{B} = (\bar{a}_1, \beta_2) \qquad \left(\text{or } \hat{\beta} = \left(\hat{\beta}_1, \hat{\beta}_2\right), \hat{\beta}_1 = \bar{a}, \hat{\beta}_2 = \beta_2 \right)$$

$$\Rightarrow \hat{B} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

We can now find the basic feasible solution $\hat{x}_B$ either by using the result $\hat{x}_B = \hat{B}^{-1}\bar{b}$ or by the transformation formulae.

$$\hat{x}_{Bi} = x_{Bi} - x_{Br}\frac{y_{ij}}{y_{rj}}, i = 1,\dots,m, i \neq r$$

and $\hat{x}_{Br} = \dfrac{x_{Br}}{y_{rj}}$ for i = r = 1, $x_i = \hat{x}_{B1}$

Now $\beta_1=\bar{b}_1$ is removed means x_3 will not be a basic feasible solution. In its plane x_4 corresponding to $\bar{a}_1$ will be a B. F. S. and $x_1 = x_{B1}$.

Using the formula

$$\hat{x}_{B1} = \frac{x_{B1}}{y_{11}} = \frac{x_3}{1} = \frac{4}{1} = 4$$

$$\hat{x}_{B2} = x_{B2} - x_{B1}\frac{y_{21}}{y_{11}} = x_4 - x_3\frac{y_{21}}{y_{11}} = 8 - 4\times\frac{1}{1} = 4$$

Hence the new B. F. S. is

$$x_1 = x_{B1} = 4, x_2 = 0, x_3 = 0, x_4 = x_{B2} = 4$$

Theorem 2.7

If a linear programming problem,

Max. $z = \bar{c}\,\bar{x}$, , s. t. $A\,\bar{x} = \bar{b}, \bar{x} \geq 0$,

has at least one optimal feasible solution, then at least one basic feasible solution must be optimal.

Proof

Let $$\bar{x}^0 = \left(x_1, x_2, ..., x_k \overbrace{0, 0, ..., 0}^{m+n-k} \right)$$

be an optimal feasible solution to the given linear programming problem which yields the optimum value

$$z^* = \sum_{j=1}^{k} c_j\, x_j \text{ . Also } \sum_{j=1}^{k} x_j\, \bar{a}_j = \bar{b} \qquad \text{.......... (1)}$$

If $\bar{a}_1, \bar{a}_2, ..., \bar{a}_k$ are linearly independent then $\bar{x}^0$ is an optimed BFS. Otherwise $\bar{a}_1, \bar{a}_2, ..., \bar{a}_k$ are linearly dependent and there exist λ_j, not all 0,

such that $\sum_{j=1}^{k} \lambda_j\, \bar{a}_j = \bar{0}$ where at least one $\lambda_j > 0$ .... (2)

Let $V = \max_{1\leq j\leq k}\left(\frac{\lambda_j}{x_j}\right)$ .... (3)

Obviously V > 0, because $x_j > 0$ and at least one $\lambda_j > 0\,(1 \leq j \leq k))$. 1

Now multiplying (2) by $\frac{1}{V}$ and subtracting from (1) we get

$$\sum_{j=1}^{k} x_j\, \bar{a}_j - \frac{1}{V}\sum_{j=1}^{k} \lambda_j\, \bar{a}_j = \bar{b}$$

$$\Rightarrow \sum_{j=1}^{k}\left(x_j - \frac{\lambda_j}{V}\right)\bar{a}_j = \bar{b} \qquad \text{.......... (4)}$$

$$\Rightarrow \hat{x} = \left(x_1 - \frac{\lambda_1}{v}, x_2 - \frac{\lambda_2}{v}, \ldots, x_k - \frac{\lambda_k}{v}, 0, 0, \ldots, 0\right)$$ is a new solution of $A\bar{x} = \bar{b}$.

From (3) $v \geq \frac{\lambda_j}{x_j} \Rightarrow x_j - \frac{\lambda_j}{v} \geq 0, j = 1, 2, \ldots, k$

Thus $\hat{x}$ is a feasible solution and since $x_j - \frac{\lambda_j}{v} = 0$ for at least one j, $\hat{x}$ contains at the most k - 1 non zero variables other variables being zero.

If the column vectors associated with the positive variables are still linearly dependent we repeat the above process and finally get the solution which is a BFS. So without loss of generality the solution $\hat{x}$ will be assumed as a basic feasible solution.

We have to prove that $\hat{x}$ is also optimum solution.

The value of the objective function corresponding to this solution $\hat{x}$ will become

$$\hat{z} = \sum_{j=1}^{k} c_j \left(x_j - \frac{\lambda_j}{\nu}\right) = \sum_{j=1}^{k} c_j x_j - \frac{1}{\nu}\sum_{j=1}^{k} c_j \lambda_j$$

or $$\hat{z} = z^* - \frac{1}{\nu}\sum_{j=1}^{k} c_j \lambda_j \qquad \text{.......... (5)}$$

(since $z^* = \sum_{j=1}^{k} c_j x_j$)

But, for optimality $\hat{z}$ must be equal to z^*. Hence $\hat{x}$ will be optimal solution if and only if we prove,

$\sum_{j=1}^{k} c_j \lambda_j = 0$ in equation (5).

We shall prove this by contradiction.

If possible, let us assume that

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

Then, there will be two possibilities :

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

2) $\sum_{j=1}^{k} c_j \lambda_j < 0$

Now, in either of these two cases we can find a real number, say r, such that

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

(in first case, r will be positive and in second case r will negative)

i. e. $\sum_{j=1}^{k} c_j (r \lambda_j) > 0$ .... (6)

Now adding $\sum_{j=1}^{k} c_j x_j$ to both sides on (6), we have

$$\sum_{j=1}^{k} c_j (r \lambda_j) + \sum_{j=1}^{k} c_j x_j > \sum_{j=1}^{k} c_j x_j$$

or $\sum_{j=1}^{k} c_j (x_j + r \lambda_j) > z^*$ .... (7)

Now, $\left(x_1 + r\lambda_1, x_2 + r\lambda_2, \ldots, x_k + r\lambda_k, \overbrace{0, 0, \ldots, 0}^{m+n-k} \right)$ is also a solution for any value of r which can be observed by multiplying equation (2) by r and adding to equation (1)

Furthermore, there exist an infinite number of choices of r for which the solution

$\left(x_1 + r\lambda_1, x_2 + r\lambda_2, \ldots, x_k + r\lambda_k, \overbrace{0, \ldots, 0}^{m+n-k} \right)$ satisfies the non - negativity restrictions also.

We now proceed to prove this statement. To satisfy the non - negativity restriction, we need

$$x_j + r\lambda_j \geq 0, j = 1, 2, \ldots, k$$

or $r\lambda_j \geq -x_j$

We have

$$\left.\begin{array}{r} r \geq -\dfrac{x_j}{\lambda_j}, \text{if } \lambda_j > 0 \\ r \leq -\dfrac{x_j}{\lambda_j}, \text{if } \lambda_j < 0 \\ r \text{ unrestricted, if } \lambda_j = 0 \end{array}\right\}$$

or

Thus, we observe that if we select r satisfying the relationship

$$\max_{\substack{j \\ (\lambda_j > 0)}} \left(-\frac{x_j}{\lambda_j} \right) \leq r \leq \min_{\substack{j \\ (\lambda_j < 0)}} \left(-\frac{x_j}{\lambda_j} \right) \qquad \text{.......... (8)}$$

then $x_j + r\lambda_j \geq 0$ for j = 1, 2, ..., k. We note that if there is no j for which $\lambda_j > 0$, then there is no lower limit for r and if there is no j for which $\lambda_j < 0$, then there is no upper limit for r.

Furthermore,

$$\max_{\substack{j \\ (\lambda_j > 0)}} \left(-\frac{x_j}{\lambda_j} \right) < 0 \text{ and } \min_{\substack{j \\ (\lambda_j < 0)}} \left(-\frac{x_j}{\lambda_j} \right) > 0$$

This proves that when r lies in the non - empty interval given by (8), then the infinite number of solutions.

$$\left(x_1 + r\lambda_1, x_2 + r\lambda_2, \ldots, x_k + r\lambda_k, \overbrace{0, 0, \ldots, 0}^{m+n-k} \right)$$

satisfy the non - negativity restrictions also.

Now, from (7) we conclude that the left hand side $\sum_{i=1}^{k} c_j \left(x_j + r\lambda_j \right)$ yields the value of the objective function which is strictly greater than the greatest value of the objective function. This contradiction proves that $\sum_{j=1}^{k} c_j \lambda_j = 0$ and hence $\hat{x}$ is optimal.

Note : By what we have proved we have the result :

If the linear programming problem :

Max. z = cx, subject to Ax = b, $x \geq 0$

has feasible solution, then it has at least one optimal basic feasible solutions.

Reduction of any feasible solution to a basic feasible solution

Example 2.4

If $x_1 = 2, x_2 = 3, x_3 = 1$, be a feasible solution of linear programming problem :

$$\text{Max. } z = x_1 + 2x_2 + 4x_3,$$

subject to $2x_1 + x_2 + 4x_3 = 11,$

$$3x_1 + x_2 + 5x_3 = 14,$$

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

then find a basic feasible solution.

Solution :

We express the above system as

$$\begin{matrix} \overline{a}_1 & \overline{a}_2 & \overline{a}_3 \end{matrix} \qquad\qquad \overline{b}$$

$$\begin{pmatrix} 2 & 1 & 4 \\ 3 & 1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 11 \\ 14 \end{pmatrix}$$

or $\quad x_1 \overline{a}_1 + x_2 \overline{a}_2 + x_3 \overline{a}_3 = \overline{b}$

But the given feasible solution is $x_1 = 2, x_2 = 3, x_3 = 1$. Hence $2\overline{a}_1 + 3\overline{a}_2 + 1\overline{a}_3 = \overline{b}$

Where $\overline{a}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \overline{a}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \overline{a}_3 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \overline{b} = \begin{bmatrix} 11 \\ 14 \end{bmatrix}$

Since the vectors $\overline{a}_1, \overline{a}_2, \overline{a}_3$ associated with the corresponding varialbes x_1, x_2, x_3 are linearly dependent, therefore one of the vectors can be expressed in terms of the remaining two.

Thus,

$\overline{a}_3 = \lambda_1 \overline{a}_1 + \lambda_2 \overline{a}_2$. So $\lambda_1 \overline{a}_1 + \lambda_2 \overline{a}_2 + \lambda_3 \overline{a}_3 = 0$, where $\lambda_3 = -1$ .... (1)

or $\quad \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \lambda_1 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

or $\quad \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 2\lambda_1 + \lambda_2 \\ 3\lambda_1 + \lambda_2 \end{bmatrix}$

which gives

$$2\lambda_1 + \lambda_2 = 4,$$

$$3\lambda_1 + \lambda_2 = 5$$

Solving these two equations we get $\lambda_1 = 1, \lambda_2 = 2$. Now substituting these values of λ_1 and λ_2 in (1), we get the linear combination

$$a_1 + 2a_2 - a_3 = 0 \text{ or } \sum_{j=1}^{k} \lambda_j a_j = 0 \qquad \text{.......... (2)}$$

Where $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = -1$

Now we have to determine which one of the three variables (x_1, x_2, x_3) should be zero.

$$\nu = \max_{1 \le j \le 3} \frac{\lambda_j}{x_j} = \max\left\{\frac{\lambda_1}{x_1}, \frac{\lambda_2}{x_2}, \frac{\lambda_3}{x_3}\right\}$$

$$= \max\left\{\frac{1}{2}, \frac{2}{3}, \frac{-1}{1}\right\} = \frac{2}{3}$$

Let $\hat{x} = \left(x_1 - \frac{\lambda_1}{\nu}, x_2 - \frac{\lambda_2}{\nu}, x_3 - \frac{\lambda_3}{\nu}\right)$

Then, $x_1 - \frac{\lambda_1}{\nu} = 2 - \frac{1}{\frac{2}{3}} = \frac{1}{2}$,

$$x_2 = \frac{\lambda_2}{\nu} = 3 - \frac{2}{\frac{2}{3}} = 0 \text{ (which was expected also),}$$

$$x_3 - \frac{\lambda_3}{\nu} = 1 - \left(\frac{-1}{\frac{2}{3}}\right) = \frac{5}{2}$$

Now this solution $\hat{x} = \left(\frac{1}{2}, 0, \frac{5}{2}\right)$ will be a basic feasible if the vectors $\bar{a}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\bar{a}_3 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$ associated with non - zero variables x_1 and x_3 are linearly Independent.

Obviously a_1 and a_3 are linearly independent.

Hence the required basic feasible solution is

$$x_1 = \frac{1}{2}, x_2 = 0, x_3 = \frac{5}{2}$$

To verify, we have $\frac{1}{2}\begin{bmatrix}1\\3\end{bmatrix}+0\begin{bmatrix}1\\1\end{bmatrix}+\frac{5}{2}\begin{bmatrix}4\\5\end{bmatrix}=\begin{bmatrix}11\\14\end{bmatrix}$

Example 2.5

Show that the feasible solution $x_1 = 1, x_2 = 0, x_3 = 1, z = 3$ to the system

$x_1 + x_2 + x_3 = 2$

$x_1 - x_2 + x_3 = 2$

$2x_1 + 3x_2 + 4x_3 = z\,(\text{Min})$ is not basic.

Solution :

First, we express the given system of constraint equations in matrix form :

$$\begin{bmatrix}1 & 1 & 1\\1 & -1 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}2\\2\end{bmatrix}$$

Therefore, according to our usual notations, we have

$$A=\begin{bmatrix}1 & 1 & 1\\1 & -1 & 1\end{bmatrix}, \bar{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}, \bar{b}=\begin{bmatrix}2\\2\end{bmatrix}$$

We show that the feasible solution $x_1 = 1, x_2 = 0, x_3 = 1$ is not basic.

So, we prove that the vectors

$$\bar{a}_1=\begin{bmatrix}1\\1\end{bmatrix} \text{ and } \bar{a}_3=\begin{bmatrix}1\\1\end{bmatrix}$$

are linearly dependent.

Since there exist non - zero scalars $\lambda_1 = 1, \lambda_2 = -1$ such that $\lambda_1\,\bar{a}_1 + \lambda_2\,\bar{a}_2 = \bar{0}$

or $$1.\begin{pmatrix}1\\1\end{pmatrix}+(-1).\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix},$$

the given feasible solution is not basic.

Theorem 2.8

Consider a L. P. P. max. $z = \bar{c}\cdot\bar{x}$, such that to $A\bar{x} = \bar{b}, \bar{x} \geq 0$.

Let $A = (\bar{a}_1, a_2, ..., \bar{a}_{a+m})$ and $B = (\beta_1, \beta_2, ..., \beta_m)$ be a non singular submatrix of A.

Assume that a non - degenerate basic feasible solution $\bar{x}_B = B^{-1}\bar{b}$ to $A\bar{x} = \bar{b}$ yields a value of the objective function $z = \bar{c}_B \bar{x}_B$. If for any colunm $\bar{a}_j$ in A but not in B we have $c_j - z_j > 0$, and if at least one $y_{ij} > 0 \, (i = 1, 2, ..., m)$ where $\bar{a}_j = \sum_{i=1}^{m} y_{ij}\beta_i$, then we can find a new basic feasible solution by replacing one of the columns in B by a_j.

Proof

We shall obtain a new basic feasible solution by replacing one of the vectors (say $\bar{a}_j$) in A but not in B by some vector in B (say β_r). Obviously,

$\bar{a}_j \neq \beta_i$ (i = 1, 2, ..., m)

Since $\bar{a}_j$ can be expressed as the linear combination of vectors in B, therefore

$$\bar{a}_j = \sum_{i=1}^{m} y_{ij}\beta_i$$

or $$\bar{a}_j = y_{1j}\beta_1 + y_{2j}\beta_2 + ... + y_{rj}\beta_r + ... + y_{mj}\beta_m \qquad ....\ (1)$$

Now, by using the replacement theorem, a_j can replace β_r and still maintains the basic matrix, provided $y_{rj} \neq 0$.

Assuming $y_{rj} \neq 0$, where $y_{rj} > 0, \bar{a}_j$ can be written as

$$\bar{a}_j = \sum_{\substack{i=1 \\ i \neq r}}^{m} y_{ij}\beta_i + y_{rj}\beta_r \qquad ....\ (2)$$

Solving the equation (2) for β_r, we obtain

$$\beta_r = \frac{1}{y_{rj}}\bar{a}_j - \sum_{\substack{i=1 \\ i \neq r}}^{m} \frac{y_{ij}}{y_{rj}}\beta_i \qquad ....\ (3)$$

Also, we have $B\bar{x}_B = \bar{b}$

or $$(\beta_1, \beta_2, ..., \beta_m)(x_{B1}, x_{B2}, ..., x_{Br}, ..., x_{Bm}) = \bar{b}$$

or $$x_{B1}\beta_1 + x_{B2}\beta_2 + ... + x_{Br}\beta_r + ... + x_{Bm}\beta_m = \bar{b}$$

or $$\sum_{\substack{i=1 \\ i \neq r}}^{m} x_{Bi}\beta_i + x_{Br}\beta_r = \bar{b} \qquad ....\ (4)$$

Substituting the value of β_r from (3) in (4), we obtain

$$\sum_{\substack{i=1\\ i\neq r}}^{m} x_{B_i}\,\beta_i + x_{B_r}\left[\frac{1}{y_{rj}}\,\bar{a}_j - \sum_{\substack{i=1\\ i\neq r}}^{m}\frac{y_{ij}}{y_{rj}}\beta_i\right] = \bar{b}$$

$$\Rightarrow \sum_{\substack{i-1\\ i\neq r}}^{m}\left(x_{Bi} - x_{Br}\frac{y_{ij}}{y_{rj}}\right)\beta_i + \frac{x_{Br}}{y_{rj}}a_j = \bar{b} \qquad \text{.......... (5 a)}$$

or $$\sum_{\substack{i=1\\ i\neq r}}^{m}\hat{x}_{Bi}\,\beta_i + \hat{x}_{Br}\,\bar{a}_j = \bar{b} \qquad \text{.......... (5 b)}$$

Where $\hat{x}_{Bi} = x_{Bi} - x_{Br}\dfrac{y_{ij}}{y_{rj}}, i=1,2,\ldots,m; i\neq r$, ... (6 a)

$\hat{x}_{Br} = \dfrac{x_{Br}}{y_{rj}}(\text{for } i=r)$ .... (6 b)

Comparison of (5 b) with (4) indicates that the new basic solution of $A\bar{x}=\bar{b}$ is given by

$$\hat{x}_B = \left(\hat{x}_{Bi}, \hat{x}_{Br}\right), i=1,2,\ldots,m; i\neq r$$

$$= \left(\hat{x}_{B1}, \hat{x}_{B2}, \ldots, \hat{x}_{Br}, \ldots., \hat{x}_{Bm}\right)$$

$$= \left(x_{B1} - x_{Br}\frac{y_{1j}}{y_{rj}}, x_{B2} - x_{Br}\frac{y_{2j}}{y_{rj}}, \ldots., \frac{x_{Br}}{y_{rj}}, \ldots., x_{Bm} - x_{Br}\frac{y_{mj}}{y_{rj}}\right)$$

and other non - basic components are zero.

For the new basic solution to be feasible, we require

$\hat{x}_{Bi} \geq 0, i=1,2,\ldots.,m$

Hence $x_{Bi} - x_{Br}\dfrac{y_{ij}}{y_{rj}} \geq 0, i=1,2,\ldots,m, i\neq r$ and .... (7 a)

$\dfrac{x_{Br}}{y_{rj}} \geq 0$ .... (7 b)

We see that (7 b) holds as $y_{rj}>0$ and since we start with a non - degenerate basic feasible solution, $x_{Bi}>0, i=1,2,....,m$. If $y_{rj}>0$ and $y_{ij}\leq 0 (i\neq r)$, then (7 a) is satisfied. If $y_{rj}>0$ and $y_{ij}>0 (i\neq r)$, then equation (7 a) is satisfied only when

$$\frac{x_{Bi}}{y_{ij}}-\frac{x_{Br}}{y_{rj}}\geq 0 \qquad \text{(dividing (7 a) by } y_{ij}>0\text{)}$$

or $$-\frac{x_{Bi}}{y_{rj}}\geq -\frac{x_{Bi}}{y_{ij}}$$

or $$\frac{x_{Br}}{y_{rj}}\leq \frac{x_{Bi}}{y_{ij}}$$

or $$\frac{x_{Br}}{y_{rj}}=\underset{i}{\text{Min}}\left[\frac{x_{Bi}}{y_{ij}}\right]$$

This, if we select r such that

$$\nu=\frac{x_{Br}}{y_{rj}}=\underset{i}{\text{Min}}\left[\frac{x_{Bi}}{y_{ij}}, y_{ij}>0\right] \qquad \text{.......... (8)}$$

then column β_r will be removed from basis matrix B to replace a_j so that the new basic solution will be feasible. This completes the proof.

Note

1) We denote the new non - singular matrtx, obtained from B by replacing β_r with $\bar{a}_j$ by

$$\hat{B}=\left(\hat{B}_1,\hat{B}_2,...,\hat{B}_m\right), \text{ where}$$

$$\hat{B}_i=\beta_i, i\neq r, \hat{B}_r=\bar{a}_j$$

2) If the minimum in (8) is not unique, the new basic solution will be degenerate. In this case, the number of positive basic variables will be less than m.

The procedure in above theorem can be explained by the following numerical example.

Example 2.6

Given the non - degeneate basic feasible solution $x_3 = 4$ and $x_4 = 8$ to the following LP problem

Max. $z = x_1 + 2x_2$, subject to

$$x_1 + 2x_2 + x_3 = 4$$

$$x_1 + 4x_2 + x_4 = 8$$

obtain the new basic feasible solution.

Solution :

The given basic feasible solution can be expressed as $B x_B = \bar{b}$

$$\Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 8 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}$$

Here, we have

$$x_B = \begin{pmatrix} x_{B1} \\ x_{B2} \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \bar{b} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}$$

$$\begin{matrix} \quad \bar{a}_1 & \bar{a}_2 & \beta_1 & \beta_2 \end{matrix}$$

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{pmatrix}, \bar{x} = \begin{pmatrix} 0 \\ 0 \\ 4 \\ 8 \end{pmatrix}$$

The $\bar{y}_j$'s for every column $\bar{a}_j$ in A but not in B are

$$\bar{y}_1 = B^{-1}\bar{a}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} y_{11} \\ y_{21} \end{pmatrix}$$

$$\bar{y}_2 = B^{-1}\bar{a}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} y_{12} \\ y_{22} \end{pmatrix}$$

Since $y_{11} = 1, y_{21} = 1$ are > 0, we can insert $\bar{a}_1$ in B. We now select β_r for replacement by $\bar{a}_1$ which corresponds to the value of suffix r determined by the minimum ratio rule :

$$\frac{x_{Br}}{y_{r1}} = \underset{i}{\text{Min}} \left[\frac{x_{Bi}}{y_{i1}}, y_{i1} > 0 \right]$$

Therefore,

$$\frac{x_{Br}}{y_{r1}} = \text{Min}\left[\frac{x_{B1}}{y_{11}}, \frac{x_{B2}}{y_{21}}\right]$$

$$\Rightarrow \frac{x_{Br}}{y_{r1}} = \text{Min}\left[\frac{4}{1}, \frac{8}{1}\right] = \frac{4}{1}$$

$$\Rightarrow \frac{x_{Br}}{y_{r1}} = \frac{x_{B1}}{y_{11}} \Rightarrow r = 1$$

Hence we remove β_1.

The new basis matrix becomes

$$\hat{B} = \left(\hat{\beta}_1, \hat{\beta}_2\right) = (\bar{a}_1, \beta_2) \qquad \text{(because } \bar{a}_1 \text{ is replaced by } \beta_1\text{)}$$

$$= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

Now we can find the new basic feasible solution $\hat{x}_B$ either by using the result $\hat{x}_B = \hat{B}^{-1}\bar{b}$ or using the transformation formulae (7 a) and (7 b) of Theorem 2.8.

Hence the new basic feasible solution is :

$$\hat{x}_{B1} = \frac{x_{B1}}{y_{11}} = \frac{4}{1} = 4$$

$$\hat{x}_{B2} = x_{B2} - x_{B1} = \frac{y_{21}}{y_{11}} = 8 - 4 \times \frac{1}{1} = 4$$

So that the solution to the original system of equations becomes

$$x_1 = x_{B1} = 4, x_2 = 0, x_3 = 0, x_4 = x_{B2} = 4$$

we note that, if we had inserted $\bar{a}_2$ instead of $\bar{a}_1$, the new basic feasible solution would have been degenerate. We have developed the procedure for obtaining a new basic feasible solution. Now we determine the value of the objective function corresponding to this new basic feasible solution. We verify, whether $\hat{z} > z$ where $\hat{z}$ denotes the new value of the objective function. For this, we prove the following theorem.

Theorem 2.9

Assume that we have a non - degenerate basis feasible solution $\bar{x}_B = B^{-1}\bar{b}$ to $A\bar{x} = \bar{b}$ which gives a value for the objective function $z = \bar{c}_B \bar{x}_B$. Assume further that we have obtained a new basic feasible solution $\hat{x}_B = \hat{B}^{-1}\bar{b}$ to $A\bar{x} = \bar{b}$ by replacing one of the columns in B by a column $\bar{a}_j$ (for which $y_{rj} > 0$) in A but not in B. If $c_j - z_j > 0$, the new value (denoted by $\hat{z}$) of the objective function will be greater than z, where $z_j = \bar{c}_B \bar{y}_j$ and $\bar{y}_j = B^{-1}\bar{a}_j$.

Proof

The value of the objection function for the original basic feasible solution is

$$z = \bar{c}_B \bar{x}_B$$

$$= (c_{B1}, c_{B2}, ..., c_{Bm})\ (x_{B1}, x_{B2}, ..., x_{Bm})$$

or $$z = \sum_{i=1}^{m} c_{Bi} x_{Bi} \qquad \text{.......... (A)}$$

The new value is given by

$$\hat{z} = \hat{c}_B \hat{x}_B$$

or $$\hat{z} = \sum_{i=1}^{m} \hat{c}_{Bi} \hat{x}_{Bi} = \sum_{\substack{i=1 \\ i \neq r}}^{m} \hat{c}_{Bi} \hat{x}_{Bi} + \hat{c}_{Br} \hat{x}_{Br}$$

where $\hat{c}_{Bi} = c_{Bi}\ (i \neq r)$, $\hat{c}_{Br} = c_j$

Therefore, $$\hat{z} = \sum_{\substack{i=1 \\ i \neq r}}^{m} c_{Bi} \hat{x}_{Bi} + c_j \hat{x}_{Br}$$

Subsituting the values of new variables $\hat{x}_{Bi}$ and $\hat{x}_{Br}$ from (7 a) and (7 b) of Theorem 2.8 into the last expression, we get

$$\hat{z} = \sum_{\substack{i-1 \\ i \neq r}}^{m} c_{Bi} \left(x_{Bi} - x_{Br} \frac{y_{ij}}{y_{rj}} \right) + c_j \frac{x_{Br}}{y_{rj}} \qquad \text{........... (B)}$$

Since the term for which i = r is $c_{Br}\left(x_{Br} - x_{Br} \frac{y_{rj}}{y_{rj}} \right) = 0$

we can include it in the summation (B) without changing $\hat{z}$, so that

$$\hat{z} = \sum_{i=1}^{m} c_{Bi}\left(x_{Bi} - x_{Br}\frac{y_{ij}}{y_{rj}}\right) + c_j\frac{x_{Br}}{y_{rj}}$$

$$= \sum_{i=1}^{m} c_{Bi}\, x_{Bi} - \frac{x_{Br}}{y_{rj}}\sum_{i=1}^{m} c_{Bi}\, y_{ij} + \frac{x_{Br}}{y_{rj}} c_j$$

$$= z - \frac{x_{Br}}{y_{rj}} z_j + \frac{x_{Br}}{y_{rj}} c_j$$

$$= z + \left(c_j - z_j\right)\frac{x_{Br}}{y_{rj}}$$

$$= z + \left(c_j - z_j\right)\nu, \text{ where } \nu = \frac{x_{Br}}{y_{rj}} \qquad \text{.......... (C)}$$

Now, from (C) we observe that the new value $\hat{z}$ of the objective function is the original value z plus the quantity $\left(c_j - z_j\right)\nu$. Since $\nu > 0$, and $c_j - z_j$ is greater than 0. The value of the objective function is improved.

Example 2.7

In worked example (2.6) show that the new value of the objective function is improved.

Solution :

Since $c_1 = 1, c_2 = 2, c_3 = 0, c_4 = 0$, then the original solution $x_3 = 4, x_4 = 8, x_1 = x_2 = 0$ gives

$$z = 1 \times 0 + 2 \times 0 + 0 \times 4 + 0 \times 8 = 0$$

In the new basis feasible solution x_1 replaces x_3

Since $z_1 = c_B\, y_1 = (0, 0)\begin{pmatrix}1\\1\end{pmatrix} = 0$

and since $c_1 - z_1 = 1 - 0 > 0$, $\hat{z}$ should exceed z (= 0). From (C) we get

$$\hat{z} = z + \left(c_1 - z_1\right)\frac{x_{B_1}}{y_{11}}$$

$$\hat{z} = 0 + 4\,(1 - 0)$$

$$= 4 \quad > z = 0$$

Theorem 2.10

If we select the vector $\bar{a}_k$ to replace β_r in B the suffix k can be selected by means of

$c_k - z_k = \underset{j}{Max}(c_j - z_j), c_j - z_j > 0$, so that the value of the objective function

z is increased as much as possible for the new basic feasible solution.

Proof

In the previsous Theorem we have obtained the improved value of z given by

$$\hat{z} = z + \frac{x_{Br}}{y_{rj}}(c_j - z_j)$$

Thus to give maximum value of $\hat{z}$ we should select that value of j for which the term.

$\frac{x_{Br}}{y_{rj}}(c_i - z_j)$ is maximum.

But the computational difficulty arises while obtaining $Max.\frac{x_{Br}}{y_{rj}}(c_j - z_j)$, because we have to compute $\frac{x_{Br}}{y_{rj}}$ for each a_j having $c_j - z_j > 0$ by the rule

$$\frac{x_{Br}}{y_{rj}} = \underset{j}{Min}\left[\frac{x_{Bi}}{y_{ij}}, y_{ij} > 0\right]$$

But the change in objective function depends on

$\frac{x_{Br}}{y_{rj}}$ and $c_j - z_j$ both.

Thus to avoid large number of computations of $\frac{x_{Br}}{y_{rj}}$, we can neglect the value of $\frac{x_{Br}}{y_{rj}}$.

Hence the most convenient and time saving rule for choosing the vector $\bar{a}_k$ to enter the basis B consists of selecting the largest $c_j - z_j$. This is equivalent to choosing the vector $\bar{a}_k$ to replace β_r by means of

$$c_k - z_k = \underset{j}{Max}(c_j - z_j), \text{ for } c_j - z_j > 0.$$

Note

The following are the advantages of using the above test.

1. The choice of vector $\overline{a}_k$ to enter the basis B by using above criteria gives the greatest possible increase in z in each step.
2. More than m iterations will not be needed to reach the optimal basic feasible solution.
3. It saves a time by giving the required solution in the least number of steps.

Definition 1 : Slack Variable

If the constraint has '$\leq$' sign then in order to make it an equality we have to add something positive to the left side of constraint. The non-negative variable which is added to the left hand side of the constraint to convert it into equation is called slack variable.

e.g. $x_1 + x_2 \leq 3$ then $x_1 + x_2 + x_3 = 3$ and x_3 is slack variable.

Surplus Variable

If a constraint has '$\geq$' sign then in order to make it an equality we have to subtract something non-negative from left hand side of inequality.

Definition

The positive variable which is subtracted from the left hand side of the constraint to convert it into equation is called surplus variable.

e.g. $x_1 + x_2 \geq 3$ then $x_1 + x_2 - x_3 = 3$ and variable x_3 is surplus variable.

Conversion of given LPP into standard form of LPP

Step 1

Convert constraints into equations except non-negativity of variable.

Step 2

Make right side of each constraint non-negative.

(multiply equation by (–1) if necessary)

e.g. $-x_1 + x_2 = -3 \equiv x_1 - x_2 = 3$

Step 3

Make all variables non-negative if variable *x* is unrestricted in sign write $x = x' - x''$ where $x', x'' \geq 0$.

Step 4

Convert objective function in maximization form.

$$\text{Min } f(x) \equiv \text{Max}[-f(x)]$$

Example

Express the following LPP in standard form.

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

Subject to

$$2x_1 + 3x_2 + 4x_3 \geq -4$$

$$3x_1 + 5x_2 + 2x_3 \geq 7$$

$x_1, x_2 \geq 0$, x_3 is unrestricted in sign.

Step 1

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

$$3x_1 + 5x_2 + 2x_3 - x_5 = 7$$

Step 2

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

$$3x_1 + 5x_2 + 2x_3 - x_5 = 7$$

Step 3

x_3 is unrestricted. $\therefore\ x_3 = x_3^{'} - x_3^{''}$

$$\text{Min } z = x_1 - 2x_2 + \left(x_3^{'} - x_3^{''}\right)$$

s.t. $\quad -2x_1 - 3x_2 - 4\left(x_3^{'} - x_3^{''}\right) + x_4 = 4$

$$3x_1 + 5x_2 + 2\left(x_3^{'} - x_3^{''}\right) - x_5 = 7$$

$$x_1, x_2, x_3^{'}, x_3^{''}, x_4, x_5 \geq 0$$

Step 4

$$\text{Min } z = x_1 - 2x_2 + \left(x_3^{'} - x_3^{''}\right)$$

$$\equiv \text{Max } z* = -x_1 + 2x_2 - \left(x_3^{'} - x_3^{''}\right)$$

Thus standard form is

$$\text{Max } z* = -x_1 + 2x_2 - x_3^{'} + x_3^{''}$$

Subject to

$$-2x_1 - 3x_2 - 4x_3' + 4x_3'' + x_4 = 4$$

$$3x_1 + 5x_2 + 2x_3' - 2x_3'' - x_5 = 7$$

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

Example 2.8

Solve the L. P. problem.

Max. $z = 3x_1 + 5x_2 + 4x_3$

subject to $2x_1 + 3x_2 \leq 8$

$$2x_2 + 5x_3 \leq 10$$

$$3x_1 + 2x_2 + 4x_3 \leq 15$$

and $x_1, x_2, x_3 \geq 0$

Solution :

The inequalities are converted into equalities by introduction of slack variables x_4, x_5 and x_6 as follows.

$$2x_1 + 3x_2 + 0.x_3 + x_4 = 8$$

$$0x_1 + 2x_2 + 5x_3 + x_5 = 10$$

$$3x_1 + 2x_2 + 4x_3 + x_6 = 15$$

Take $x_1 = 0, x_2 = 0, x_3 = 0$

Hence $x_4 = 8$ and $x_5 = 10, x_6 = 15$ which is the initial basic feasible solution.

Now we construct a starting simplex table. Here we compute Δ_j for all zero variables x_j , j = 1, 2, 3 by the formula.

$$\Delta_j = C_j - C_B Y_j$$

$$\Delta_1 = C_1 - C_B Y_1$$

$$\Delta_1 = 3 - (0,0,0)(2,0,3) = 3$$

$$\Delta_2 = C_2 - C_B Y_2$$

$$\Delta_2 = 5 - (0,0,0)(3,2,2) = 5$$

$$\Delta_3 = C_3 - C_B Y_3$$

$$\Delta_3 = 4 - (0,0,0)(0,5,2) = 4$$

Since all Δ_j are not less than or equal to zero therefore the solution is not optimal. So we proceed to the next step.

To find incoming vector :

Since $\Delta_2 = 3$ is max. of $\Delta_1, \Delta_2, \Delta_3$ therefore $\alpha_2 (= y_2)$ is incoming vector.

Starting simplex table 1

B	c_B	x_B	Y_1 (α_1)	Y_2 (α_2)	Y_3 (α_3)	Y_4 (β_1)	Y_5 (β_2)	Y_6 (β_3)	min ratio $\frac{x_B}{y_2}$
Y_4	0	8	2	[3]	0	1	0	0	$\frac{8}{3}$ →
Y_5	0	10	0	2	5	0	1	0	5
Y_6	0	15	3	2	4	0	0	1	$\frac{15}{4}$
$z = c_B x_B$		x_j	0	0	0	8	10	15	
= 0		c_j	3	5	4	0	0	0	
		Δ_j	3	5	4	x	x	x	
				↑		↓			

To find outgoing vector

Since α_2 is incoming vector therefore we consider the ratio

$$\frac{x_{B1}}{Y_2} = \left(\frac{x_{B1}}{Y_{12}}, \frac{x_{B2}}{Y_{22}}, \frac{x_{B3}}{Y_{32}}\right)$$

$$\text{i. e. } \frac{x_{B1}}{Y_2} = \left[\frac{8}{3}, 5, \frac{15}{2}\right]$$

$$\text{We have } \frac{x_{Br}}{Y_{r2}} = \min_i \left\{\frac{x_{Bi}}{Y_{i2}}, Y_{i2} > 0\right\}$$

$$= \min_i \left\{ \frac{x_{B1}}{Y_{12}}, \frac{x_{B2}}{Y_{22}}, \frac{x_{B3}}{Y_{32}} \right\} = \frac{8}{3}$$

Hence r = 1

i. e. β_1 is the outgoing vector.

Since α_2 is incoming vector and β_1 is outgoing vector, therefore the key element is $y_{12} (= a_{12})$ as shown in table 1 which is equal to 3.

In order to bring β_1 in place α_2 we make the following intermediate tables.

	x_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Y_4	8	2	$\boxed{3}$	0	1	0	0
Y_5	10	0	2	5	0	1	0
Y_6	15	3	2	4	0	0	1

Divide key element by 3 to get unity at this position and then subtract 2 times of the first row (obtained after dividing by 3) from the second and third row.

	x_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Y_2	$\frac{8}{3}$	$\frac{2}{3}$	1	0	$\frac{1}{3}$	0	0
Y_5	$\frac{14}{3}$	$-\frac{4}{3}$	0	5	$-\frac{2}{3}$	1	0
Y_6	$\frac{29}{3}$	$\frac{5}{3}$	0	4	$-\frac{2}{3}$	0	1

Now we construct second simplex table in which $\beta_1 (Y_4)$ is replaced by $\alpha_2 = (y_2)$.

Second simplex table 2

B	c_B	x_B	Y_1	Y_2 (β_1)	Y_3	Y_4	Y_5 (β_2)	Y_6 (β_3)	min ratio $\frac{x_B}{y_3}$
Y_2	5	$\frac{8}{3}$	$\frac{2}{3}$	1	0	$\frac{1}{3}$	0	0	--
Y_5	0	$\frac{14}{3}$	$-\frac{4}{3}$	0	$\boxed{5}$	$-\frac{2}{3}$	1	0	$\frac{14}{15} \rightarrow$min
Y_6	0	$\frac{29}{3}$	$\frac{5}{3}$	0	4	$-\frac{2}{3}$	0	1	$\frac{29}{12}$
$z = c_B x_B$		x_j	0	$\frac{8}{3}$	0	0	$\frac{14}{3}$	$\frac{29}{3}$	
		c_j	3	5	4	0	0	0	
		Δ_j	$-\frac{1}{3}$	x	4	$-\frac{5}{3}$	x	x	

$\uparrow$ incoming vector (under Y_3) $\qquad$ $\downarrow$ outgoing vector (under Y_5)

To test the optimality of the solution compute Δ_j for all zero variables x_1, x_3 and x_4.

$$\Delta_1 = c_1 - c_B\, Y_1 = 3 - (5,0,0)\left(\frac{2}{3}, -\frac{4}{3}, \frac{5}{3}\right)$$

$$\Delta_1 = c_1 - c_B\, Y_1 = 3 - \frac{10}{3} = -\frac{1}{3}$$

$$\Delta_3 = c_3 - c_B y_3 = 4 - (5,0,0)(0,5,4) = 4 - 0 = 0$$

$$\Delta_4 = c_4 - c_B\, Y_4 = 0 - (5,0,0)\left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right)$$

$$\Delta_4 = -\frac{5}{3}$$

Since all Δ_j are not less than or equal to zero, therefore this solution is also not optimal.

Since $\Delta_3 = 4$ is maximum of the Δ_j's, $\alpha_3 = (Y_3)$ is the incoming vector.

Also $$\frac{x_{Br}}{Y_{r3}} = \min_i \left[\frac{x_{Bi}}{Y_{i3}}, Y_{i3>0}\right]$$

$$= \min\left[\frac{x_{B2}}{Y_{23}}, \frac{x_{B3}}{Y_{33}}\right] \text{ (since } Y_{13} = 0\text{)}$$

$$= \min\left[\frac{14}{15}, \frac{29}{12}\right] = \frac{14}{15} = \frac{x_{B2}}{Y_{23}}$$

$$\Rightarrow r = 2$$

Therefore $\beta_2 (= y_5)$ is the outgoing vector and $y_{23} = a_{23} = 5$ is the key element.

In order to bring y_3 in place of $\beta_2 (= y_5)$ we make the following intermediate table.

	x_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Y_2	$\frac{8}{3}$	$\frac{2}{3}$	1	0	$\frac{1}{3}$	0	0
Y_5	$\frac{14}{3}$	$-\frac{4}{3}$	0	$\boxed{5}$	$-\frac{2}{3}$	1	0
Y_6	$\frac{29}{3}$	$\frac{5}{3}$	0	4	$-\frac{2}{3}$	0	1

Divide the key element by 5 to get 1 at this position, then subtract 4 times of the second row thus obtained from the third row.

	x_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Y_2	$\frac{8}{3}$	$\frac{2}{3}$	1	0	$\frac{1}{3}$	0	0
Y_5	$\frac{14}{15}$	$-\frac{4}{5}$	0	1	$-\frac{2}{15}$	$\frac{1}{5}$	0
Y_6	$\frac{89}{15}$	$\frac{41}{15}$	0	0	$-\frac{2}{15}$	$-\frac{4}{5}$	1

The third simplex table in which $\beta_2 (= Y_5)$ is replaced by Y_3 is as follows

Table 3

B	c_B	x_B	Y_1	Y_2 $(=\beta_1)$	Y_3 $(=\beta_2)$	Y_4	Y_5	Y_6 $(=\beta_3)$	min ratio $\frac{x_B}{y_1}$
Y_2	5	$\frac{8}{3}$	$\frac{2}{3}$	1	0	$\frac{1}{3}$	0	0	4
Y_5	4	$\frac{14}{15}$	$-\frac{4}{15}$	0	1	$-\frac{2}{15}$	$\frac{1}{5}$	0	$-\frac{7}{2}$ neg.
Y_6	0	$\frac{89}{15}$	$\boxed{\frac{41}{15}}$	0	0	$-\frac{2}{15}$	$-\frac{4}{05}$	1	$\frac{89}{41}$ min→
		x_j	0	$\frac{8}{3}$	$\frac{14}{15}$	0	0	$\frac{89}{15}$	
		c_j	3	5	4	0	0	0	
		Δ_j	$\frac{11}{15}$	x	x	$-\frac{17}{15}$	$-\frac{4}{5}$	x	

↑ Incoming vector (under Y_1) ↓ Outgoing vector (under Y_6)

To test the optimality of the solution again compute Δ_j for all zero variables x_1, x_4 and x_5.

$$\Delta_1 = c_1 - c_B\, y_1 = 3 - (5,4,0)\left(\frac{2}{3}, -\frac{4}{15}, \frac{41}{15}\right)$$

$$= 3 - \left(\frac{10}{3} - \frac{16}{15}\right) = \frac{11}{15}$$

$$= 3 - \left(\frac{50-16}{15}\right) = \frac{45-34}{15}$$

$$\Delta_4 = c_4 - c_B \frac{y_4}{4} = 0 - (5,4,0)\left(\frac{1}{3}, -\frac{2}{15}, -\frac{2}{15}\right)$$

$$\Rightarrow \quad \Delta_4 = \left(-\frac{5}{3} - \frac{8}{15}\right) = -\frac{17}{15}, \Delta_5 = c_5 - c_B\, y_5 = 0 - (5,4,0)\left(0, \frac{1}{5}, \frac{4}{5}\right) = -4/5$$

Also $\Delta_5 = c_5 - c_e\, y_5 = -\frac{4}{5}$

Since all the Δ_j's are not less than or equal to zero, therefore the solution is not optimal.

Since Δ_1 is maximum of the Δ_j's, it follows that, $\alpha_1 (= Y_1)$ is the incoming vector.

Also $$\frac{x_{Br}}{Y_{r1}} = \min_i \left[\frac{x_{Bi}}{Y_{i1}}, Y_{i1} > 0 \right]$$

$$= \min_i \left[\frac{y_{B1}}{Y_{11}}, \frac{x_{B3}}{Y_{31}} \right] \qquad (\because Y_{21} \text{ is negative})$$

$$= \min_i \left[4, \frac{89}{41} \right] = \frac{89}{41}$$

$$\Rightarrow r = 3.$$

i. e. $\beta_3 (= Y_6)$ is the outgoing vector and $Y_{31} = a_{31} = \frac{41}{15}$ is the key element.

Again in order to bring Y_1 in place of $\beta_3 (= Y_6)$ we make the following intermediate table.

	x_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Y_2	$\frac{8}{3}$	$\frac{2}{3}$	1	0	$\frac{1}{3}$	0	0
Y_3	$\frac{14}{15}$	$-\frac{4}{15}$	0	1	$-\frac{2}{15}$	$\frac{1}{5}$	0
Y_6	$\frac{89}{15}$	$\boxed{\frac{41}{15}}$	0	0	$-\frac{2}{15}$	$-\frac{4}{5}$	1

Divide the key element by $\frac{41}{15}$ to get 1 at this position, then subtract $\frac{2}{3}$ times of the third row from the first row and adding $\frac{4}{15}$ times of the third row to the second row we have,

	x_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Y_2	$\frac{50}{41}$	0	1	0	$\frac{15}{41}$	$\frac{8}{41}$	$-\frac{10}{41}$
Y_3	$\frac{62}{41}$	0	0	1	$-\frac{6}{41}$	$\frac{5}{41}$	$\frac{4}{41}$
Y_6	$\frac{89}{15}$	1	0	0	$-\frac{2}{41}$	$-\frac{12}{41}$	$\frac{15}{41}$

The fourth simplex table in which $\beta_3 \left(=Y_5\right)$ is replaced by y_1 is as follows.

B	c_B	x_B	Y_1 β_3	Y_2 β_1	Y_3 β_1	Y_4	Y_5	Y_6	min ratio
Y_2	5	$\frac{50}{41}$	0	1	0	$\frac{15}{41}$	$\frac{8}{41}$	$-\frac{10}{41}$	
Y_3	4	$\frac{62}{41}$	0	0	1	$-\frac{6}{41}$	$\frac{5}{41}$	$\frac{4}{41}$	
Y_1	3	$\frac{89}{41}$	1	0	0	$-\frac{2}{41}$	$-\frac{12}{41}$	$\frac{15}{41}$	
$z=c_B x_B$		x_j	$\frac{89}{41}$	$\frac{50}{41}$	$\frac{62}{41}$	0	0	0	
= 765/41		c_j	3	5	4	0	0	0	
		Δ_j	x	x	x	$-\frac{45}{41}$	$-\frac{24}{41}$	$-\frac{11}{41}$	

To test the optimality of the solution again compate Δ_j for all zero variables x_4, x_5 and x_6.

$$\Delta_4 = c_4 - c_B \frac{1}{4} = 0 - (5,4,3)\left(\frac{15}{41}, -\frac{5}{41}, -\frac{2}{41}\right) = -\frac{45}{41}$$

$$\Delta_5 = c_5 - c_B \frac{1}{5} = 0 - (5,4,3)\left(\frac{8}{41}, \frac{5}{41}, -\frac{12}{41}\right) = -\frac{24}{41}$$

$$\Delta_6 = c_6 - c_B \frac{1}{6} = 0 - (5,4,3)\left(-\frac{10}{41}, \frac{4}{41}, \frac{15}{41}\right) = -\frac{11}{41}$$

Since all the Δ_j's for zero variables are negative so, this solution is optimal.

Hence $x_1 = \frac{89}{41}, x_2 = \frac{50}{41}, x_3 = \frac{62}{41}$

and $\max. z = \frac{765}{41}$

Computational Procedure for Simplex Method

Example $\text{Max } z = 3x_1 + 2x_2$

Subject to $x_1 + x_2 \le 4$

$x_1 - x_2 \le 2, \quad x_1, x_2 \ge 0$

Answer

Step 1

Convert the given LPP into a standar form.

$\text{Max } z = 3x_1 + 2x_2 + 0x_3 + 0x_4$

$x_1 + x_2 + x_3 = 4$

Subject to $x_1 - x_2 + x_4 = 2, \quad x_1, x_2, x_3, x_4 \ge 0$

Step 2

Construct starting simplex table. Variable which form identity matrix in starting simplex table are basic variables, c_B represent cost of basic variables.

Basic variable	$c_B \rightarrow$ cost of B.V. c_B	x_B	3 x_1	2 x_2	0 x_3	0 x_4
x_3	0	4	1	1	1	0
x_4	0	2	1	– 1	0	1

Step 3

Calculate $\Delta_j = c_B \cdot x_j - c_j$

$\Delta_1 = c_B \cdot x_1 - c_1$

$= (0)(1) + (0)(1) - 3$

$= -3$

$\Delta_2 = c_B \cdot x_2 - c_2$

$= (0)(1) + (0)(-1) - 2$

$= -2$

$\Delta_3 = \Delta_4 = 0$

Step 4 : Optimality Test

(i) If all $\Delta_j \geq 0$ the solution is optimal. Alterative optimal solutio will exist if any Δ_j corresponding to non basic x_j is also zero.

(ii) If corresponding to any – ve Δ_j, all elements of the column x_j are – ve or zero (≤ 0), then the solution under test is unbounded.

(iii) If at least one $\Delta_j < 0$ then solution is not optimal and therefore proceed to improve the solution in the next step.

Step 5

Choose incoming and outgoing variable.

Let $\Delta_k = \underset{j}{\text{Min}}\{\Delta_j\} < 0$

The correspoding variable x_k is **incoming varable.**

Outgoing variable is decided by minimum ratio (component wise) rule.

If $$\frac{x_{Br}}{x_{kr}} = \underset{i}{\text{Min}}\left\{\frac{x_{Bi}}{x_{ki}} / x_{ki} > 0\right\}$$

Then x_{Br} is outgoing variable from the set of basic variables.

$$\Delta_k = \underset{j}{\text{Min}}\{\Delta_j\}$$

Since

$$\underset{j}{\text{Min}}\{\Delta_j\} = \text{Min}\{-3, -2, 0, 0\} = -3$$

The variable corresponding to $\Delta_1 = -3$ is x_1. Therefore x_1 is incoming variable and x_1 becomes basic variable.

Consider component wise ratio of the values of basic variables i.e. x_B and coefficient of incoming variable x_1 and take its Minimum.

$$\underset{k}{\text{Min}}\left\{\frac{x_{Bk}}{x_{1k}}\right\} = \text{Min}\left\{\frac{4}{1}, \frac{2}{1}\right\} = 2$$

Corresponds to x_4 and therefore x_4 is outgoing variable.

Thus x_1 is incoming and x_4 is outgoing variable.

B.V.	c_B	c_j x_B	3 (x_1)	2 x_2	0 x_3	0 x_4	Min Ratio $\frac{x_B}{x_1}$
x_3	0	4	1	1	1	0	$\frac{4}{1}=4$
← (x_4)	0	2	1	– 1	0	1	$\frac{2}{1}=\boxed{1}$
		Δ_j	– 3 ↑	– 2	0	0	

Step 6

In order to make x_1 as basic variable perform elementary row operations to convert column corresponding to variable x_1 as unit vector. Here operation $R_1 - R_2$ will make column corresponding to variable x_1 as unit vector. The position 1 in the unit vector depends upon the position of incoming variable in basic variables.

B.V.	c_B	x_B	3 x_1	2 x_2	0 x_3	0 x_4
x_3	0	2	0	2	1	– 1
x_1	3	2	1	– 1	0	1

Repeat step 4, 5 and 6.

B.V.	c_B	c_j x_B	3 x_1	2 (x_2)	0 x_3	0 x_4	Min ratio
← x_3	0	2	0	2	1	– 1	$\frac{2}{2}=\boxed{1}$
x_1	3	2	1	– 1	0	1	---
		Δ_j →	0	– 5	0	3	

Step 4 : $\Delta_2 < 0$

Therefore, variable x_2 is incoming variable. Component wise ratio $\frac{x_B}{x_2}$ is {1, –}. Minimum ratio corresponds to x_3 and x_3 is outgoing variable. Now make column corresponding to x_2 as unit vector.

B.V.	c_B	x_B	3 x_1	2 x_2	0 x_3	0 Min x_4 ratio
x_2	2	1	0	1	$\frac{1}{2}$	$-\frac{1}{2}$
x_1	3	3	1	0	$\frac{1}{2}$	$\frac{1}{2}$
		Δ_j	0	0	$\frac{3}{2}$	$\frac{1}{2}$

Since $\Delta_j \geq 0 \quad \forall j$ the solution $x_2 = 1$ and $x_1 = 3$ is an optimal solution and optimal value.

$$\text{Max } z = 3x_1 + 2x_2 = 3(3) + 2(1) = 11$$

Example 2.9

Solve by simplex method the following L. P. problem.

Minimize $\quad z = x_1 - 3x_2 + 2x_3$

Subject to $\quad 3x_1 - x_2 + 2x_3 \leq 7$

$$-2x_1 + 4x_2 \leq 12$$

$$-4x_1 + 3x_2 + 8x_3 \leq 10$$

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

Solution :

First we convert the problem of minimization to maximization problem by taking objective function $z' = -z$.

max. $\quad z' = -z = -x_1 + 3x_2 - 2x_3$

Now the equations obtained by introducing slakc variables x_4, x_5, x_6 are as follows.

$$3x_1 - x_2 + 2x_3 + x_4 = 7$$

$$-2x_1 + 4x_2 + 0x_3 + x_5 = 12$$

$$-4x_1 + 3x_2 + 8x_3 + x_6 = 10$$

Taking $x_1 = x_2 = x_3 = 0$ we get $x_4 = 7, x_5 = 12, x_6 = 10$ which is the starting B. F. S.

Starting simplex table

B	c_B	x_B	Y_1 (α_1)	Y_2 (α_2)	Y_3 α_3	Y_4 β_1	Y_5 β_2	Y_6 β_3	min ratio $\frac{x_{Bi}}{Y_{12}}$
Y_4	0	7	3	-1	2	1	0	0	$-7 \rightarrow$ neg.
Y_5	0	12	-2	4	0	0	1	0	$3 \rightarrow$ min
Y_6	0	10	-4	3	8	0	0	1	$\frac{10}{3}$
$z^1 = c_B x_8$		x_j	0	0	0	7	12	10	
= 0		c_j	-1	3	-2	0	0	0	
		Δ_j	-1	3	-2	x	x	x	

$$\Delta_1 = c_1 = c_B y_1 = -1 - (0,0,0)(3,-2,-4) = -1$$

$$\Delta_2 = c_2 - c_B\, y_2 = 3 - (0,0,0)(-1,4,3) = 3$$

$$\Delta_3 = c_3 - c_B\, y_3 = -2 - (0,0,0)(2,0,8) = -2$$

Since all the Δ_j are not less than or equal to zero therefore the solution is not optimal.

Δ_2 is maximum.

Hence the incoming vector is $\alpha_2 (= y_2)$ and by mini ratio rule outgoing vector is $\beta_2 (= y_5)$.

Therefore key element $= y_{22} = a_{22} = 4$

In order to b ring $\alpha_2 (= y_2)$ in place of $\beta_2 (= y_5)$ the inter mediate table is as follows.

	x_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Y_4	7	3	-1	2	1	0	0
Y_5	12	-2	4	0	0	1	0
Y_6	10	-4	3	8	0	0	1

	x_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Y_4	10	$\frac{5}{2}$	0	2	1	$\frac{1}{4}$	0
Y_2	3	$-\frac{1}{2}$	1	0	0	$\frac{1}{4}$	0
Y_6	1	$-\frac{5}{2}$	0	8	0	$-\frac{3}{4}$	1

Second simplex table

B	c_B	x_B	Y_1	Y_2 β_2	Y_3	Y_4 β_1	Y_5	Y_6 β_3	min ratio $\frac{x_B}{Y_1}$
Y_4	0	10	$\frac{5}{2}$	0	2	1	$\frac{1}{4}$	0	4 min
Y_2	3	3	$-\frac{1}{2}$	1	0	0	$\frac{1}{4}$	0	-6 neg.
Y_6	0	1	$-\frac{5}{2}$	0	8	0	$-\frac{3}{4}$	1	$-\frac{2}{5}$ neg
		x_j	0	3	0	10	0	1	
		c_j	-1	3	-2	0	0	0	
		Δ_j	$\frac{1}{2}$	x	-2	x	$-\frac{3}{4}$	x	
			↑			↓			

$$\Delta_1 = c_1 - c_B y_1 = -1 - (0,3,0)\left(\frac{5}{2}, -\frac{1}{2}, -\frac{5}{2}\right) = \frac{1}{2}$$

$$\Delta_3 = c_3 - c_B y_3 = -2 - (0,3,0)(2,0,8) = -2$$

$$\Delta_5 = c_5 - c_B y_5 = 0 - (0,3,0)\left(\frac{1}{4}, \frac{1}{4}, -3\right) = -\frac{3}{4}$$

Since all the Δ_j are not less than or equal to zero the solution is not optimal.

Here $\Delta_1 = \frac{1}{2}$ is maximum.

Therefore y_1 is the incoming, vector and by the minimal ratio rate we find that $\beta_1 (= y_4)$ as the outgoing vector.

Therefore key element $= y_{11} = \frac{5}{2}$.

In order to to bring y_1 in place of β_1 the inter mediate table is as follows

	x_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Y_4	10	$\frac{5}{2}$	0	2	1	$\frac{1}{4}$	0
Y_2	3	$-\frac{1}{2}$	1	0	0	$\frac{1}{4}$	0
Y_6	1	$-\frac{5}{2}$	0	8	0	$-\frac{3}{4}$	1

	x_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Y_1	4	1	0	$\frac{4}{5}$	$\frac{2}{5}$	$\frac{1}{10}$	0
Y_2	5	0	1	1	$\frac{1}{5}$	$\frac{3}{10}$	0
Y_6	11	0	0	13	$\frac{5}{2}$	$-\frac{1}{2}$	1

Third simplex table

B	c_B	x_B	Y_1 β_1	Y_2 β_2	Y_3	Y_4	Y_5	Y_6	min ratio β_3
Y_1	-1	4	1	0	$\frac{4}{5}$	$\frac{2}{5}$	$\frac{1}{10}$	0	
Y_2	3	5	0	1	1	$\frac{1}{2}$	$\frac{3}{10}$	0	
Y_6	0	11	0	0	13	$\frac{5}{2}$	$-\frac{1}{2}$	1	
$z' = c_B\, x_B$		x_j	4	5	0	0	0	11	
= 11		c_j	-1	3	-2	0	0	0	
		Δ_j	x	x	$-\frac{21}{5}$	$-\frac{11}{10}$	$-\frac{41}{40}$	x	

$$\Delta_3 = c_3 - c_B\, Y_3 = -2 - (-1,3,0)\left(\frac{4}{5},1,13\right) = -\frac{21}{5}$$

$$\Delta_4 = c_4 - c_B\, Y_4 = 0 - (-1,3,0)\left(\frac{2}{5},\frac{1}{2},\frac{5}{2}\right) = -\frac{11}{10}$$

$$\Delta_5 = c_5 - c_B\, Y_5 = -0 - (-1,3,0)\left(\frac{1}{10},\frac{3}{8},-\frac{19}{8}\right) = -\frac{41}{40}$$

Since all Δ_j's for all non zao variables are negative so this solution is optimal.

Optimal solution is

$$x_1 = 4, x_2 = 5, x_3 = 0$$

and max. $z' = 11$

Hence $\min z = -11$

Example 2.10

Using simplex algorithm to solve the problem.

max. $z = 2x_1 + 5x_2 + 7x_3$

subject to $3x_1 + 2x_2 + 4x_3 \le 100$

$x_1 + 4x_2 + 2x_3 \le 100$

$$x_1 + x_2 + 3x_3 \le 100$$

$$x_1, x_2, x_3 \ge 0$$

Solution :

The equations obtained by introducing slack variables x_4, x_5, x_6 are as follows.

$$3x_1 + 2x_2 + 4x_3 + x_4 = 100$$

$$x_1 + 4x_2 + 2x_3 + x_5 = 100$$

$$x_1 + x_2 + 3x_3 + x_6 = 100$$

Take $x_1 = x_2 = x_3 = 100$

Therefore starting B. F. S. is

$x_4 = 100, x_5 = 100, x_6 = 100$

Starting simplex table

B	c_B	x_B	Y_1 α_1	Y_2 α_2	Y_3 α_3	Y_4 β_1	Y_5 β_2	Y_6 β_3	min ratio
Y_4	0	100	3	2	4	1	0	0	25 min
Y_5	0	100	1	4	2	0	1	0	50
Y_6	0	100	1	1	3	0	0	1	$\frac{100}{3}$
$z' = c_B x_B$		x_j	0	0	0	100	100	100	
= 0		c_j	2	5	7	0	0	0	
		Δ_j	2	5	7	x	x	x	

$\uparrow$ in (column Y_3) $\downarrow$ out (column Y_4)

$$\Delta_1 = c_1 - c_B y_1 = 2 - (0,0,0)(3,1,1) = 2$$

$$\Delta_2 = c_2 - c_B y_2 = 5 - (0,0,0)(2,4,1) = 5$$

$$\Delta_3 = c_3 - c_B y_3 = 7 - 0 = 7$$

Since all Δ_j are not less than or equal to zero for zero variables, so the solution is not optimal.

Since $\Delta_3 = 7$ is maximum therefore $\alpha_3 (= y_3)$ is the incoming vector.

By the min ratio rule

$$\min\left\{\frac{x_{Bi}}{y_{i3}}, y_{i3} > 0\right\} = \frac{100}{4} = 25 \text{ , for i = 1}$$

Therefore $\beta_1 (= y_4)$ is the outgoing vector. Therefore the key element is $y_{13} = a_{13} = 4$. In order to brings β_1 in place of α_3 we divide the first row by 4 and then subtract 2 and 3 times of this row from the second and third rows respectively.

Thus the second simplex table is as follows.

B	c_B	x_B	Y_1	Y_2	Y_3 β_1	Y_4	Y_5 β_2	Y_6 β_6	min ratio $\frac{x_B}{y_2}$
Y_3	7	25	$\frac{3}{4}$	$\frac{1}{2}$	1	$\frac{1}{4}$	0	0	50
Y_5	0	50	$-\frac{1}{2}$	3	0	$-\frac{1}{2}$	1	0	$\frac{50}{3} \rightarrow$
Y_6	0	25	$-\frac{3}{4}$	$-\frac{1}{2}$	0	$-\frac{3}{4}$	0	1	-50 neg.
		x_j	0	0	25	0	50	25	
		c_j	2	5	7	0	0	0	
		Δ_j	$-\frac{13}{4}$	$\frac{3}{2}$	x	$-\frac{7}{4}$	x	x	

$\uparrow$ incoming vector (under Y_2); $\downarrow$ (under Y_5)

For above simplex table

$$\Delta_1 = c_1 - c_B\, y_1 = 2 - (7,0,0)\left(\frac{3}{4}, -\frac{1}{2}, -\frac{5}{4}\right) = 2 - \frac{21}{4}$$

$$\Delta_1 = -\frac{13}{4}$$

$$\Delta_2 = c_2 - c_B y_2 = +5 - (7,0,0)\left(\frac{1}{2}, 3, -\frac{1}{2}\right) = 5 - \frac{1}{2} = \frac{3}{2}$$

$$\Delta_4 = c_\alpha\ c_B y_4 = 0 - (7,0,0)\left(\frac{1}{4}, -\frac{1}{2}, -\frac{3}{4}\right) = -\frac{7}{4}$$

Since all Δ_j are not less than or equal to zero so the solution is not optimal.

Here $\Delta_2 = \frac{3}{2}$ is max.

Therefore y_2 is incoming vector and by min ratio rule we find that $\beta_2 (= y_5)$ is the outgoing vector. Key element is 3. Intermidiate table is :

	x_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Y_3	25	$\frac{3}{4}$	$\frac{1}{2}$	1	$\frac{1}{4}$	0	0
Y_5	50	$-\frac{1}{2}$	3	0	$-\frac{1}{2}$	1	0
Y_6	25	$-\frac{5}{4}$	$-\frac{1}{2}$	0	$-\frac{3}{4}$	0	1

The third simplex table is as follows.

B	C_B	X_B	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Min
Y_3	7	$\frac{50}{3}$	$\frac{5}{6}$	0	1	$\frac{1}{3}$	$-\frac{1}{6}$	0	
Y_2	5	$\frac{50}{3}$	$-\frac{1}{6}$	1	0	$-\frac{1}{6}$	$\frac{1}{3}$	0	
Y_6	0	$\frac{100}{3}$	$-\frac{4}{3}$	0	0	$-\frac{5}{6}$	$\frac{1}{6}$	1	
		x_j	0	$\frac{50}{3}$	$\frac{50}{3}$	0	0	$\frac{100}{3}$	
		c_j	2	5	7	0	0	0	
		Δ_j	-3	x	x	$-\frac{3}{2}$	$-\frac{1}{2}$	x	

$$\Delta_1 = C_1 - C_B\, Y_1 = z - (7,5,0)\left(\frac{5}{6}, -\frac{1}{6}, -\frac{4}{3}\right) = -3$$

$$\Delta_4 = C_4 - C_B\, Y_4 = 0 - (7,5,0)\left(\frac{1}{3}, -\frac{1}{6}, -1\right) = -\frac{3}{2}$$

$$\Delta_5 = C_5 - C_B\, Y_5 = 0 - (7,5,0)\left(-\frac{1}{6}, \frac{1}{3}, \frac{1}{2}\right) = -\frac{1}{2}$$

Since all Δ_j for zero variables are negative, this solution is optimal.

Optimal solution is $x_1 = 0, x_2 = \frac{50}{3}, x_3 = \frac{50}{3}$ and Max. $z = 200$.

Complete solution with all computational steps is conveniently represented in the following example.

Example :

Solve Max $z = 7x_1 + 5x_2$

Subject to $x_1 + 2x_2 \le 6$, $4x_1 + 3x_2 \le 12$, $x_1, x_2 \ge 0$

Solution :

Max $z = 7x_1 + 5x_2$

Subject to $x_1 + 2x_2 + x_3 = 6$, $4x_1 + 3x_2 + 0x_3 + x_4 = 12$, $x_1, x_2, x_3, x_4 \ge 0$

		c_j	7	5	0	0	Min
B.V.	c_B	x_B	x_1	x_2	x_3	x_4	ratio $\frac{x_B}{x_i}$
x_3	0	3	1	2	1	0	6
$\leftarrow x_4$	0	12	4	3	0	1	[3]
		Δ_j	$-7 \uparrow$	-5	0	0	
x_3	0	3	0	$\frac{5}{4}$	1	$-\frac{1}{4}$	
x_1	7	3	1	$\frac{3}{4}$	0	$\frac{1}{4}$	
		$\Delta_j \rightarrow$	0	$+\frac{1}{4}$	0	$\frac{7}{4}$	

Since $\Delta_j \geq 0 \quad \forall j$ the solution is optimal solution.

$x_1 = 3, x_2 = 0$ and Max $z = 7(3) + 5(0) = 21$.

Artificial Variable Technique

If starting simplex table do not contain identity matrix, we introduce new type of variables called artificial variables. These variables are fictitious and donot have any physical meaning. This is only a device to introduce identity matrix in starting simplex table and to get basic feasible solution so that simplex method may be adopted. Artificial variables are eliminated from the simplex table as and when they become zero.

Two Phase Simplex Method

The process of eliminating artificial variables is performed in phase I and phase II is used to get an optimal solution.

Computational Procedure of Two Phase Simplex Method

Phase I

In this phase the simplex method is applied to LPP with artificial variables leading to a final simplex table containing a bsic feasible solution (BFS) to the original problem.

Step 1

Assign a cost – 1 to each artificial variable and cost 0 to all other variables.

Step 2

Solve by simplex method until either of three possibilities do arise.

(i) If Max $z^* < 0$, given original problem does not have any feasible solution.

(ii) If Max $z^* = 0$ and atleast one artificial variable appears in the optimal basis (basic variable in last simplex table) at zero level then proceed to Phase II.

(iii) If Max $z^* = 0$ and no artificial variable appears in the optimal basis proceed to Phase II.

Phase II

Assign the actual cost to the variables in objective function and zero cost to every artificial variable that appears in the basis. This new objective function is now maximized by simplex method with last simplex table of phase I as starting simplex table with actual cost values.

Example 1

Solve the following problem

$$\text{Max } z = x_1 + x_2$$

Subject to $\quad 2x_1 + x_2 \geq 4$

$$x_1 + 7x_2 \geq 7, \quad x_1, x_2 \geq 0$$

Solution :

Convert the given problem into standard LPP.

$$\text{Max } z = -x_1 - x_2$$

s.t. $2x_1 + x_2 - x_3 = 4\ ,\ x_1 + 7x_2 - x_4 = 7$

i.e. $\begin{bmatrix} 2 & 1 & -1 & 0 \\ 1 & 7 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$

Since coefficient matrix donot contain identity matrix, we have to solve this problem by two phase method by introducing artificial variables.

Phase I $\quad$ $\text{Max } z* = -1a_1 - 1a_2$

Subject to $\quad$ $2x_1 + x_2 - x_3 + a_1 = 4$

$x_1 + 7x_2 - x_4 + a_2 = 7\ ,\ x_1, x_2, x_3, x_4, a_1, a_2 \geq 0$

B.V.	c_B	x_B	0 x_1	0 x_2	0 x_3	0 x_4	–1 a_1	–1 a_2	Min Rato
a_1	– 1	4	2	1	– 1	0	1	0	4
$\leftarrow a_2$	– 1	7	1	$\boxed{7}$	0	– 1	0	1	$\boxed{1}$
		Δ_j	– 3	– 8	1	1	0	0	
$\leftarrow a_1$	– 1	3	$\frac{13}{7}$	0	– 1	$\frac{1}{7}$	1	$-\frac{1}{7}$	$\boxed{\frac{21}{13}}$
x_2	0	1	$\frac{1}{7}$	1	0	$-\frac{1}{7}$	0	$\frac{1}{7}$	7
			$-\frac{13}{7}\uparrow$	0	1	$-\frac{1}{7}$	0	$\frac{8}{7}$	
x_1	0	$\frac{21}{13}$	1	0	$-\frac{7}{13}$	$\frac{1}{13}$	$\frac{7}{13}$	$-\frac{1}{13}$	
x_2	0	$\frac{10}{13}$	0	1	$\frac{1}{13}$	$-\frac{2}{13}$	$-\frac{1}{13}$	$\frac{2}{13}$	
			0	0	0	0	0	1	

Since $\Delta_j \geq 0 \;\; \forall j$, an optimum basic feasible solution to the auxiliary LPP has been attained.

$$x_1 = \frac{21}{13}, \;\; x_2 = \frac{10}{13}, \;\; x_3 = x_4 = a_1 = a_2 = 0 .$$

By step 2 (iii) proceed to Phase II.

Phse II

Remove column of a_1 and a_2 from last simplex table. Starting simplex table will be last simplex table of phase I. Whereas objective function is a function given in original problem.

Max $z = -x_1 - x_2$

B.V.	c_B	c_j / x_B	-1 / x_1	-1 / x_2	0 / x_3	0 / x_4
x_1	-1	$\frac{21}{13}$	1	0	$-\frac{7}{13}$	$\frac{1}{13}$
x_2	-1	$\frac{10}{13}$	0	1	$\frac{1}{13}$	$-\frac{2}{13}$
		$\Delta_j \rightarrow$	0	0	$\frac{6}{13}$	$\frac{1}{13}$

Since $\Delta_j \geq 0 \;\; \forall j$, an optimum BFS has been attained.

$$x_1 = \frac{23}{13}, \;\; x_2 = \frac{10}{13}$$

$$\text{Min } z = x_1 + x_2$$

$$= \frac{23}{13} + \frac{10}{13} = \frac{33}{13}$$

Example 2

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

Subject to $-2x_1 + x_2 + 3x_3 = 2$

$2x_1 + 3x_2 + 4x_3 = 1, \;\; x_1, x_2, x_3 \geq 0$

Solution :

Though constraints are in the form of equations coefficient matrix do not contain identity matrix and therefore one has to introduce artificial variables and solve by two phase simplex method.

Phase I

$$\text{Max } z* = -a_1 - a_2$$

s.t. $-2x_1 + x_2 + 3x_3 + a_1 = 2$

$2x_1 + 3x_2 + 4x_3 + a_2 = 1, \quad x_1, x_2, x_3, a_1, a_2 \geq 0$

B.V.	c_B	c_j / x_B	0 / x_1	0 / x_2	0 / x_3	– 1 / a_1	– 1 / a_2	Min ratio
a_1	– 1	2	– 2	1	3	1	0	$\frac{21}{3}$
$\leftarrow a_2$	– 1	1	2	3	$\boxed{4}$	0	1	$\boxed{\frac{1}{4}}$
		Δ_j	0	– 4	– 7 ↑	0	0	
a_1	– 1	$\frac{5}{4}$	$-\frac{7}{2}$	$-\frac{5}{4}$	0	1	$-\frac{3}{4}$	
x_3	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	1	0	$\frac{1}{4}$	
			$\frac{7}{2}$	$\frac{5}{4}$	0	0	$\frac{7}{4}$	

Since all $\Delta_j \geq 0$, an optimum BFS to the LPP has been attained.

But $\quad \text{Max } z* = -a_1 - a_2 = -\frac{5}{4} < 0$

Therefore (by step 2(i) of phase I) original problem does not possess any feasible solution.

Alternatively example 1 can be solved as follows.

Example 2.11

Solve the following L. P. problem

Min. $z = x_1 + x_2$

subject to $2x_1 + x_2 \geq 4$

$x_1 + 7x_2 \geq 7$, $x_1, x_2 \geq 0$

Solution :

First we convert the problem of minimization to the maximization problem by taking the objective function $z' = -z$ i. e.

Max. $z' = -z = -x_1 - x_2$

Introduction of surplus variables x_3 and x_4 in the given inequalities yields.

$$2x_1 + x_2 - x_3 = 4$$

$$x_1 + 7x_2 - x_4 = 7$$

Here we can not get the starting B. F. S. so we introduce the artificial variables (positive) x_5 and x_6.

The above equations may be written as

$$2x_1 + x_2 - x_3 + x_5 = 4$$

$$x_1 + 7x_2 - x_4 + x_6 = 7$$

The problem will be solved in two phases.

Phase : 1

This phase consists of the removal of artificial variables.

Taking $x_1 = x_2 = x_3 = 0, x_4 = 0$ we get $x_5 = 4$ and $x_6 = 7$.

We construct the first table as follows.

Table 1

	x_B	Y_1	Y_2	Y_3	Y_4	$A_1(\beta_1)$	$A_2(\beta_2)$
A_1	4	2	1	-1	0	1	0
A_2	7	1	7	0	-1	0	1→
	x_j	0	0	0	0	4	7
			↑				↓

First we shall remove the artificial variable vector (columns) A_1 and A_2 from the basis matrix. In place of artificial variable vector the entering vector should be so chosen that the revised solution is non negative (B. F.) solution.

We can remove A_2 and introduce y_2 in its place in the basic matrix. For this we divide the second row by 7 and then subtract it from the first row. Thus we get the following table.

It maybe seen that if y_1, y_3, y_4 is entered in place of A_2 then the revised solution is not non negative. So we can not enter either of them, in place of A_2. Since artificial variable x_6 becomes zero, we forget about A_2 for ever and will not consider it in any other table.

	x_B	Y_1	Y_2 (β_2)	Y_3	Y_4	A_1 (β_1)	A^*_2
A_1	3	$\frac{13}{7}$	0	-1	$\frac{1}{7}$	1	$-\frac{1}{7}\rightarrow$
Y_2	1	$\frac{1}{7}$	1	0	$-\frac{1}{7}$	0	$\frac{1}{7}$
	x_j	0	1	0	0	3	0
		$\uparrow$					

Now we proceed to remove A_1 and introduce y_1 in its place in basic matrix. For this we mutiply first row by $\frac{7}{13}$ and subtract $\frac{1}{7}$ times of this new row from the second row. Thus we get the following table.

Table 2

		x_B	Y_1 (β_1)	Y_2 (β_2)	Y_3	Y_4	A_1*
	y_1	$\frac{21}{13}$	1	0	$-\frac{7}{13}$	$\frac{1}{13}$	$\frac{7}{13}$
	y_2	$\frac{10}{13}$	0	1	$\frac{1}{13}$	$-\frac{14}{91}$	$-\frac{1}{13}$
		x_j	$\frac{21}{13}$	$\frac{10}{13}$	0	0	0

Since the artificial variable x_5 becomes zero we forget about A_1 and will not consider it again.

Thus we get the following solution in phase (1)

$$x_1 = \frac{21}{13}, x_2 = \frac{10}{13}, x_3 = 0, x_4 = 0$$

Which is the B. F. S. with which we proceed to get the optimal solution by simplex method.

Phase (II)

The starting simplex table

B	c_B	x_B	Y_1 (β_1)	Y_2 (β_2)	Y_3	Y_4	Min. ratio
Y_1	-1	$\frac{21}{13}$	1	0	$-\frac{7}{13}$	$\frac{1}{13}$	
Y_2	-1	$\frac{10}{13}$	0	1	$\frac{1}{13}$	$-\frac{14}{91}$	
$z' = c_B x_B$		x_j	$\frac{21}{13}$	$\frac{10}{13}$	0	0	
$= -\frac{31}{13}$		c_j	- 1	-1	0	0	
		Δ_j	x	x	$-\frac{6}{13}$	$-\frac{7}{91}$	

$$\Delta_3 = c_3 - c_B y_3 = 0 - (-1, -1)\left(-\frac{7}{13}, \frac{1}{13}\right) = -\frac{6}{13}$$

$$\Delta_4 = c_4 - c_B y_4 = 0 - (-1, -1)\left(\frac{1}{13}, -\frac{14}{91}\right) = -\frac{7}{91}$$

Since Δ_j s for all zero variables are negative so the solution is optimal.

Therefore the optimal solution is

$$x_1 = \frac{21}{13}, x_2 = \frac{10}{13} \text{ and}$$

$$\text{Min. } z = -\text{ max. } z' = \frac{31}{13}$$

Example 2.12

Solve the following L. P. Problem

Max. $z = x_1 + 2x_2 + 3x_3 - x_4$

Subject to $x_1 + 2x_2 + 3x_3 = 15$

$2x_1 + x_2 + 5x_3 = 20$

$x_1 + 2x_2 + x_3 + x_4 = 10$

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

Solution :

In order to get an identity matrix we need two more columns of the unit matrix as one column of unit matrix (coeff. of x_4) is present in the constraints.

Thus we need only two artificial variables in the first two contraints. Introducing the artificial variables x_5 and x_6 we have,

$x_1 + 2x_2 + 3x_3 + 0.x_4 + x_5 = 15$

$2x_1 + x_2 + 5x_3 + 0.x_4 + x_6 = 20$

$x_1 + 2x_2 + x_3 + x_4 = 10$

Phase (1)

Taking $x_1 = x_2 = x_3 = 0$ we get $x_4 = 10, x_5 = 15, x_6 = 20$.

First table

	x_B	Y_1	Y_2	Y_3	Y_4 (β_3)	A_1 (β_1)	A_2 (β_2)
A_1	15	1	2	3	0	1	0
A_2	20	2	1	5	0	0	1 $\rightarrow$
Y_4	10	1	2	1	1	0	0
	x_j	0	0	0	10	15	20
				$\uparrow$			$\downarrow$

First we remove the artificial variable vector A_2 and introduce y_3 in its place.

For this we divide the second row by 5 and subtract it 3 and one times of it from the first and third rows respectively.

Thus we get the following table.

Second Table

	x_B	Y_1	Y_2	Y_3	Y_4 (β_3)	A_1 (β_1)	A_2
A_1	3	$-\frac{1}{5}$	$\frac{7}{5}$	0	0	1	$-\frac{3}{5}\rightarrow$
Y_3	4	$\frac{2}{5}$	$\frac{1}{5}$	1	0	0	$\frac{1}{5}$
Y_4	6	$\frac{3}{5}$	$\frac{9}{5}$	0	1	0	$-\frac{1}{5}$
	x_j	0	0	4	6	3	0

$\uparrow$

Now the artificial variable $x_6 = 0$ so we shall not consider it again. Again we remove the artificial variable vector A_1 and introduce y_2 in its place. For this we multiply first row by $\frac{5}{7}$ and then subtract its $\frac{1}{5}$ and $\frac{9}{5}$ times from the second and third rows.

Thus we get the following table.

	x_B	Y_1	Y_2 (β_1)	Y_3 (β_2)	Y_4 (β_3)	A_1	
Y_2	$\frac{15}{7}$	$-\frac{1}{7}$	1	0	0	$\frac{5}{7}$	
Y_3	$\frac{25}{7}$	$\frac{3}{7}$	0	1	0	$-\frac{1}{7}$	
Y_4	$\frac{15}{7}$	$\frac{6}{7}$	0	0	1	$-\frac{9}{7}$	
	x_j	0	$\frac{15}{7}$	$\frac{25}{7}$	$\frac{15}{7}$	0	

Here the artifical variable $x_5 = 0$. We shall not consider it in the other table.

Thus we get the following B. F. S. with which we can proceed, for the optimal solution by simplex method.

$$x_1 = 0, x_2 = \frac{15}{7}, x_3 = \frac{25}{7}, x_4 = \frac{15}{7}$$

Phase (II)

The starting simplex table is as follows.

B	c_B	x_B	Y_1	Y_2 β_1	Y_3 β_2	Y_4 β_3	min ratio $\frac{x_B}{y_1}$
Y_2	2	$\frac{15}{7}$	$-\frac{1}{7}$	1	0	0	- 14 (neg.)
Y_3	3	$\frac{25}{7}$	$\frac{3}{7}$	0	1	0	$\frac{25}{3}$
Y_4	- 1	$\frac{15}{7}$	$\frac{6}{7}$	0	0	1	$\frac{5}{2}$(min)$\rightarrow$
		x_j	0	$\frac{15}{7}$	$\frac{25}{7}$	$\frac{15}{7}$	
		c_j	1	2	3	-1	
		Δ_j	$\frac{6}{7}$	x	x	x	
			$\uparrow$			$\downarrow$	

$$\Delta_1 = c_1 - c_B Y_1 = 1 - (2, 3, -1)\left(-\frac{1}{7}, \frac{3}{7}, \frac{6}{7}\right) = \frac{6}{7}$$

Since all Δ_j are not less than or equal to zero so the solution is not optimal.

Here y_1 is the incoming vector and by minimum ratio rule we find that y_4 is the outgoing vector.

Therefore key element $y_{31} = \frac{6}{7}$.

In order to bring y_1 in place of y_4 multiply third row by $\frac{7}{6}$ and then add its $\frac{1}{7}$ times in first row and subtract $\frac{3}{7}$ times from the second row.

The second simplex table is as follows.

B	c_B	x_B	Y_1 β_3	Y_2 β_1	Y_3 β_2	Y_4	min ratio
Y_2	2	$\frac{5}{2}$	0	1	0	$\frac{1}{6}$	
Y_3	3	$\frac{5}{2}$	0	0	1	$-\frac{1}{2}$	
Y_1	1	$\frac{5}{2}$	1	0	0	$\frac{7}{6}$	
$z = c_B x_B$		x_j	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{2}$	0	
= 15		c_j	1	2	3	-1	
		Δ_j	x	x	x	-1	

$$\Delta_4 = c_4 - c_B\, y_4 = 1 - (2,3,1)\left(\frac{1}{6}, -\frac{1}{2}, \frac{7}{6}\right) = -1$$

Since Δ_4 for zero variable is negative so the solution is optimal.

Optimal solution is

$$x_1 = \frac{5}{2}, x_2 = \frac{5}{2}, x_3 = \frac{5}{2} \text{ and } \max. z = 15 .$$

Example 2.13

Using simplex algorithm solve the L. P. problem

Min. $z = 4\,x_1 + 8\,x_2 + 3\,x_3$

Subject to $\quad x_1 + x_2 \geq 2$

$2\,x_1 + x_3 \geq 5$

$x_1, x_2, x_3 \geq 0$

Solution :

First we convert the problem of minimization to maximization problem by taking $z' = -z$.

$$\max.\ z' = -z = -4\,x_1 - 8\,x_2 - 3\,x_3$$

Introducing the surplus variables x_4, x_5 the equations obtained are

$x_1 + x_2 + 0.x_3 - x_4 = 2$

$2x_1 + 0x_2 + x_3 - x_5 = 5$

The columns of x_2 and x_3 form a unit matrix. Therefore there is no need to introduce the artificial variables.

Taking $x_1 = 0, x_4 = 0, x_5 = 0$ we have

$x_2 = 2, x_3 = 5$ as starting B. F. S.

Starting simplex table

B	c_B	x_B	Y_1 (α_1)	Y_2 (β_1)	Y_3 (β_2)	Y_4 (α_4)	Y_5 (α_5)	min ratio $\frac{x_B}{y_1}$
Y_2	- 8	2	1	1	0	-1	0	2 min $\rightarrow$
Y_3	- 3	5	2	0	1	0	-1	$\frac{5}{2}$
		x_j	0	2	5	0	0	
		c_j	-4	-8	-3	0	0	
		Δ_j	10	x	x	-8	-3	
			$\uparrow$	$\downarrow$				

$$\Delta_1 = c_1 - c_B y_1 = -4 - (-8, -3)(1, 2) = 10$$

$$\Delta_4 = c_4 - c_B y_4 = 0 - (-8, -3)(-1, 0) = -8$$

$$\Delta_5 = c_5 - c_B y_5 = 0 - (-8, -3)(0, -1) = -3$$

Since all Δ_j s are not less than or equal to zero so the solution is not optimal.

$\text{Max.}\Delta_j = 10 = \Delta_1$

$\therefore$ Entering vector is $\alpha_1 (= y_1)$ and by minimum ratio rule we find that outgoing vector is $\beta_1 (= y_2)$.

Therefore key element is $y_{11} = 1$.

In order to bring α_1 in place of β_1 we subtract 2 times of the first row from the second row.

Second simplex table is

B	c_B	x_B	Y_1 (β_1)	Y_2	Y_3 (β_2)	Y_4	Y_5	min ratio $\frac{x_B}{y_4}$
Y_1	-4	2	1	1	0	-1	0	-2 neg.
Y_3	- 3	1	0	-2	1	2	-1	$\frac{1}{2}$ min $\rightarrow$
		x_j	2	0	1	0	0	
		c_j	- 4	- 8	- 3	0	0	
		Δ_j	x	- 10	x	2	-3	
					$\downarrow$	$\uparrow$		

$$\Delta_2 = c_2 - c_B y_2 = -8 - (-4, -3)(1, -2) = -10$$

$$\Delta_4 = c_4 - c_B y_4 = 0 - (-4, -3)(-1, 2) = 2$$

$$\Delta_5 = c_5 - c_B y_5 = 0 - (-4, -3)(0, -1) = -3$$

Since all Δ_j s are not less than or equal to zero this solution is not optimal.

Since Max $\Delta_j = \Delta_4$, the incoming vector is y_4 and by the minimum ratio rule we find that the outgoing vector is $y_3 (= \beta_2)$.

Key element = 2

In order to bring y_4 in place of y_3 we divide the second row by 2 and then add it to the first row.

Third simplex table is

B	c_B	x_B	Y_1 (β_1)	Y_2	Y_3	Y_4 (β_2)	Y_5	min ratio
Y_1	- 4	$\frac{5}{2}$	1	0	$\frac{1}{2}$	0	$-\frac{1}{2}$	
Y_4	0	$\frac{1}{2}$	0	-1	$\frac{1}{2}$	1	$-\frac{1}{2}$	
		x_j	$\frac{5}{2}$	10	0	$\frac{1}{2}$	0	
		c_j	-4	-8	-3	0	0	
		Δ_j	x	-8	-1	x	-2	

$$\Delta_2 = c_2 - c_B y_2 = -8 - (-4,0)(0,-1) = -8$$

$$\Delta_3 = c_3 - c_B y_3 = -3 - (-4,0)\left(\frac{1}{2},\frac{1}{2}\right) = -1$$

$$\Delta_5 = c_5 - c_B y_5 = 0 - (-4,0)\left(-\frac{1}{2},-\frac{1}{2}\right) = -2$$

Since all Δ_j's are negative, this solution is optimal.

So the optimal solution is

$$x_1 = \frac{5}{2}, x_2 = 0, x_3 = 0$$

and $\min z = -(\max . z') = 10$

DUALITY THEOREM

INTRODUCTION

We have considered the L. P. Problems in which by minimum ratio rule we get only one vector to be deleted from the basis. But there are the L. P. Problems where we get more than one vector which may be deleted from the basis.

Thus if $\min\left\{\frac{x_{Bi}}{y_{ik}}, y_{ik} > 0\right\}$ (α_k is incoming vector)

occurs at $i = i_1, i_2, .. \; i_s$

i. e. minimum occurs for more than one value of i then the problem is to select the vector to be deleted from the basis (If we choose one vector say β_i (i is one of $i_1, i_2, ..., i_s$) and delete it from the basis then the next solution may be a degenerate B. F. S. Such problem is called problem of degeneracy.

It is observed that when the simplex method is applied to a degenerate B. F. S. to get a new B. F. S., the value of the objective function may remain unchanged i. e. the value of the objective function is not improved.

The procedure for such problems of degeneracy is as follows.

Let $\min_i \left\{\frac{(X_{Bi})}{y_{ik}}, y_{ik} > 0\right\}$ occur at $i = i_1, i_2, ..., i_s$

where $\alpha_k = y_k$ is the incoming vector.

Let $I_1 = \{i_1, i_2, ..., i_s\}$

1) Renumber the columns of the table starting with the columns in the basis. Let $\overline{y}_1, \overline{y}_2, ...$ etc. be the new numbers of columns. Let $\overline{y}_t$ be the new number of entering vector y_k i. e. $y_k = \overline{y}_t$.

2) Calculate $\min\left\{\frac{\overline{y}_{i1}}{y_{ik}}\right\} \forall i \in I_1$. If minimum is unique then delete the corresponding vector from the basis.

If minimum is not unique then proceed to the next step.

3) Calculate $\min_i \left\{ \frac{\bar{y}_{i2}}{y_{ik}} \right\} \forall\, i \in I_2$ where I_2 is the set of all those values of $i \in I_1$, for which there is a tie in I_2. Clearly $I_2 \subset I_1$.

In this case if minimum is unique then correspondng vector is deleted from the basis. If in this case also, minimum is not unique proceed to the next step.

4) Compute $\min_i \left\{ \frac{\bar{y}_{i3}}{y_{ik}} \right\} \forall\, i \in I_3$ where I_3 is the set of those values of $i \in I_2$ for which there is a tie in (3) clearly $I_3 \subset I_2 \subset I_1$.

Proceeding in this way we can get a unique minimum value of i i. e. the unique vector to be deleted from the basis.

Example 1

Solve the L. P. Problem

Max. $$z = \frac{3}{4}x_1 - 150\,x_2 + \frac{1}{50}x_3 - x_4$$

Subject to $$\frac{1}{4}x_1 - 60\,x_2 - \frac{1}{25}x_3 + 9\,x_4 \le 0$$

$$\frac{1}{2}x_1 - 90\,x_2 - \frac{1}{50}x_3 + 3\,x_4 \le 0$$

$$x_3 \le 1$$

and $x_1, x_2, x_3, x_4 \ge 0$

Solution :

Introducing the slack variables in the constraints we get the following equalities

$$\frac{1}{4}x_1 - 60\,x_2 - \frac{1}{25}x_3 + 9\,x_4 + x_5 = 0$$

$$\frac{1}{2}x_1 - 90\,x_2 - \frac{1}{50}x_3 + 3\,x_4 + x_6 = 0$$

$$x_3 + x_7 = 1$$

Taking $x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 0$ we have

$$x_5 = 0, x_6 = 0, x_7 = 1$$

Which is the starting B. F. S.

Starting simplex table

B	c_B	x_B	$\bar{y}_4$ y_1	$\bar{y}_5$ y_2	$\bar{y}_6$ y_3	$\bar{y}_7$ y_4	$\bar{y}_1$ $y_5\,(\beta_1)$	$\bar{y}_2$ $y_6\,(\beta_2)$	$\bar{y}_3$ $y_7\,(\beta_3)$	Min ratio $\frac{x_B}{y_1}$
y_5	0	0	$\frac{1}{4}$	-60	$-\frac{1}{25}$	9	1	0	0	0
y_6	0	0	$\frac{1}{2}$	- 90	$-\frac{1}{50}$	3	0	1	0	0
y_7	0	1	0	0	1	0	0	0	1	-
$z = c_B\, x_B$		x_j	0	0	0	0	0	0	1	
= 0		c_j	$\frac{3}{4}$	-150	$\frac{1}{50}$	-6	0	0	0	
		Δ_j	$\frac{3}{4}$	-150	$\frac{1}{50}$	-6	0	0	x	

↑ (under y_1) ↓ (under y_6)

$$\Delta_1 = c_1 - c_B\, y_1 = \frac{3}{4} - (0,0,0)\left(\frac{1}{4},\frac{1}{2},0\right) = \frac{3}{4}$$

$$\Delta_2 = c_2 - c_B\, y_2 = -150 - (0,0,0)(-60,-90,0) = -150$$

$$\Delta_3 = c_3 - c_B\, y_3 = \frac{1}{50} - (0,0,0)\left(-\frac{1}{25},-\frac{1}{50},1\right) = \frac{1}{50}$$

$$\Delta_4 = c_4 - c_B\, y_4 = -6 - (0,0,0)(9,3,0) = -6$$

Since all Δ_j are not less than as equal to zero therefore the solution is not optimal

and $\max \Delta_j = \frac{3}{4} = \Delta_1$

Therefore incoming, vector is y_1 and

$\min_i \left\{ \frac{x_{Bi}}{y_{il}}, y_{ij} > 0 \right\}$ is not unique.

This minimum is o and occurs for i = 1 and i = 2.

This problem is a problem of degeneracy.

Therefore to select the vector to be deleted from the basic we proceed as follows.

1) First of all we renumber the columns of above table as follows.

Let $\overline{y}_1 = y_5, \overline{y}_2 = y_6, \overline{y}_3 = y_7$

$\overline{y}_4 = y_1 = \overline{y}_5 = y_2, \overline{y}_6 = y_3, \overline{y}_7 = y_4$

2) Since minimum ratio occurs for
i = 1 and i = 2 it follows that

$I_1 = \{1,2\}$

Incoming vector is $y_1 = \overline{y}_4, k = 4$ for i = 1, 2

$$\min_{i \in I_1} \left\{ \frac{\overline{y}_{i1}}{\overline{y}_{i4}} \right\} = \min \left\{ \frac{\overline{y}_{11}}{\overline{y}_{14}}, \frac{\overline{y}_{21}}{\overline{y}_{24}} \right\}$$

$$= \min i \left\{ \frac{1}{\left(\frac{1}{4}\right)}, \frac{0}{\left(\frac{1}{2}\right)} \right\} = \min\{4, 0\}$$

$$= 0 = \frac{\overline{y}_{21}}{\overline{y}_{24}}$$

This minimum is unique and occur for i = 2. Therefore the vector to be deleted (i. e. the outgoing vector) from the basis is $\overline{y}_2 (= \beta_2) = y_6$.

Therefore key element is $y_{21} = \frac{1}{2}$.

Therefore in older to bring y_1 in place of y_6 we divide the second row by $\frac{1}{2}$ and then subtract $\frac{1}{4}$ times of this row from the first row.

Second simple table

B	c_B	x_B	y_1 (β_2)	y_2	y_3	y_4	y_5 (β_1)	y_6	y_7 (β_3)	Min ratio $\frac{x_B}{y_3}$
y_5	0	0	0	-15	$-\frac{3}{100}$	$\frac{15}{2}$	1	$-\frac{1}{2}$	0	--
y_1	$\frac{3}{4}$	0	1	-180	$-\frac{1}{25}$	6	0	2	0	--
y_7	0	1	0	0	1	0	0	0	1	1 (Min) →
		x_j	0	0	0	0	0	0	1	
		c_j	$\frac{3}{4}$	-150	$\frac{1}{50}$	-6	0	0	0	
		Δ_j	0	-15	$\frac{1}{20}$	$-\frac{21}{2}$	0	$-\frac{3}{2}$	x	

↑ Incoming Vector (under y_3) ↓ Outgoing Vector (under y_7)

$$\Delta_2 = c_2 - c_B\, y_2 = -150 - \left(0, \frac{3}{4}, 0\right)(-15, -180, 0) = -15$$

$$\Delta_3 = c_3 - c_B\, y_3 = +\frac{1}{50} - \left(0, \frac{3}{4}, 0\right)\left(-\frac{3}{100}, -\frac{1}{25}, 1\right) = \frac{1}{20}$$

$$\Delta_4 = c_4 - c_B\, y_4 = -6 - \left(0, \frac{3}{4}, 0\right)\left(\frac{15}{2}, 6, 0\right) = -\frac{21}{2}$$

$$\Delta_6 = c_6 - c_B\, y_6 = -\frac{3}{2}$$

Since all Δ_j are not less than or equal to zero therefore the solution is not optimal.

$$\text{Max. } \Delta_j = \frac{1}{20} = \Delta_3$$

Therefore incoming vector is $\frac{1}{3}$ and by minimum ratio rule we find that the outgoing vector is $y_7\left(=\beta_2\right)$.

(In considering $\frac{x_B}{y_B}$ we need not consider the ratios $\frac{x_{B1}}{y_{13}}$ and $\frac{x_{B2}}{y_{23}}$ since $y_{13}=-\frac{3}{100}$ and $y_{23}=-\frac{1}{25}$ are both negative.)

Therefore key element $y_{33}=1$.

In order to bring y_3 in place at $y_7\left(\beta_3\right)$ we add $\frac{3}{100}$ and $\frac{1}{25}$ times of the third row in the first and second rows respatively.

The third simplex table

B	c_B	x_B	y_1 (β_2)	y_2	y_3 (β_3)	y_4	y_5 (β_1)	y_6	y_7	
y_5	0	$\frac{3}{100}$	0	-15	0	$\frac{15}{2}$	1	$-\frac{1}{2}$	$\frac{3}{100}$	
y_1	$\frac{3}{4}$	$\frac{1}{25}$	1	-180	0	6	0	2	$\frac{1}{25}$	
y_3	$\frac{1}{50}$	1	0	0	1	0	0	0	1	
		x_j	$\frac{1}{25}$	0	1	0	$\frac{3}{100}$	0	0	
		c_j	$\frac{3}{4}$	-150	$\frac{1}{50}$	-6	0	0	0	
		Δ_j	x	-15	x	$-\frac{21}{2}$	x	$-\frac{3}{2}$	$-\frac{1}{20}$	

$$\Delta_2=c_2-c_B\,y_2=-150-\left(0,\frac{3}{4},\frac{1}{50}\right)(-15,-180,0)=-15$$

$$\Delta_4=c_4-c_B\,y_4=-6-\left(0,\frac{3}{4},\frac{1}{50}\right)\left(\frac{15}{2},6,0\right)=-\frac{21}{2}$$

$$\Delta_6 = c_6 - c_B\, y_6 = 0 - \left(0, \frac{3}{4}, \frac{1}{50}\right)\left(-\frac{1}{2}, 2, 0\right) = -\frac{3}{2}$$

$$\Delta_7 = c_7 - c_B\, y_7 = 0 - \left(0, \frac{3}{4}, \frac{1}{50}\right)\left(\frac{3}{100}, \frac{1}{25}, 1\right) = -\frac{1}{20}$$

Since all $\Delta_j \le 0$ thesefore the solution is optimal and the optimal solution is

$$x_1 = \frac{1}{25}, x_2 = 0, x_3 = 1, x_4 = 0$$

and $\quad$ $\text{Max. } z = \frac{1}{20}$

DUALITY

Introduction

Every L. P. Problem is associated with another L. P. Problem called the dual of the problem. Consider a L. P. Problem

Max. $\quad z = c_1 x_1 + c_2 x_2 + ... + c_n x_n$

Subject to $\quad a_{11} x_1 + a_{12} x_2 + ... + a_{1n} x_n \le b_1$

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

....

$a_{m1} x_1 + a_{m2} x_2 + ... + a_{mn} x_n \le b_m$ .... (i)

and $\quad x_1, x_2, ..., x_n \ge 0$,

where the signs of all parameters a, b, c are orbitary.

Then the dual of this problem is defined as

Mini $\quad z^* = b_1 w_1 + b_2 w_2 + ... + b_m w_m$

Subject to $\quad a_{11} w_1 + a_{21} w_2 + ... + a_{m1} w_m \ge c_1$

$a_{12} w_1 + a_{22} w_2 + ... + a_{m2} w_m \ge c_2$

....

and $\quad a_{1n} w_1 + a_{2n} w_2 + ... + a_{mn} w_n \ge c_n$

and $w_1, w_2, ..., w_m \ge 0$

where $w_1, w_2, ..., w_m$ are called the dual variables.

Also problem (1) is called the primal problem.

In a matrix notation a L. P. Problem is

Max. $z = \bar{c}\,\bar{x}$

Subject to $A\,\bar{x} \le \bar{b}$

and $\bar{x} \ge 0$

and its dual is defined as

Min $z^* = b'\,w$

Subject to $A'\,w \ge c'$

and $w \ge 0$

Where $w = \begin{bmatrix} w_1 \\ w_2 \\ \cdots \\ w_m \end{bmatrix}$

and $A', \bar{b}', \bar{c}'$ are the transposes of the matrices $A, \bar{b}$ and $\bar{c}$ respectively.

It is obvious from the definition that the dual of the dual is the primal itself.

It is important to note that we can write the dual of a problem if all its constraints involve the sign $\le$.

If the constraint has a sign $\ge$ then multiply both the sides by - 1 and makes the sign $\le$.

If the constraint has a sign = for ex. $\sum_{j=1}^{n} a_{ij}\,x_j = b_i$ .... (3)

then we can replace it by two constraints involving two inequalities i. e.

$$\sum_{j=1}^{n} a_{ij}\,x_j \le b_i \qquad \text{........... (4)}$$

$$\sum_{j=1}^{n} a_{ij}\,x_j \ge b_i \qquad \text{........... (5)}$$

5) may be written as

$$-\sum_{j=1}^{n} a_{ij}\,x_j \le -b_i$$

Standard form of the primal

The L. P. Problem is in standard primal form if

1) It is a problem of maximization and

2) All the constraints involve the sign $\leq$.

Relationship between two problems (Primal and dual)

The two problems (primal and the dual) are related to each other in the following manner.

1) If one is a maximization problem then the other is a minimization problem.

2) If one of them has a finite optimal solution then the other problem also has a finite optimal solution.

3) From the final simplex table of one problem the solution of the other an be read from the Δ_j row below the columns of slack and surplus variables as follows.

   The Δ_j's $\left(\Delta_j = c_j - z_j = c_j - c_B\, y_j\right)$ with the sign changed for the slack vectors in the optimal (final) simplex table for the primal are the values of the corresponding optimal dual variables in the final simplex table for the dual problem.

4) The optimal values of the objective functions in both the problems are the same that is $\text{Max.} Z_x = \text{Min}\, Z_w$.

5) If one problem has an unbounded solution then other has no feasible solution.

Example 2

Write the dual of the problem

Mini. $z = 3x_1 + x_2$

Subject to $2x_1 + 3x_2 \geq 2$

$x_1 + x_2 \geq 1$

and $x_1, x_2 \geq 0$

Solution :

First we write the problem in standard primal form as follows.

Max. $z' = -3x_1 - x_2$ where $z' = -z$

Such that $-2x_1 - 3x_2 \leq -2$

and $-x_1 - x_2 \leq -1$

and $x_1, x_2 \geq 0$

which may be written as

$$\text{Max } z' = [-3, -1]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Such that $\begin{bmatrix} -2 & -3 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \le \begin{bmatrix} -2 \\ -1 \end{bmatrix}$

and $x_1, x_2 \ge 0$

The dual of the given problem is given by

Mini. $z^* = [-2, -1]\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$

such that $\begin{bmatrix} -2 & -1 \\ -3 & -1 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \ge \begin{bmatrix} -3 \\ -1 \end{bmatrix}$

and $w_1, w_2 \ge 0$

or mini. $z^* = -z\,w_1 - w_2$

such that $-2w_1 - w_2 \ge -3$

$-3w_1 - w_2 \ge -1$

Example 3

Write the dual of the problem

miz. $z = 2x_2 + 5x_3$

such that $x_1 + x_2 \ge 2$

$2x_1 + x_2 + 6x_3 \le 6$

$x_1 - x_2 + 3x_3 = 4$

and $x_1, x_2, x_3 \ge 0$.

Solution :

First we write the given problem in standard primal form as follows.

1) The objective function is changed from minimization to maximization.

i. e. Max $z' = -2x_2 - 5x_3$ where $z' = -z$

2) The sign of first constraint is changed to $\le$ by multiplying both sides by - 1 and

3) The third constraint is replaced by two constraints.

$x_1 - x_2 + 3x_3 \le 4$

and $\quad x_1 - x_2 + 3x_3 \geq 4$

The second may be written as

$$-x_1 + x_2 - 3x_3 \leq -4$$

Thus the given problem in standard primal form is as follows.

Max. $z' = 0x_1 - 2x_2 - 5x_3$

subject to $\quad -x_1 - x_2 \leq 2$

$$2x_1 + x_2 + 6x_3 \leq 6$$

$$x_1 - x_2 + 3x_3 \leq 4$$

$$-x_1 + x_2 - 3x_3 \leq -4$$

and $x_1, x_2, x_3 \geq 0$

i. e. $\text{Max. } z' = [0, -2, -5]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$

such that $\begin{bmatrix} -1 & -1 & 0 \\ 2 & 1 & 6 \\ 1 & -1 & 3 \\ -1 & 1 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \leq \begin{bmatrix} 2 \\ 6 \\ 4 \\ -4 \end{bmatrix}$

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

Therefore the dual of the given problem is given by

$$\text{Mini } z^* = [2, 6, +4, -4]\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix}$$

such that $\begin{bmatrix} -1 & 2 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 0 & 6 & 3 & -3 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} \geq \begin{bmatrix} 0 \\ -2 \\ -5 \end{bmatrix}$

and $w_1, w_2, w_3, w_4 \geq 0$

or Min. $z^* = 2w_1 + 6w_2 + 4w_3 - 4w_4$

such that $-w_1 + 2w_2 + w_3 - w_4 \geq 0$

$-w_1 + w_2 - w_3 + w_4 \geq -2$

$0w_1 + 6w_2 + 3w_3 - 3w_4 \geq -5$

and $w_1, w_2, w_3, w_4 \geq 0$

Example 4

Apply the simplex method to solve the following

Max. $z = 30x_1 + 23x_2 + 29x_3$

s. t. $6x_1 + 5x_2 + 3x_3 \leq 26$

$4x_1 + 2x_2 + 5x_3 \leq 7$

and $x_1, x_2, x_3 \geq 0$ .... (1)

Also read the solution of the dual of the above problem from the final table.

Solution :

Introducing the slack variables x_4 and x_5 ,we have

$6x_1 + 5x_2 + 3x_3 + x_4 = 26$

$4x_1 + 2x_2 + 5x_3 + x_5 = 7$

Taking $x_1 = x_2 = x_3 = 0$ we have $x_4 = 26$ and $x_5 = 7$,

which is the starting B. F. S.

Starting Simplex Table

B	c_B	x_B	y_1 (α_1)	y_2 (α_2)	y_3 (α_3)	y_4 (β_1)	y_5 (β_2)	Min. Ratio. $\frac{X_B}{y_1}$
y_4	0	26	6	5	3	1	0	$\frac{13}{3}$
y_5	0	7	4	2	5	0	1	$\frac{7}{4}\rightarrow$Min
$z=c_Bx_B$		x_j	0	0	0	26	7	
$= 0$		c_j	30	23	29	0	0	
		Δ_j	30	23	29	x	x	

$\uparrow$ Incoming $\qquad\qquad$ $\downarrow$ Outgoing

$$\Delta_1=c_1-c_By_1=30-(0,0)(0,4)=30$$

Similarly $\Delta_2=23,\Delta_3=29$

Since all Δ_j are not less than or equal to zero therefore the solution is not optimal.

Max. $\Delta_j=30=\Delta_1$

Hence $\alpha_1(=y_1)$ is incoming vector and by minimum ratio rule we find that $y_5(=\beta_2)$ is outgoing vector.

Hence the key element $y_{21}=a_{21}=4$.

Second simplex table

B	c_B	x_B	y_1 (β_2)	y_2	y_3	y_4 (β_1)	y_5	Min. Ratio. $\frac{X_B}{y_2}$
y_4	0	$\frac{31}{2}$	0	2	$-\frac{9}{2}$	1	$-\frac{3}{2}$	$\frac{31}{4}$
y_1	30	$\frac{7}{4}$	1	$\frac{1}{2}$	$\frac{5}{4}$	0	$\frac{1}{4}$	$\frac{7}{8}\rightarrow$
$z = c_B x_B$		x_j	$\frac{7}{4}$	0	0	$\frac{31}{2}$	0	
$= \frac{105}{2}$		c_j	30	23	29	0	0	
		Δ_j	x	8	$-\frac{17}{2}$	x	$-\frac{15}{2}$	
			$\downarrow$	$\uparrow$				

$$\Delta_2 = c_2 - c_B\, y_2 = 23 - (0,30)\left(2, \frac{1}{2}\right) = 8$$

$$\Delta_3 = c_3 - c_B\, y_3 = 29 - (0,30)\left(-\frac{9}{2}, \frac{5}{4}\right) = -\frac{17}{2}$$

$$= 29 - 37.5 = -8.5$$

$$\Delta_5 = c_5 - c_B\, y_5 = (0,30)\left(-\frac{3}{2}, \frac{1}{4}\right) = -\frac{15}{2}$$

Since all Δ_j are not less than or equal to zero so the solution is not optimal. Here y_2 is insuming vector and y_1 is out going vector.

The key element is $y_{22} = \frac{1}{2}$

Final simplex table

B	c_B	x_B	y_1	y_2	y_3	y_4	y_5	Min. Ratio.
y_4	0	$\frac{17}{2}$	- 4	0	$-\frac{19}{2}$	1	$-\frac{5}{2}$	
y_2	23	$\frac{7}{2}$	2	1	$\frac{5}{2}$	0	$\frac{1}{2}$	
$z=c_Bx_B$		x_j	0	$\frac{7}{2}$	0	$\frac{17}{2}$	0	
$=\frac{161}{2}$		c_j	30	23	29	0	0	
		Δ_j	16	x	$-\frac{57}{2}$	x	$-\frac{23}{2}$	

$$\Delta_1=c_1-c_B\,y_1=30-(0,23)(-4,2)=-16$$

$$\Delta_3=c_3-c_B\,y_3=29-(0,23)\left(-\frac{19}{2},\frac{5}{2}\right)=-\frac{57}{2}$$

$$\Delta_5=c_5-c_B\,y_5=0-(0,23)\left(-\frac{5}{2},\frac{1}{2}\right)=-\frac{23}{2}$$

Since all Δ_j are $\leq$ 0 the solution isoptimal.

Therefore optimal solution is

$x_1=0, x_2=\frac{7}{2}, x_3=0$ and $\max. z=\frac{161}{2}$.

To write the dual of the problem.

The given problem may be written as :

Max. $z=[30,23,29]\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$,

such that $\begin{bmatrix}6&5&3\\4&2&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}26\\7\end{bmatrix}$

and $x_1,x_2,x_3\geq 0$

Therefore the dual of the given problem is given by

$$\text{Mini } z^* = [26, 7]\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

such that $\begin{bmatrix} 6 & 4 \\ 5 & 2 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \geq \begin{bmatrix} 30 \\ 23 \\ 29 \end{bmatrix}$

where $w_1, w_2 \geq 0$

OR

Min. $z^* = 26\,w_1 + 7\,w_2$

s. t. $6\,w_1 + 4\,w_2 \geq 30$

$5\,w_1 + 2\,w_2 \geq 23$

$3\,w_1 + 5\,w_2 \geq 29$ .... (2)

where $w_1, w_2 \geq 0$

The dual problem (2) may be written as

Max. $z_1^* = -\,26\,w_1 - 7\,w_2$

s. t. $6\,w_1 + 4\,w_2 - w_3 + w_6 = 30$

$5\,w_1 + 2\,w_2 - w_4 + w_7 = 23$

$3\,w_1 + s\,w_2 - w_5 + w_8 = 29$

and $w_1, w_2, \ldots, w_8 \geq 0$

Where w_3, w_4, w_5 are surplus variables and w_6, w_9, w_8 are the artificial variables.

Now we obtain the solution of the above problem by simplex method.

~ ~ ~ ~ ~ EXERCISE ~ ~ ~ ~ ~

1) Solve the L.P. Problem

Max. $z = 3x_1 + 5x_2 + 4x_3$

Subject to, $2x_1 + 3x_2 \leq 8$

$2x_2 + 5x_3 \leq 10$

$3x_1 + 2x_2 + 4x_3 \leq 15$

and $x_1, x_2, x_3 \geq 0$

2) Solve by simplex method the following L. P. Problem

Minimize $z = x_1 - 3x_2 + 2x_3$

Subject to, $3x_1 - x_2 + 2x_3 \leq 7$

$-x_1 + 4x_2 \leq 12$

$-4x_1 + 3x_2 + 8x_3 \leq 16$

$x_1, x_2, x_3 \geq 0$

3) Solve the following L. P. Problem

Minimize $z = x_1 + x_2$

Subject to, $2x_1 + x_2 \geq 4$

$x_1 + 7x_2 \geq 7$

$x_1, x_2 \geq 0$

4) Using the simplex method to solve the following L. P. Problem

Max. $z = x_1 + 2x_2 + 3x_3 - x_4$

Subject to, $x_1 + 2x_2 + 3x_3 = 15$

$2x_1 + x_2 + 5x_3 = 20$

$x_1 + 2x_2 + x_3 + x = 10$

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

5) Using the simplex method solve the L.P. Problem

Min. $z = 4x_1 + 8x_2 + 3x_3$

Subject to, $x_1 + x_2 \geq 2$

$2x_1 + x_3 \geq 5$

$x_1, x_2, x_3 \geq 0$

6) Using the simplex method solve the following.

Max. $z = 2x_1 + 5x_2 + 7x_3$

Subject to, $3x_1 + 2x_2 + 4x_3 \leq 100$

$x_1 + 4x_2 + 2x_3 \leq 100$

$x_1 + x_2 + 3x_3 \leq 100$

$x_1, x_2, x_3 \geq 0$

7) Solve the following L. P. Problem.

Max. $z = \frac{3}{4}x_1 - 150x_2 + \frac{1}{50}x_3 - x_4$

Subject to, $\frac{x_1}{4} - 60x_2 - \frac{1}{25}x_3 + 4x_4 \leq 0$

$\frac{x_1}{2} - 90x_2 - \frac{1}{50}x_3 + 3x_4 \leq 0$

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

8) Use the simplex method to solve the following

Max. $z = 30x_1 + 23x_2 + 29x_3$

Subject to, $6x_1 + 5x_2 + 3x_3 \leq 26$

$4x_1 + 2x_2 + 5x_3 \leq 7$

and $x_1, x_2, x_3 \geq 0$

Also read the solution of the dual of the above problem from the final table.

9) Use two phase simplex method to solve.

Minimize $z = 3x_1 + 2x_2 + x_3 + x_4$

Subject to, $4x_1 + 5x_2 + x_3 - 3x_4 = 5$

$2x_1 - 3x_2 - 4x_3 + 5x_4 = 7$

$x_j \geq 0, c_j = 1, 2, 34$

10) Solve the following L.P.P.

Maximize $z = 3x_1 + 4x_2$

Subject to, $x_1 + 4x_2 \leq 8$

$x_1 - 2x_2 \leq 4$

$x_1, x_2 \geq 0$

11) Solve the following L.P.P.

Maximize $z = 2x_1 + x_2$

Subject to, $4x_1 + 3x_2 \leq 12$

$4x_1 + x_2 \leq 8$

$4x_1 - x_2 \leq 8$

$x_1, x_2 \geq 0$

12) Solve the following L.P.P.

Max. $z = 5x_1 + 3x_2$

Subject to, $x_1 + x_2 \leq 2$

$5x_1 + 2x_2 \leq 10$

$3x_1 + 8x_2 \leq 12$

$x_1 \geq 0, x_2 \geq 0$

13) Solve by L.P.P.

Max. $z = 22x_1 + 30x_2 + 25x_3$

Subject to, $2x_1 + 2x_2 \leq 100$

$2x_1 + x_2 + x_3 \leq 100$

$x_1 + 2x_2 + 2x_3 \leq 100$

$x_1, x_2, x_3 \geq 0$

14) Solve by L.P.P.

Max. $z = 5x_1 - 2x_2 + 3x_3$

Subject to, $2x_1 + 2x_2 - x_3 \geq 2$

$3x_1 - 4x_2 \leq 3$

$x_2 + 3x_3 \leq 5$

$x_1, x_2, x_3 \geq 0$

15) Solve by L.P.P.

Max. $z = x_1 + 15x_2 + 2x_3 + 5x_4$

Subject to, $3x_1 + 2x_2 + x_3 + x_4 \leq 6$

$2x_1 + x_2 + x_6 + 4x_4 \leq 4$

$2x_1 + 6x_2 - 8x_3 + 4x_4 = 0$

$x_1 + 3x_2 - 4x_3 + 3x_4 = 0$

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

❑ ❑ ❑ ❑

UNIT 03

DUALITY THEOREM

DUALITY IN LINEAR PROGRAMMING PROBLEM

Definition : Primal Problem

$$\text{Max } z_x = \sum_{i=1}^{n} c_i x_i \left(= \bar{c}^T \bar{x}\right)$$

$$\text{s.t. } A\bar{x} \le \bar{b},\ \bar{x} \ge \bar{0},\ A_{m\times n}$$

Definition : Dual Problem

$$\text{Min } z_w = \sum_{i=1}^{m} b_i w_i \left(= \bar{b}^T \bar{w}\right)$$

$$\text{s.t. } A^T\bar{w} \ge \bar{c},\ \bar{w} \ge \bar{0}$$

($\bar{x}$ has n components, $\bar{w}$ has m components)

General Rules for converting any primal to its dual

Step 1 : Convert the objective function into max form (MIn z = – (Max – z)).

Step 2 : If the constraint has '$\ge$' then multiply the constraint by (–1)

Step 3 : If the constraint has '=' then replace this constraint by two constraints '$\le$' and '$\ge$' e.g.

$x_1 + x_2 = 2 \equiv x_1 + x_2 \le 2$ and $x_1 + x_2 \ge 2$.

Step 4 : Every unrestricted variable is replaced by the difference of two non-negative variables e.g. x_1 is unrestricted.

$$x_1 = x_1^{*} - x_1^{**},\ x_1^{*}, x_1^{**} \ge 0$$

Step 1 to 4 gives Standard primal LPP.

Max $z = \bar{c}^T \bar{x}$

s.t. $A\bar{x} \le \bar{b},\ \bar{x} \ge \bar{0}$

Step 5 : Dual of above primal LPP is obtained

(i) $A \to A^T$

(ii) Interchange $\bar{b},\ \bar{c}$.

(iii) $\le \ \to \ \ge$

(iv) Minimize objective function.

Example : Max $z = 3x_1 + 2x_2$

s.t. $x_1 + 3x_2 \le 5$

$x_1 - x_2 \le 7$, $x_1, x_2 \ge 0$

Answer : Max $z_x = \bar{c}^T \bar{x} = c_1 x_1 + c_2 x_2$

s.t. $A\bar{x} \le \bar{b}$, $\bar{x} \ge \bar{0}$

Primal : Max $z = 3x_1 + 2x_2 = [3 \quad 2]\begin{bmatrix} x_1 \\ x_1 \end{bmatrix} = \bar{c}^T \bar{x}$

s.t. $\begin{bmatrix} 1 & 3 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \le \begin{bmatrix} 5 \\ 7 \end{bmatrix}$, $\bar{x} \ge \bar{0}$

Dual : Min $z_w = [5 \quad 7]\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$

$\begin{bmatrix} 1 & 1 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \ge \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, $w_1, w_2 \ge 0$

Example : Write dual of following LPP

Max $z = 2x_1 + 3x_2 - x_3$

s.t. $x_1 + x_2 - 3x_3 \le 8$

$x_1 - x_2 + x_3 \le 4$, $x_1, x_2, x_3 \ge 0$

Answer : Max $z = \begin{bmatrix} 2 & 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

s.t. $\begin{bmatrix} 1 & 1 & -3 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \le \begin{bmatrix} 8 \\ 4 \end{bmatrix}$, $x_1, x_2, x_3 \ge 0$

Dual LPP Min $z_w = 8w_1 + 4w_2$

s.t. $\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \ge \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$, $w_1, w_2 \ge 0$

Primal : Max $z = \bar{c}^T \bar{x}$

s.t. $A\bar{x} \le \bar{b}$, $\bar{x} \ge \bar{0}$

Dual : Min $z_w = \bar{b}^T \bar{w}$

s.t. $A^T \bar{w} \ge \bar{c}$, $\bar{w} \ge \bar{0}$

Example : Find the dual of the following Primal.

Min $z_x = 2x_2 + 5x_3$

s.t. $x_1 + x_2 \ge 2$, $2x_1 + x_2 + 6x_3 \le 6$

$x_1 - x_2 + 3x_3 = 4$, $x_1, x_2, x_3 \ge 0$

Answer : Max $z_x^{'} = -2x_2 - 5x_3$ $\left(z_x^{'} = -z_x\right)$

$-x_1 - x_2 \le -2$, $2x_1 + x_2 + 6x_3 \le 6$

$x_1 - x_2 + 3x_3 \le 4$, $-\left(x_1 - x_2 + 3x_3\right) \le -4$, $x_1, x_2, x_3 \ge 0$

Max $z_x^{'} = -2x_2 - 5x_3$

s.t. $-x_1 - x_2 \le -2$

$2x_1 + x_2 + 6x_3 \le 6$

$x_1 - x_2 + 3x_3 \le 4$

$-x_1 + x_2 - 3x_3 \le -4$, $x_1, x_2, x_3 \ge 0$

Standard : Max $z_x^{'} = \begin{bmatrix} 0 & -2 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

s.t. $\begin{bmatrix} -1 & -1 & 0 \\ 2 & 1 & 6 \\ 1 & -1 & 3 \\ -1 & 1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \le \begin{bmatrix} -2 \\ 6 \\ 4 \\ -4 \end{bmatrix}$, $x_1, x_2, x_3 \ge 0$

Dual : Min $z_w = -2w_1 + 6w_2 + 4w_3 - 4w_4$

s.t. $\begin{bmatrix} -1 & 2 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 0 & 6 & 3 & -3 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} \ge \begin{bmatrix} 0 \\ -2 \\ -5 \end{bmatrix}$, $w_1, w_2, w_3, w_4 \ge 0$

Min $z_w = -2w_1 + 6w_2 + 4\left(w_3 - w_4\right)$

$-w_1 + 2w_2 + 1\left(w_3 - w_4\right) \ge 0$

$-w_1 + w_2 + 1\left(w_3 - w_4\right) \ge -2$

$6w_2 + 3\left(w_3 - w_4\right) \ge -5$, $w_1, w_2, w_3, w_4 \ge 0$

Let $w_3^{'} = w_3 - w_4$ then $w_3^{'}$ is unrestricted.

$\Rightarrow \quad \text{Min } z_w = -2w_1 + 6w_2 + 4w_3^{'}$

s.t. $\quad -w_1 + 2w_2 + w_3^{'} \ge 0$

$-w_1 + w_2 - w_3^{'} \ge -2$

$6w_2 + 3w_3^{'} \ge -5, \; w_2, w_2 \ge 0$

$w_3^{'}$ is unrestricted.

Observation : Third constraint in primal is equation. Third variable in its dual is unrestricted in sign.

Example : Find dual of

$\text{Min } z_x = 2x_1 + 3x_2 + 4x_3$

s.t. $\quad 2x_1 + 3x_2 + 5x_3 \ge 2, \; 3x_1 + 4x_2 + 6x_3 \le 5$

$x_1, x_2 \ge 0, \; x_3$ unrestricted.

Answer : $\quad \text{Max } z_x^{'} = -2x_1 - 3x_2 - 4x_3$

s.t. $\quad -2x_1 - 3x_2 - 5x_3 \le -2$

$3x_1 + 4x_2 + 6x_3 \le 5, \; x_2, x_2 \ge 0$

$x_3 = x_4 - x_5, \; x_4, x_5 \ge 0$

$\text{Max } z_x^{'} = -2x_1 - 3x_2 - 4(x_4 - x_5)$

$-2x_1 - 3x_2 - 5(x_4 - x_5) \le -2$

$3x_1 + 4x_2 + 6(x_4 - x_5) \le 5$

Standard Primal : $\text{Max } z_x^{'} = \begin{bmatrix} -2 & -3 & -4 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_4 \\ x_5 \end{bmatrix}$

$$\begin{bmatrix} -2 & -3 & -5 & 5 \\ 3 & 4 & 6 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_4 \\ x_5 \end{bmatrix} \le \begin{bmatrix} -2 \\ 5 \end{bmatrix}, \; x_1, x_2, x_4, x_5 \ge 0$$

Its dual is

$$\text{Min } z_w = -2w_1 + 5w_2$$

$$\text{s.t.} \quad \begin{bmatrix} -2 & 3 \\ -3 & 4 \\ -5 & 6 \\ 5 & -6 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \ge \begin{bmatrix} -2 \\ -3 \\ -4 \\ 4 \end{bmatrix}, \; w_1, w_2 \ge 0$$

$$\text{Min } z_w = -2w_1 + 5w_2$$

$$-2w_1 + 3w_2 \ge -2, \; -3w_1 + 4w_2 \ge -3$$

$$-5w_1 + 6w_2 \ge -4 \text{ and } 5w_1 - 6w_2 \ge 4 \Rightarrow 5w_1 - 6w_2 = 4$$

Observation : 3rd variable in primal is unrestricted. 3rd constraint in its dual is an equation.

Theorem : The dual of the dual of a given primal is the primal.

Proof : Consider a primal

$$\text{Max } z_x = \bar{c}^T \bar{x}$$

$$\text{s.t.} \quad A\bar{x} \le \bar{b}, \; \bar{x} \ge \bar{0} \qquad \text{.... (I)}$$

Dual of the above primal is

$$\text{Min } z_w = \bar{b}^T \bar{w}$$

$$\text{s.t.} \quad A^T \bar{w} \ge \bar{c}, \; \bar{w} \ge \bar{0} \qquad \text{.... (II)}$$

The corresponding primal is,

$$\text{Max } -z_w = -\bar{b}^T \bar{w}$$

$$\text{s.t.} \quad -A^T \bar{w} \le -\bar{c}, \; \bar{w} \ge \bar{0} \qquad \text{... (III)}$$

Observe that (II) and (III) are same.

Consider dual of (III)

$$\text{Min } z_u = -\bar{c}^T \bar{u}$$

s.t. $\left(-A^T\right)^T \bar{u} \geq -\bar{b}$, $\bar{u} \geq \bar{0}$ ..... (IV)

Standard form of (IV) is,

$$\text{Max } (-z_u) = -(-\bar{c})^T \bar{u} = \bar{c}^T \bar{u}$$

s.t. $-A\bar{u} \geq -\bar{b}$, $\bar{u} \geq 0 \Rightarrow +A\bar{u} \leq \bar{b}$, $\bar{u} \geq 0$

Thus we have,

$$\text{Max } z_u' = \bar{c}^T \bar{u},\ A\bar{u} \leq \bar{b},\ \bar{u} \geq 0 \quad \text{..... (IV)}$$

Observe that (I) $\equiv$ (V)

Thus dual of dual is primal.

Theorem : If x is any FS to primal problem and w is any FS to the dual problem then,

$$\bar{c}^T \bar{x} \leq \bar{b}^T \bar{w}$$

i.e. $$\sum_{i=1}^{n} c_i x_i \leq \sum_{i=1}^{m} b_i w_i$$

Proof : Primal is Max $z_x = \bar{c}^T \bar{x}$ s.t. $A\bar{x} \leq \bar{b}$, $\bar{x} \geq \bar{0}$

Dual is Min $z_w = \bar{b}^T \bar{w}$ s.t. $A^T \bar{w} \geq \bar{c}$, $\bar{w} \geq 0$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{1n} \end{bmatrix}, \quad \bar{x} \geq \bar{0} \qquad A_{m\times n}\ \bar{x}_{n\times 1} = b_{m\times 1}$$

i.e $$\sum_{j=1}^{n} a_{ij} x_j \leq b_i,\ i = 1, 2, 3, ..., n \quad \text{.... (1)}$$

$$A^T\bar{w} \geq \bar{c} \Rightarrow \begin{bmatrix} a_{11} & a_{21} & a_{31} & \cdots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \cdots & a_{m2} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{m3} \\ \vdots & \vdots & \vdots & & \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix}_1 \geq \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

$$a_{1k}w_1 + a_{2k}w_2 + a_{3k}w_3 + \ldots. + a_{mk}w_m \geq c_k$$

$$\sum_{p=1}^{m} a_{pk}w_p \geq c_k \text{ , } k = 1, 2, 3, \ldots.., n \qquad \ldots\ldots (2)$$

From (1) and (2) we have,

$$\sum_{i=1}^{n} c_i x_i \leq \sum_{i=1}^{n}\left[\sum_{p=1}^{m} a_{pi}w_p\right]x_i = \sum_{p=1}^{m} w_p\left(\sum_{i=1}^{n} a_{pi}x_i\right)$$

$$\sum_{i=1}^{n} c_i x_i \leq \sum_{p=1}^{m} w_p\left(\sum_{j=1}^{n} a_{pj}x_j\right) \leq \sum_{p=1}^{m} w_p b_p \qquad \text{(by 1)}$$

Thus we have,

$$\bar{c}\cdot\bar{x} \leq \bar{b}\cdot\bar{w}$$

$$\bar{c}\cdot\bar{x} = (c_1c_2\ldots c_n)\cdot(x_1x_2\ldots x_n) = \sum_{i=1}^{n} c_i x_i = \bar{c}^T\bar{x}$$

$$\bar{c}^T\bar{x} = \bar{b}^T\bar{w}$$

Theorem : If $\hat{\bar{x}}$ is a FS to the primal and $\hat{\bar{w}}$ is a FS to its dual such that $\bar{c}\cdot\hat{\bar{x}} = \bar{b}\cdot\hat{\bar{w}}$ then $\hat{\bar{x}}$ is an optimal solution to the primal and $\hat{\bar{w}}$ is an optimal solution to the dual.

Proof : We know that if $\bar{x}$ is a FS to the primal and $\hat{\bar{w}}$ is a FS to its dual then $\bar{c}\cdot\bar{x} \leq \bar{b}\cdot\hat{\bar{w}}$.

Thus $\bar{c}\cdot\bar{x} \leq \bar{b}\cdot\hat{\bar{w}} = \bar{c}\cdot\hat{\bar{x}} \Rightarrow \bar{c}\cdot\bar{x} \leq \bar{c}\cdot\hat{\bar{x}}$

If $\bar{x}$ is a FS to the primal then, $\bar{c}\cdot\bar{x} \leq \bar{c}\cdot\hat{\bar{x}} \Rightarrow \bar{c}\cdot\hat{\bar{x}}$ is maximum.

$\Rightarrow \hat{\bar{x}}$ is an optimal solution to the primal.

Similarly if $\bar{w}$ is any FS to its dual $\bar{c}\cdot\hat{\bar{x}} \leq \bar{b}\cdot\bar{w}$.

But $\bar{c}\cdot\hat{\bar{x}} = \bar{b}\cdot\hat{\bar{w}}$

$\Rightarrow$ $\bar{b}\cdot\hat{\bar{w}} \leq \bar{b}\cdot\bar{w} \Rightarrow \bar{b}\cdot\hat{\bar{w}}$ is minimum.

$\Rightarrow$ $\hat{\bar{w}}$ is an optimum solution to the dual.

DUALITY THEOREM :

If $\bar{x}_0\left(\bar{w}_0\right)$ is an optimum solution to the primal (dual) then there exist a feasible solution $\bar{w}_0\left(\bar{x}_0\right)$ to the dual s.t. $\bar{c}^T\bar{x}_0=\bar{b}^T\bar{w}_0$.

Proof : Primal Max $z_x=\bar{c}^T\bar{x}$ s.t. $A\bar{x}\le\bar{b}$, $\bar{x}\ge\bar{0}$

Consider Max $z_x=\bar{c}^T\bar{x}$ s.t. $A\bar{x}+I\bar{x}_5=\bar{b}$

$A_{m\times n}$, $I_{m\times m}$ identity $\underset{m\times(n+m)}{\begin{bmatrix}A & I\end{bmatrix}}\begin{bmatrix}\bar{x}\\ \bar{x}_s\end{bmatrix}_{(n+m)\times 1}=\bar{b}$

$A=\begin{bmatrix}B & C\end{bmatrix}$ where $|B|\ne 0$ then $\bar{x}_B=\bar{B}^{-1}\bar{b}$.

Let $\bar{x}_0=\begin{bmatrix}\bar{x}_B\\ \bar{0}\end{bmatrix}$ be an optimum solution to the primal where $\bar{x}_B\in R^m$, $\bar{0}\in R^{n-m}$ then $\bar{x}_B=\bar{B}^{-1}\bar{b}$.

Therefore $z=\bar{c}^T\bar{x}_0=\bar{c}_B^T\bar{x}_B$ where $\bar{c}_B$ is cost vector corresponding to $\bar{x}_B$.

$$\Delta_j=\bar{c}_B^T\bar{x}_j-c_j=\bar{c}_B^T\bar{B}^{-1}\bar{a}_j-c_j,\ j = 1, 2, 3, ..., n$$

$$=\bar{c}_B^T\bar{B}^{-1}e_j-0,\ j = n + 1, ... n + m$$

Since $\bar{x}_0$ is optimal $\Delta_j\ge 0$.

$\therefore$ $\bar{c}_B^T\bar{B}^{-1}a_j-c_j\ge 0$, j = 1, 2, 3, ..., n

$\bar{c}_B^T\bar{B}^{-1}e_j\ge 0$, j = n + 1, n + 2, ... n + m

$\bar{c}_B^T\bar{B}^{-1}a_j\ge c_j$, j = 1, 2, 3, ..., n

$$\begin{bmatrix}\bar{c}_B^T\bar{B}^{-1}a_1 & \bar{c}_B^T\bar{B}^{-1}a_2 & \cdots & \bar{c}_B^T\bar{B}^{-1}a_n\end{bmatrix}\ge\begin{bmatrix}c_1 & c_2 & c_3 & \cdots & c_n\end{bmatrix}$$

$\bar{c}_B^T\bar{B}^{-1}A\ge\bar{c}^T$ and $\bar{c}_B^T\bar{B}^{-1}e_j\ge 0$, j = n + 1, ... n + m

Put $\bar{c}_B^T\bar{B}^{-1}=\bar{w}_0^T$ (say) $\bar{w}_0\in R^m$

Then $\bar{w}_0^TA\ge\bar{c}^T$ or $A^T\bar{w}_0\ge\bar{c}$.

Since $\bar{c}_B^T \bar{B}^{-1} e_j \geq 0$, $\bar{c}_B^T \bar{B}^{-1} \geq \bar{0}$ i.e. $\bar{w}_0^T \geq \bar{0}$

Thus $A^T \bar{w}_0 \geq \bar{c}$, $\bar{w}_0 \geq \bar{0}$

i.e. $\bar{w}_0$ is feasible solution to the dual.

$$\bar{b}^T \bar{w}_0 = \bar{w}_0{}^T \bar{b} = \bar{c}_B^T \bar{B}^{-1} \bar{b} = \bar{c}_B^T \bar{x}_B$$

Since $\bar{b}^T \bar{w}_0 = \bar{c}_B^T \bar{x}_B$

$\bar{w}_0$ is an optimum solution to the dual.

Similarly starting from dual problem we can reach to primal solution.

Theorem : If k^{th} constraint in the primal is an equality then the dual variable w_k is unrestricted in sign.

Proof : Primal

$$\text{Max } z_x = \bar{c}^T \bar{x}$$

s.t.
$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots.. + a_{1n}x_n \leq b_1$$
$$\vdots$$
$$a_{k1}x_1 + a_{k2}x_2 + a_{k3}x_3 + \ldots.. + a_{kn}x_n \leq b_k$$
$$-a_{k1}x_1 - a_{k2}x_2 - a_{k3}x_3 - \ldots.. - a_{kn}x_n \leq -b_k$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots.. + a_{mn}x_n \leq b_m$$
$$x_1, x_2, x_3, \ldots.., x_n \geq 0$$

Dual of above primal will be,

$$\text{Min } z_w = b_1w_1 + b_2w_2 + \ldots. + b_kw_k^{'} - b_kw_k^{''} + b_{k+1}w_{k+1} + \ldots + b_mw_m$$

s.t.
$$\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{k1} & -a_{k1} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{k2} & -a_{k2} & \cdots & a_{m2} \\ a_{13} & a_{23} & \cdots & a_{k3} & -a_{k3} & \cdots & a_{m3} \\ \vdots & \vdots & & \vdots & \vdots & & \\ a_{1n} & a_{2n} & \cdots & a_{kn} & -a_{kn} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k^{'} \\ w_k^{''} \\ \vdots \\ w_m \end{bmatrix} \geq \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{bmatrix}$$

$$w_1, w_2, w_3, \ldots, w_k', w_k'', \ldots, w_m \geq 0$$

$$\text{Min } z_w = b_1w_1 + b_2w_2 + \ldots + b_k\left(w_k' - w_k''\right) + \ldots + b_mw_m$$

s.t. $$a_{11}w_1 + a_{21}w_2 + \ldots + a_{k1}\left(w_k' - w_k''\right) + \ldots + a_{m1}w_m \geq b_1$$

$$a_{12}w_1 + a_{22}w_2 + \ldots + a_{k2}\left(w_k' - w_k''\right) + \ldots + a_{m2}w_m \geq b_2$$

$$\vdots$$

$$a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{kn}\left(w_k' - w_k''\right) + \ldots + a_{mn}w_m \geq b_n$$

$$w_1, w_2, \ldots, w_k', w_k'', \ldots, w_m \geq 0$$

Put $w_k = w_k' - w_k''$ then w_k is unrestricted.

Thus we have,

$$\text{Min } z_w = \sum_{i=1}^{m} b_iw_i$$

s.t. $$a_{11}w_1 + a_{21}w_2 + \ldots + a_{k1}w_k + \ldots + a_{m1}w_m \geq c_1$$

$$a_{12}w_1 + a_{22}w_2 + \ldots + a_{k2}w_k + \ldots + a_{m2}w_m \geq c_2$$

$$\vdots$$

$$a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{kn}w_k + \ldots + a_{mn}w_m \geq c_n$$

$w_1, w_2, \ldots, w_{k-1}, w_{k+1}, \ldots, w_m \geq 0$, w_k unrestricted k^{th} variable in dual is unrestricted in sign.

Theorem : If p^{th} variable in primal is unrestricted in sign then p^{th} constraint of the dual is an equation.

Proof : $$\text{Max } z_x = c_1x_1 + c_2x_2 + \ldots + c_px_p + \ldots + c_nx_n$$

s.t. $$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1p}x_p + \ldots + a_{1n}x_n \leq b_1$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2p}x_p + \ldots + a_{2n}x_n \leq b_2$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots + a_{mp}x_p + \ldots + a_{mn}x_n \leq b_n$$

$x_1, x_2, \ldots, x_{p-1}, x_{p+1}, \ldots, x_n \geq 0$, x_p unrestricted.

Since x_p is unrestricted write

$$x_p = x_p' - x_p'' \quad \text{s.t.} \quad x_p' \geq 0,\ x_p'' \geq 0$$

Then primal becomes,

$$\text{Max } z_x = c_1x_1 + ... + c_p\left(x_p' - x_p''\right) + ... + c_nx_n$$

s.t. $$a_{11}x_1 + a_{12}x_2 + ... + a_{1p}\left(x_p' - x_p''\right) + ... + a_{1n}x_n \leq b_1$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + ... + a_{mp}\left(x_p' - x_p''\right) + ... + a_{mn}x_n \leq b_m$$

$$x_1, x_2, ..., x_{p-1}, x_p', x_p'', ..., x_n \geq 0$$

The dual problem is,

$$\text{Max } z_w = b_1w_1 + b_2w_2 + ... + b_mw_m$$

s.t. $$\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1p} & a_{2p} & & a_{mp} \\ -a_{1p} & -a_{2p} & & -a_{mp} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & & a_{mn} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_m \end{bmatrix} \geq \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \\ -c_p \\ \vdots \\ c_n \end{bmatrix}$$

i.e. $$a_{11}w_1 + a_{21}w_2 + a_{31}w_3 + ... + a_{m1}w_m \geq c_1$$

$$a_{12}w_1 + a_{22}w_2 + a_{32}w_3 + ... + a_{m2}w_m \geq c_2$$

$$\vdots$$

$$a_{1p}w_1 + a_{2p}w_2 + a_{3p}w_3 + ... + a_{mp}w_m \geq c_p$$

$$-a_{1p}w_1 - a_{2p}w_2 - a_{3p}w_3 - ..... - a_{mp}w_m \geq -c_p$$

$$\vdots$$

$$a_{1n}w_1 + a_{2n}w_2 + a_{3n}w_3 + ... + a_{mn}w_m \geq c_n$$

p and $(p + 1)^{th}$ constraint implies.

$$a_{1p}w_1 + a_{2p}w_2 + a_{3p}w_3 + ... + a_{mp}w_m = c_p$$

Thus p^{th} constraint in the dual is an equation.

INTERGER LINEAR PROGRAMMING

INTRODUCTION

There are certain decision problems where decision variables make sense only if they have integer values in the solution. For example, it does not make sense saying 1.5 men working on a project or 1.6 machines in a workshop. The integer solution to the problem can, however, be obtained by rounding off the optimum value of the variables to the nearest integer value. This approach can be easy in terms of economy of effort in time and cost that might be required to derive an integer solution but this solution may not satisfy all the given constraints. Secondly, the value of the objective function so obtained may not be optimal value. All such difficulties can be avoided if the given problem, where an integer solution is required, is solved by integer programming techniques.

Types of Interger Programming Problems

There are two types of integer programming problems.

i) Linear integer programming problems.

ii) Non - linear integer programming problems.

In this unit we are going to learn the methods of solving linear integer programming problems. linear integer programming problems can be classified into three categories :

i) Pure (all) integer programming problems in which all decision variables are required to have integer values.

ii) Mixed integer programming problems in which some, but not all, of the decision variables are required to have integer values.

iii) Zero - one integer programming problems in which all decision variables must have integer values of 0 or 1.

The pure integer programming problem in its standard form can be stated as follows :

Maximize $z = c_1 x_1 + c_2 x_2 + c_3 x_3 + ... + c_n x_n$

Subject to the constraints

$a_{11} x_1 + a_{12} x_2 + a_{13} x_3 + ... + a_{1n} x_n = b_1$

$a_{21} x_1 + a_{22} x_2 + a_{23} x_3 + ... + a_{2n} x_n = b_2$

....

$$a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + ... + a_{mn}x_n = b_m$$

and $x_1, x_2, x_3, ..., x_n \geq 0$ and are integers.

Here we shall discuss two methods.

i) Gomory's cutting plane method and

ii) Branch and Bound method for solving integer programming problems.

GOMORY'S ALL INTEGER CUTTING PLANE METHOD

Gomory's cutting plane method was developed by R. E. Gomory in 1956 to solve integer linear programming problems using the dual simplex method. It is based on the generation of a sequence of linear inequalities called a 'cut'. This 'cut' cuts out a part of the feasible region of the corresponding L. P. problem while leaving out the feasible region of the integer linear programming problem. The hyperplane boundary of a cut is called the cutting plane.

Gomory's algorithm has the following properties :

i) Additional linear constraints never cut - off that portion of the original feasible solution space which contain a feasible integer solution to the original problem.

ii) Each new additional constraint (or hyperplane) cuts - off the current non - integer optimal solution to the linear programming problem.

Method for constructing additional constraint (cut)

Gomory's method begins by solving the linear programming (LP) problem without taking into consideration the integer value requirement of the decision variables. If the solution so obtained in an integer i. e. all variables in the x_B column (also called basis) of the simplex table assume non - negative integer values, the current solution is the optimal solution to the given integer LP problem. But if some of the basic variables do not have non - negative integer value, an additional linear constraint called the Gomory constraint (or cut) is generated. This linear constraint (or cutting plane), is added to the bottom of the optimal simplex table so that the solution no longer remains feasible. The new problem is then solved by using the dual simplex method. If the optimed solution so obtained in again non - integer, another cutting plane is generated. The procedure is repeated until all basis variables assume non - negative integer values.

The procedure for developing a cut

Select one of the rows, called source row for which basic variable is non - integer. The desired cut is developed by considering only fractional parts of the coefficients in source row.

Suppose the basic variable x_r has the largest fractional value among all basic variables. Then the r^{th} constraint equation (row) from the simplex table can be rewritten as ,

$$x_{B_r} = b_r = 1.x_r + \left(a_{r1}x_1 + a_{r2}x_2 + ...\right)$$

$$= x_r + \sum_{j \neq r} a_{rj}x_j \qquad .... (i)$$

Where $x_j = (j = 1, 2, 3, \dots)$ represents all the non - basic variables in the r^{th} constraint except the variables x_r and $b_r = (x_{B_r})$ is the non - integer value of varialbe x_r. Let us decompose the coefficients of x_j and x_{B_r} into integer and non - negative fractional parts in equation (i).

$$[x_{B_r}] + f_r = (1+0)\,x_r + \sum_{j \neq r} \{[a_{r_j}] + f_{r_j}\} x_j \qquad \text{.......... (ii)}$$

Where $[x_{Br}]$ and $[a_{ij}]$ denote the largest integer obtained by trucating the fractional part from x_{B_r} and a_{rj} respectively. Rearranging equation (ii) we get,

$$f_r + \{[x_{Br}] - x_r - \sum [a_{rj}] x_j\} = \sum_{j \neq r} f_{rj}\, x_j \qquad \text{.......... (iii)}$$

Where f_r is strictly positive fraction $(0 < f_r < 1)$ while $0 \leq f_{rj} \leq 1$. We may write equation (iii) in the form of following inequality.

$$f_r \leq \sum_{j \neq r} f_{rj}\, x_j$$

i. e. $\sum f_{rj}\, x_j = f_r + s_g$ or $-f_r = s_g - \sum_{j \neq r} f_{rj}\, x_j$ .... (iv)

Where S_g is a non - negative slack variable and is called the Gomory slack variable. Equation (iv) represents Gomory's cutting plane constraint. This constraint create an additional row along with a column for the new variable S_g.

Steps of Gormory's all integer programming algorithm

Step - 1

Initialization : Formulate the standard integer LP problem. If there are any non - integer coefficients in the constraint equations, convert them into integer coefficients. Solve it by simplex method, ignoring the integer requirement of variables.

Step - 2

Test of optimality

a) Examine the optimal solution. If all basic variables (i. e. $x_{Bi} = b_i \geq 0$) have integer values, the integer optimal solution has been derived and the procedure should be terminated. The current optimal solution obtained in step 1 is the optimal basic feasible solution to the integer linear programming.

b) If one or more basic variables with integer requirements have non - integer solution values, then go to step 3.

Step - 3

Generate cutting plane : Choose a row r corresponding to a variable x_r which has the largest fractional value f_r and generate the cutting plane (a Gomory constraint) as explained earlier in equation (iv)

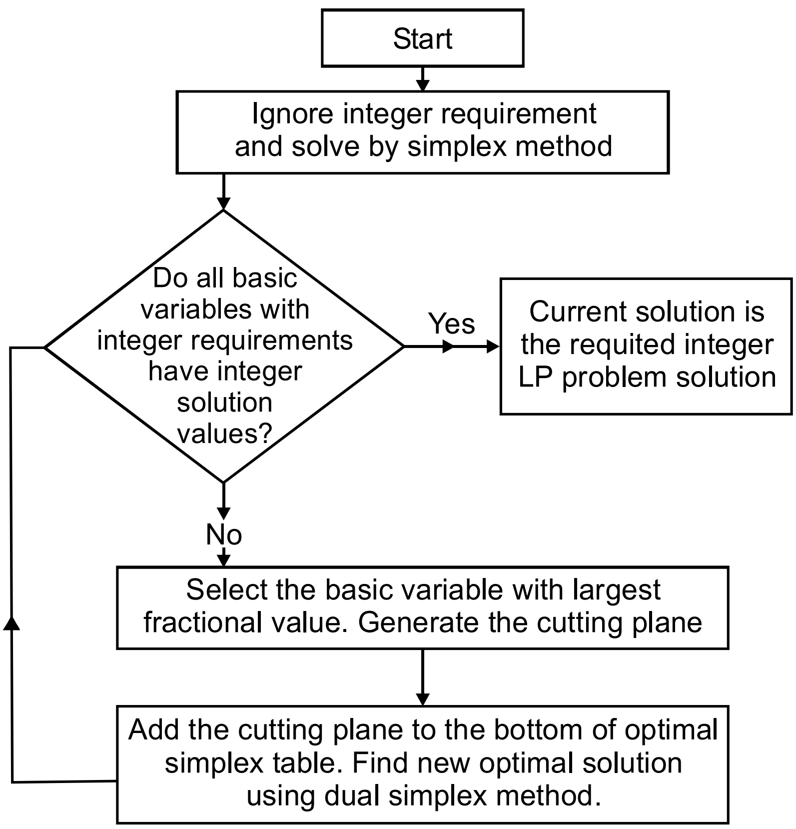
$$-f_r = s_g - \sum_{j \neq r} f_{rj} x_j$$

where $0 \leq f_{rj} < 1$ and $0 < f_r < 1$.

If there are more than one variables with the same largest fraction, then choose the one that has the smallest contribution to the maximization LP problem or the largest cost to the minimization LP problem.

Step - 4

Obtain the new solution : Add the cutting plane generated in step 3 to the bottom of the optimal simplex table as obtained in step. 3. Find a new optimal solution by using the dual simplex method i. e. choose a variable to enter into the new solution having the smallest ratio $\{\Delta_j / a_{ij}; a_{ij} < 0\}$ and return to step 2.



The process is repeated until all basic variables with integer requirements assume non - negative integer values.

The procedure for solving an ILP problem can be explained through a flow chart given above.

EXAMPLES

1) Solve the following integer programming problem using Gomory's cutting plane algorithm.

Maximize $z = x_1 + x_2$

Subject to

$3x_1 + 2x_2 \leq 5$

$x_2 \leq 2$

and $x_1, x_2 \geq 0$ and are integers.

Answer :

Step : 1

Introducing the slack variables we get,

Maximize $z = x_1 + x_2 + 0s_1 + 0s_2$

Subject to

$3x_1 + 2x_2 + s_1 = 5$

$x_2 + s_2 = 2$

and $x_1, x_2, s_1, s_2 \geq 0$

The optimum solution to the LPP is given below.

		C_j	1	1	0	0	
Basic Variables	**Coeffts of Basic variables** C_B	**Values of Basic variables** $b = X_B$	**Variables** x_1	x_2	s_1	s_2	**Min Ratio** x_B / x_k
s_1	0	5	3	2	1	0	5 / 2
$\leftarrow s_2$	0	2	0	[1]	0	1	[2/1]
	$z = c_B x_B = 0$	$\Delta_j = z_j - c_j = c_B x_j - c_j \rightarrow$	-1	-1 $\uparrow$	0	0	
$\leftarrow s_1$	0	1	[3]	0	1	-2	[1/3]

$\rightarrow x_2$	1	2	0	1	0	1	2/0
	$z = c_B x_B = 2$	$\Delta_j = z_j - c_j \rightarrow$	-1 $\uparrow$	0	0	-1	
$\rightarrow x_1$ x_2	1 1	1 / 3 2	1 0	0 1	1/3 0	-2/3 1	
	$z = 7/3$	$\Delta_j = z_j - c_j \rightarrow$	0	0	1/3	1/3	$\Delta_j \geq 0$

The optimal solution is $x_1 = \frac{1}{3}, x_2 = 2$ and Max. $z = \frac{7}{3}$.

Step : 2

In the current optimal solution, all the basic variables in the basis are not integers and the solution is not acceptable. Since both decision variables x_1 and x_2 are assumed to take an integer value, a pure integer cut is developed under the assumption that all the variables are integers. We go to next step.

Step : 3

Since x_1 is the only basic variable whose value is a non - negative fraction, we shall consider the fist row for generating the Gomory cut. Considering x_1 - equation as the source row we write.

$$\frac{1}{3} = x_1 + 0.x_2 + \frac{1}{3}s_1 - \frac{2}{3}s_2 \qquad (x_1 \text{ - source row})$$

The factoring of the x_1 - source row yields

$$\left(0 + \frac{1}{3}\right) = (1+0)\,x_1 + \left(0 + \frac{1}{3}\right)s_1 + \left(-1 + \frac{1}{3}\right)s_2$$

Observe that each of the non - integer coefficient is factored into integer and fractional parts in such a manner that the fractional part the fractional part is strictly positive.

Rearrange the equation so that all of the integer coefficients appear on the left hand side. This gives

$$\frac{1}{3} + (s_2 - x_1) = \frac{1}{3}s_1 + \frac{1}{3}s_2$$

Therefore $\frac{1}{3} \leq \frac{1}{3}s_1 + \frac{1}{3}s_2$

Thus complete Gomorian constraint can be written as

$$\frac{1}{3}+g_1=\frac{1}{3}s_1+\frac{1}{3}s_2 \text{ or } -\frac{1}{3}=g_1-\frac{1}{3}s_1-\frac{1}{3}s_2$$

Where g_1 is the new non - negative (integer) slack variable.

By adding the Gomory cut at the bottom of the simplex table, the new table so obtained is given below.

		$c_j \rightarrow$	1	1	0	0	0	
Basic Variables	**Coeffts of Basic variables**	**Values of Basic variables**	**Variables** x_1	x_2	s_1	s_2	g_1	
x_1	1	1 / 3	1	0	1/3	-2/3	0	
x_2	1	2	0	1	0	1	0	
g_1	0	- 1 / 3	0	0	-1/3	-1/3	1	

Step - 4

Apply the dual simplex method to find the new optimal solution.

		$c_j \rightarrow$	1	1	0	0	0	
Basic Variables	**Coeffts of Basic variables**	**Values of Basic variables**	**Variables** x_1	x_2	s_1	s_2	g_1	
x_1	1	1 / 3	1	0	1/3	-2/3	0	
x_2	1	2	0	1	0	1	0	
$\leftarrow g_1$	0	- 1 / 3	0	0	-1/3	-1/3	1	
$z=\frac{7}{2}$	$z_j - c_j =$		0	0	1/3 $\uparrow$	1/3	0	
x_1	1	0	1	0	0	-1	1	
x_2	1	2	0	1	0	1	0	
s_1	0	1	0	0	1	1	-3	
$z=2$	$\Delta = z_j - c_j \rightarrow$		0	0	0	0	1	

Since all $\Delta_j \geq 0$, the solution is optimal solution. Thus $x_1=0, x_2=2, s_1=1$ and max. z = 2. This solution satisfies the integer requirement.

2) Solve the following integer programming problem using Gomory's cutting plane algorithm.

Maximize $z = 2x_1 + 20x_2 - 10x_3$

Subject to $2x_1 + 20x_2 + 4x_3 \le 15$

$6x_1 + 20x_2 + 4x_3 = 20$

and x_1, x_2, x_3 are non - negative integers.

Also show that it is not possible to obtain a feasible integer solution by using the method of simplex rounding off.

Answer :

Adding slack variable s_1 in the first constraint and artificial variable in the second constraint the problem is stated in the standard form as :

Maximize $z = 2x_1 + 20x_2 - 10x_3 + 0s_1 - MA_1$

subject to

$2x_1 + 20x_2 + 4x_3 + s_1 = 15$

$6x_1 + 20x_2 + 4x_3 + A_1 = 20$

and $x_1, x_2, s_1, A_1 \ge 0$ and are integers.

The optimal solution of the problem ignoring the integer requirement using the simplex method (Big M technique) is obtained in the following table.

Basic Variables	Coeffts of Basic variables	Values of Basic variables	x_1	x_2	x_3	s_1	A_1	Min Ration
		c_j	2	20	-10	0	-M	
$\leftarrow s_1$	0	15	2	[20]	4	1	0	[15/20]
A_1	- M	20	6	20	4	0	1	20 / 20
$Z = -20M$	$z_j - c_j \rightarrow$		-6M-2	-20 M-20	-4M+10	0	0	
x_2	20	3/4	1/10	1	1/5	1/20	0	15/2
$\leftarrow A_1$	- M	5	4	0	0	-1	1	[5/4]
$z = 15 - 5M$	$z_j - c_j \rightarrow$		- 4 M $\uparrow$	0	14	M + 1	0	

x_2	20	5 / 8	0	1	1/5	3/40	-1/40	
x_1	2	5 / 4	1	0	0	-1/4	1/4	
$z=15$	$z_j - c_j \rightarrow$		0	0	14	1	M	$\Delta_j \geq 0$

The non - integer optimal solution is $x_1 = 5/4, x_2 = 5/8, x_3 = 0$ and Max. z = 15. Then the rounded off solution will be $x_1 = 1, x_2 = 0, x_3 = 0$ and Max z = 2. This solution does not satisfy the second constraint $6x_1 + 20x_2 + 4x_3 = 20$. Hence it is not possible to obtain an integer optimal solution by simply rounding off the values of the variables.

To obtain the integer valued solution, we proceed to construct Gomory's constraint (fractional cut). Since the fractional part of the value of $x_2 = (0 + 5/8)$ is more that the fractional part of $x_1 = (1 + 1/4)$, the x_2 - row is selected for constructing the fractional cut as given below.

$$\frac{5}{8} = 0.x_1 + 1.x_2 + \frac{1}{5}x_3 + \frac{3}{40}s_1$$

$$\left(0 + \frac{5}{8}\right) = (1+0)x_2 + \left(0 + \frac{1}{5}\right)x_3 + \left(0 + \frac{3}{40}\right)s_1$$

On rearranging above equation we obtain the Gomory's fractional cut as,

$$-\frac{5}{8} = g_1 - \frac{1}{5}x_3 - \frac{3}{40}s_1 \qquad \text{(Cut I)}$$

Adding this additional constraint at the bottom of optimal simplex table, we get

		c_j	2	20	-10	0	0
Basic Variables	**Coeffts of Basic variables**	**Values of Basic variables**	**Variables**				
			x_1	x_2	x_3	s_1	g_1
x_2	20	5 / 8	0	1	1 / 5	3 / 40	0
x_1	2	5 / 4	1	0	0	- 1/4	0
$\leftarrow g_1$	0	- 5 / 8	0	0	-1/5	-3/40	1
z = 15	$z_j - c_j \rightarrow$		0	0	14	1 $\uparrow$	0

Here $$\max\left\{\frac{0}{0}, \frac{0}{0}, \frac{14}{(-1/5)}, \frac{1}{(-3/40)}\right\}$$

$$= \max\left\{-,-,-70,-\frac{40}{3}\right\}$$

$$= -\frac{40}{3} \text{ Therefore we must enter the variable } s_1.$$

Thus s_1 is the entering variable whereas g_1 is outgoing variable. Here we are applying dual simplex method.

		c_j	2	20	-10	0	0
Basic Variables	**Coeffts of Basic variables**	**Values of Basic variables**	**Variables**				
			x_1	x_2	x_3	s_1	g_1
x_2	20	0	0	1	0	0	1
x_1	2	10 / 3	1	0	2 / 3	0	-10/3
s_1	0	25 / 3	0	0	8/3	1	-40/3
z = 20 / 3	$z_j - c_j \rightarrow$		0	0	34/3	0	40/3

The solution is optimal but is still non - integer solution. Therefore one more fractioned but should be added. Consider x_1 - row for censtructing the cut.

$$\left(3+\frac{1}{3}\right)=(1+0)x_1+\left(0+\frac{2}{3}\right)x_3+\left(-4+\frac{2}{3}\right)g_1$$

We obtain Gomory's fractional cut as,

$$-\frac{1}{3}=g_2-\frac{2}{3}x_3-\frac{2}{3}g_1 \qquad \text{(Cut - II)}$$

Adding this constraint to the optiomal simplex table the new table becomes

		c_j	2	20	-10	0	0	0
Basic Variables	**Coeffts of Basic variables**	**Values of Basic variables**	**Variables**					
			x_1	x_2	x_3	s_1	g_1	g_2
x_2	20	0	0	1	0	0	1	0
x_1	2	$\frac{10}{3}$	1	0	$\frac{2}{3}$	0	$-\frac{10}{3}$	0
s_1	0	$\frac{25}{3}$	0	0	$\frac{8}{3}$	1	$-\frac{40}{3}$	0

$\leftarrow g_2$	0	$-\frac{1}{3}$	0	0	$-\frac{2}{3}$	0	$-\frac{2}{3}$	1
$z=\frac{20}{3}$	$z_j - c_j$		0	0	$\frac{34}{3}$	0	$\frac{40}{3}$	0
		Ratio	-	-	$\frac{34/3}{-2/3}$ = - 17 ↑		$\frac{40/3}{-2/3}$ - 20	-

Maximum ratio = - 17. Remove g_2 from the basis and enter variable x_3 into the basis by applying the dual simplex method.

		c_j	2	20	-10	0	0	0
Basic Variables	**Coeffts of Basic variables**	**Values of Basic variables**	**Variables**					
			x_1	x_2	x_3	s_1	g_1	g_2
x_2	20	0	0	1	0	0	1	0
x_1	2	3	1	0	0	0	-4	0
s_1	0	7	0	0	0	1	-16	4
x_3	-10	1/2	0	0	1	0	1	-3/2
	z = 1							

The above optimal solution is still non - integer because variable x_3 doex not have integer value. Thus a first fractional cut will have to be constructed with the help of x_3 - row and the required Gomory's fractional cut is

$$-\frac{1}{2}=g_3-\frac{1}{2}g_2 \qquad \text{(Cut III)}$$

Additing this cut to the bottom of above table we get a new table. Apply the dual simplex method.

		c_j	2	20	-10	0	0	0	0
Basic Variables variables	**Coeffts of Basic variables**	**Values of Basic**	**Variables**						
			x_1	x_2	x_3	s_1	g_1	g_2	g_3
x_2	20	0	0	1	0	0	1	0	0
x_1	2	3	1	0	0	0	-4	0	0
s_1	0	7	0	0	0	1	-16	4	0
x_3	-10	1/2	0	0	1	0	1	-3/2	0
$\leftarrow g_3$	0	-1/2	0	0	0	0	0	[-1/2]	1
z = 1	$z_j - c_j \rightarrow$		0	0	0	0	2	15	0
	Ratio $\frac{z_j - c_j}{5^{th}\text{ row}} \rightarrow$		-	-	-	-	-	-30	-
								↑	

Max. ratio = - 30 and therefore remove variable g_3 and enter variable g_2 into the basis By applying the dual simplex method, we get the new optimal solution as shown in the following table.

		c_j	20	20	-10	0	0	0	0
Basic Variables variables	**Coeffts of Basic variables**	**Values of Basic**	**Variables**						
			x_1	x_2	x_3	s_1	g_1	g_2	g_3
x_2	20	0	0	1	0	0	1	0	0
x_1	2	3	1	0	0	0	-4	0	0
s_1	0	3	0	0	0	1	-16	0	8
x_3	-10	2	0	0	1	0	1	0	-3
g_2	0	1	0	0	0	0	0	1	-2
$z = -14$		$z_j - c_j \rightarrow 0$	0	0	0	2	0	30	

Since all the variables in above table have assumed integer values and all $z_j - c_j \geq 0$, the solution is integer optimal solution. $x_1 = 3, x_2 = 0, x_3 = 2$ and maz x = - 14.

3) The owner of a readymade garments store sells two types of shirts - zee shirts and button - down shorts. He makes a profit of Rs. 3 and Rs. 12 per shirt on zee - shirts and Button down shirts, respectively. He has tow tailors A and B at his disposal to stitch the shirts. Tailors A and B can devote at the most 7 hours and 15 hours per day respectively. Both these shirts are to be stitched by both the tailors. Tailors A and B spend 2 hours and 5 hours, respectively in stitching one zee - shirt and 4 hours and 3 hours, respectively in stitching a Button down shirt. How many shirts of both types should be stitched in order to maximize daily profit?

a) Formulate and solve this problem as an LP problem.

b) If the optimal solution is not integer valued, use Gomory technique to derive the optimal integer solution.

Answer :

Let x_1 and x_2 are number of zee - shirts and Button down shirts to be stitched daily, respectively. Then we have to maximize profit = $3x_1 + 12x_2$ subject to the constraints.

i) Availability of time with tailor A

$$2x_1 + 4x_2 \leq 7$$

ii) Availability of time with tailor B

$$5x_1 + 3x_2 \leq 15$$

and $x_1, x_2 \geq 0$ and are integers. Thus we get,

Maximize $z = 3x_1 + 12x_2$

Subject to,

$$2x_1 + 4x_2 \leq 7$$

$$5x_1 + 3x_2 \leq 15$$

and $x_1, x_2 \geq 0$ and are integers.

Adding slack variables s_1 and s_2 the given LP problem is stated into its standard form.

Maximize $z = 3x_1 + 12x_2$

Subject to,

$$2x_1 + 4x_2 + s_1 = 07$$

$$5x_1 + 3x_2 + s_2 = 15$$

and $x_1, x_2, s_1, s_2 \geq 0$

		c_j	3	12	0	0	
Basic Variables	**Coeffts of Basic variables** C_B	**Values of Basic variables** $b = X_B$	**Variables** x_1	x_2	s_1	s_2	**Min Ratio** x_B / x_k
$\leftarrow s_1$	0	7	2	4	1	0	7/4
s_2	0	15	5	3	0	1	15 / 3
$z = 0$		$z_j - c_j \rightarrow$	-3	-12	0	0	
$\rightarrow x_2$	12	7 / 4	1 / 2	1	1 / 4	0	
s_2	0	39 / 4	7 / 2	0	- 3 / 4	1	
z = 21		$z_j - c_j \rightarrow$	3	0	3	0	$\Delta_j \geq 0$

The non - integer optimal solution is $x_1 = 0, x_2 = 7/4$ and max z = 21.

b)

To construct Gomory's fractional cut we use x_2 - rows.

$$\frac{7}{4} = \frac{1}{2}x_1 + x_2 + \frac{1}{4}s_1$$

The required fractional cut is

$$-\frac{3}{4} = g_1 - \frac{1}{2}x_1 - \frac{1}{4}s_1$$

Adding this additional constraint to the bottom of the optimal simplex and applying the dual simplex method we get the following iterations.

		c_j	3	12	0	0	0
Basic Variables	**Coeffts of Basic variables**	**Values of Basic variables**	**Variables** x_1	x_2	s_1	s_2	g_1
x_2	12	7 / 4	1 / 2	1	1 / 4	0	0
s_2	0	39 / 4	7 / 2	0	- 3 /4	1	0
$\leftarrow g_1$	0	- 3 / 4	- 1 /2	0	-1/4	0	1
	z = 21	$z_j - c_j$	3	0	3	0	0

		$\frac{z_j - c_j}{\text{row } 3}$	- 6	-	- 12	0	0
			↑				
x_2	12	1	0	1	0	0	1
s_2	0	g / 2	0	0	- 5/2	1	7
x_1	3	3 / 2	1	0	1/2	0	- 2
	$z = \frac{33}{2}$	$z_j - c_j \rightarrow$	0	0	$\frac{3}{2}$	0	6

The optimal solution is still non - integer. Therefore adding one more fractional out with the help of x_1 - row we get the following able and subsequent interations by dual simplex method.

		c_j	3	12	0	0	0	0
Basic Variables	**Coeffts of Basic variables**	**Values of Basic variables**	**Variables**					
			x_1	x_2	s_1	s_2	g_1	g_2
x_2	12	1	0	1	0	0	1	0
s_2	0	9 / 2	0	0	$-\frac{5}{2}$	1	7	0
x_1	3	3 / 2	1	0	1/2	0	-2	0
g_2	0	- 1 /2	0	0	[-1/4]	0	0	1
$z = \frac{33}{2}$		$z_j - c_j \rightarrow$	0	0	$\frac{3}{2}$	0	6	0
	Ratio $\frac{z_j - c_j}{\text{row } 4} \rightarrow$		-	-	-3	0	-	-
x_2	12	1	0	1	0	0	1	0
s_2	0	7	0	0	0	1	7	-5
x_1	3	1	1	0	0	0	-2	1
s_1	0	1	0	0	1	0	0	-2
z = 15		$z_j - c_j \rightarrow$	0	0	0	0	6	$3 \geq 0$

Since all the variables have assumed integer values and all $z_j - c_j \geq 0$, the solution is an integer optimal solution. Thus the company should produce $x_1 = 1$ zee shirt, $x_2 = 1$. Button - down shirt to yield maximum profit z = Rs. 15.

4.4 GEOMETRICAL INTERPRETATION OF GOMORY'S CUTTINGS PLANE METHOD

Let us consider the problem

Maximum $z = x_1 + x_2$

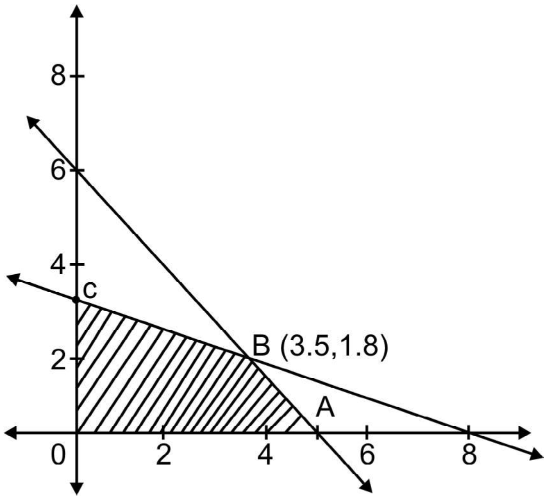
Subject to

$$2x_1 + 5x_2 \le 16$$

$$6x_1 + 5x_2 \le 30$$

$$x_1, x_2 \ge 0$$

The graphical solution of this problem is obtained in the figure with solution space represented by the convex region OABC. The optimal solution occurs at the extreme point B i. e. $x_1 = 3.5, x_2 = 1.8$, max z = 5.3. But this solution is not integer valued. While solving this problem by Gomory's method, we introduce first Gomory's constraint $-\frac{3}{10}x_3 - \frac{9}{10}x_4 \le -\frac{4}{5}$.



In order to express this constraint in terms of $x_1 \& x_2$, we use the constraints $2x_1 + 5x_2 + x_3 = 16$ and $6x_1 + 5x_2 + x_4 = 30$. Then Gomory's constraint becomes,

$$-\frac{3}{10}(16 - 2x_1 - 5x_2) - \frac{9}{10}(30 - 6x_1 - 5x_2) \le -\frac{4}{5}$$

i. e. $x_1 + x_2 \le 5\frac{1}{6}$

This constraint cuts off the feasible region and now the feasible region is reduced to somewhat less than the previous one and the procedure continues till an integer valued corner is found. Because of cuttings in the feasible region, the method was named as cutting plane method.

~ ~ ~ ~ ~ EXERCISE ~ ~ ~ ~ ~

Find the optimum integer solution of the following all integer programming problems.

1) Max $z = x_1 + x_2$

Subjct to

$$3x_1 - 2x_2 \le 5$$

$$x_1 \le 2$$

$x_1, x_2 \geq 0$ and are integers. $(\text{Ans.}: x_1 = 3, x_2 = 2, \text{max. } z = 5)$

2) Max. $z = x_1 - 2x_2$

Subjct to

$4x_1 + 2x_2 \leq 15$

$x_1, x_2 \geq 0$ and integers.

$(\text{Ans.}: x_1 = 3, x_2 = 0, \text{max. } z = 3)$

3) Max. $z = 3x_2$

Subject to,

$3x_1 + 2x_2 \leq 7$

$x_1 - x_2 \geq -2$

$x_1, x_2 \geq 0$ and integers.

$(\text{Ans.}: x_1 = 0, x_2 = 2, \text{max } z = 6)$

4) Max. $z = x_1 + 5x_2$

Subject to,

$x_1 + 10x_2 \leq 20$

$x_1 \leq 2$

$x_1, x_2 \geq 0$ and integers.

$(\text{Ans.}: x_1 = 2, x_2 = 1, \text{max } z = 7)$

5) Max. $z = 3x_1 + 4x_2$

Subject to,

$3x_1 + 2x_2 \leq 8$

$x_1 + 4x_2 \geq 10$

$x_1, x_2 \geq 0$ and are integers.

$(\text{Ans.}: x_1 = 0, x_2 = 4, \text{max } z = 16)$

6) Max. $z = 11x_1 + 4x_2$

Subject to,

$-x_1 + 2x_2 \leq 4$

$5x_1 + 2x_2 \le 16$

$2x_1 - x_2 \le 4$

$x_1, x_2 \ge 0$ and are integers.

$(Ans.: x_1 = 2, x_2 = 3, \max z = 34)$

7) Max. $z = x_1 - x_2$

Subject to,

$x_1 + 2x_2 \le 4$

$6x_1 + 2x_2 \le 9$

$x_1, x_2 \ge 0$ and are integers.

$(Ans.: x_1 = 1, x_2 = 0, \max z = 2)$

8) Max. $z = 3x_1 - 2x_2 + 5x_3$

Subject to,

$5x_1 + 2x_2 + 7x_3 \le 28$

$4x_1 + 5x_2 + 5x_3 \le 30$

$x_1, x_2, x_3 \ge 0$ and are integers.

$(Ans.: x_1 = 0, x_2 = 0, x_3 = 4, \max z = 20)$

BRANCH AND BOUND METHOD

The branch and bound method was first developed by A. H. Land and A. G. Daig and it was further studied by J.O. C. Little et. al. and other researchers. This method can be used to solve all integer, mixed integer and zero - one linear problems. This is the most general technique for the solution of integer programming problem (I.P.P.) in which a few or all the variables are constrained by their upper or lower bounds.

STEPS OF BRANCH AND BOUND ALGORITHM

Step : 1

Initialization : Consider the following all integer programming problem.

$$\left.\begin{array}{l}\text{Maximize } z=c_1 x_1+c_2 x_2+\ldots+c_n x_n \\ \text{Subject to constraints} \\ a_{11}x_1+a_{12}x_2+a_{13}x_3+\ldots+x_{1n}x_n \le b_1 \\ a_{21}x_1+a_{22}x_2+a_{23}x_3+\ldots+a_{2n}x_n \le b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+a_{m3}x_3+\ldots+a_{mn}x_n \le b_m\end{array}\right\} (LP-A)$$

Obtain the optimal solution of the given problem ignoring integer restriction on the variables.

If the solution to this LP problem (say L P - A) is infeasible or unbounded, the solution to the given all integer programming problem is also infeasible or unbounded, as the case may be,

Otherwise examine optimal feasible solution. If the answer satisfies the integer restrictions, the optimal integer solution has been obtained. If one or more basic variables do not satisfy integer requirement then go to step 2.

Step : 2

a) Let the optimal value of objective function of LP - A be z_1. This value provides an initial upper bound on objective function value for integer LP problem. Let it be denoted by z_u. The lower bound on integer LP problem can be obtained by truncating to integer all values of the varialbes. Let the lower bound be denoted by z_L.

b) Let x_k be the basic variable having largest fractional value.

c) Branch (or partition) the LP - A into two new LP sub - problems (also called nodes) based on integer value of x_k i. e. partitioning is done by adding two mutually exclusive constraints.

$x_k \le [x_k]$ and $x_k \ge [x_k]+1$

to the original LP problem. Here $[x_k]$ is the integer portition of the current non - integer value of the variable x_k. This is done to exclude the non - integer value of the variable x_k. The two new LP sub problems are as follows.

LP sub - problem B	LP sub - problem C
Max $z=\sum_{j=1}^{n} c_j . x_j$	Max $z=\sum c_j . x_j$
subject to	subject to
$\sum a_{ij} x_j = b_i$	$\sum a_{ij} x_j = b_i$

$x_k \le [x_k]$ $\qquad$ $x_k \ge [x_k]+1$

and $x_j \ge 0$ $\qquad$ and $x_j \ge 0$

Step : 3

Bound step : Obtain optimal solution of sub - problems B and C. Let the optimal value of the objective function of LP - B be z_2 and that of LP - C be z_3.

Step : 4

Examine solution of both LP - B and LP - C, which might contain optimal point.

1) Exclude a sub - problem from further consideration if it has an infeasible solution.

2) If a sub - problem yields a solution that is feasible but not an integer then for this sub - problem return to step - 2.

3) If a sub - problem yields a feasible integer solution examine the value of objective function. If this value is equal to the upper bound z_U, an optimal solution has been reached. But if it is not equal to the upper bound z_U but exceeds the lower bound z_L, this value is considered as new upper bound and return to step 2. Finally if it is less than the lower bound, terminate this branch.

Step : 5

The procedure of branching and bounding contimes until no further sub problem remains to be examined. At this stage, the integer solution corresponding to the current lower bound is the optimal all integer programming problem solution.

Examples

1) Solve the following all integer programming problem using the branch and bound method.

Maximize $z = 3x_1 + 5x_2$

Subject to the constraints

$2x_1 + 4x_2 \le 25$

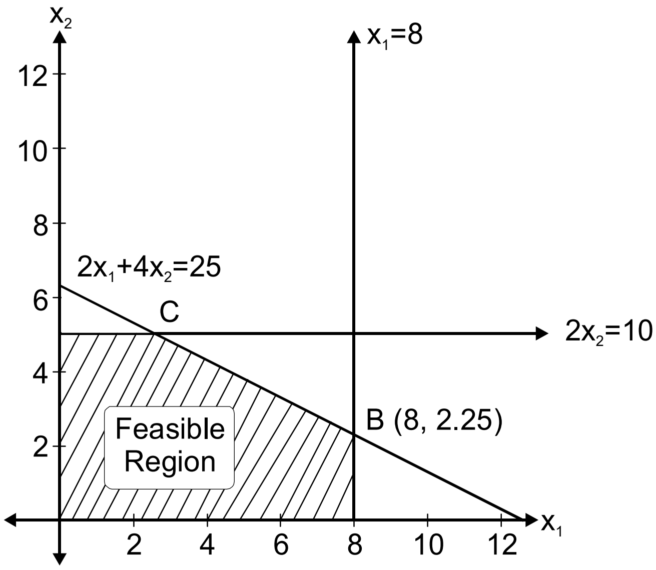
$x_1 \le 8$

$2x_2 \le 10$

and $x_1, x_2 \ge 0$ and integers.

Answer :

Relaxing the integer requirements, the optimal non - integer solution of the given integer L. P. problem obtained by the graphical method as shown below is $x_1 = 8, x_2 = 2.25$ and $z_1 = 35.25$.



The value of z_1 represents the initial upper bound, $z_u = 35.25$ on the value of the objective function i. e. the value of the objective function in the subsequent steps cannot exceed 35.25. The lower bound z_L is obtained by truncating the solution values to $x_1 = 8$ and $x_2 = 2$.

Thus $z_L = 3(8) + 5(2) = 34$

The variable $x_2 (= 2.25)$ is the only non - integer solution value and is therefore selected for dividing the given problem into two sub - problems LP - B and LP - C. Two new censtrains $x_2 \le 2$ and $x_2 \ge 3$ are created. These two constraints are added to the given problem to get two sub - problems.

LP - B

Max $z = 3x_1 + 5x_2$

Subject to,

$$2x_1 + 4x_2 \le 25$$

$$x_1 \le 8$$

$$2x_2 \le 10$$

$$x_2 \le 2$$

and $x_1, x_2 \ge 0$ and integers.

LP - C

Max. $z = 3x_1 + 5x_2$

Subject to,

$$2x_1 + 4x_2 \le 25$$

$$x_1 \le 8$$

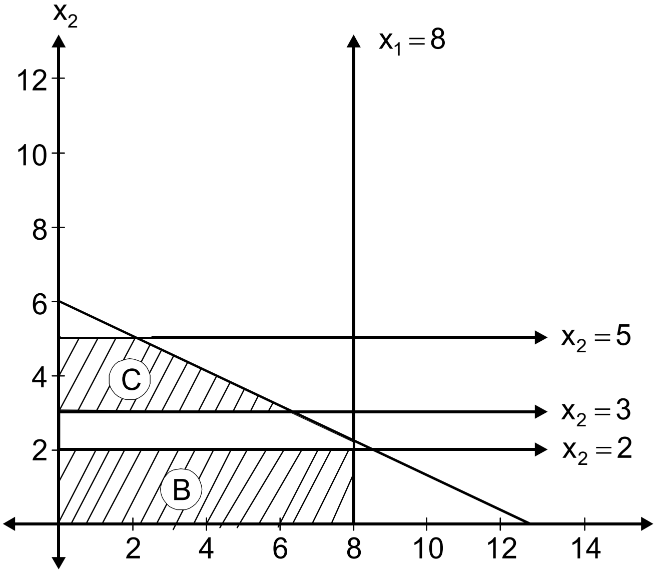
$$2x_2 \le 10$$

$$x_2 \ge 3$$

and $x_1, x_2 \ge 0$ and integer.

In sub - problem L. P. B. the constraint $2x_2 \le 10$ is redundant as $x_2 \le 2$ satisfy $2x_2 \le 10$.

Subproblem B and C are solved graphically.



B) Feasble region for sub - problem B

C) Feasible region for sub - problem C.

The solution to subproblem B is $x_1 = 8, x_2 = 2, z_2 = 34$.

The solution to subproblem C is $x_1 = 6.5, x_2 = 3, z_3 = 34.5$. Notice that both solution yield value of z lower than that of original LP problem. The value of z, establishes an upper bound on z_2 and z_3 values of sub - problems.

Since the solution of sub - problem B is an all integer, we stop the search of this sub - problem i. e. no further branching is required from node B. The value of $z_2 = 34$ becomes the new lower bound on the IP problems optiomal solution. A non - integer solution of sub - problem C and also $z_3 > z_2$, both indicate that further brancing is necessary from node C. However if $z_3 \le z_2$ then no further branching would have been required from node C. The upper bound now takes the value $z_U = z_3 = 34.5$ instead of 35.25 at node A.

The sub - problem C is now branched into two new subproblems D and E, and are obtained by adding the constraints $x_1 \le 6$ and $x_1 \ge 7$ (for problem C, $x_1 = 6.25$)

LP - D	LP - E
Max. $z = 3x_1 + 5x_2$	Max. $z = 3x_1 + 5x_2$
Subject to,	Subject to,
$2x_1 + 4x_2 \le 25$	$2x_1 + 4x_2 \le 25$
$x_1 \le 8$	$x_1 \le 8$
$2x_2 \le 10$	$2x_2 \le 10$
$x_2 \ge 3$	$x_2 \le 3$

$x_1 \le 6$ $\qquad\qquad$ $x_1 \ge 7$

and $x_1, x_2 \ge 0$ and integers. $\qquad$ and $x_1, x_2 \ge 0$ and integers.

Sub - problems D and E are solved graphically.

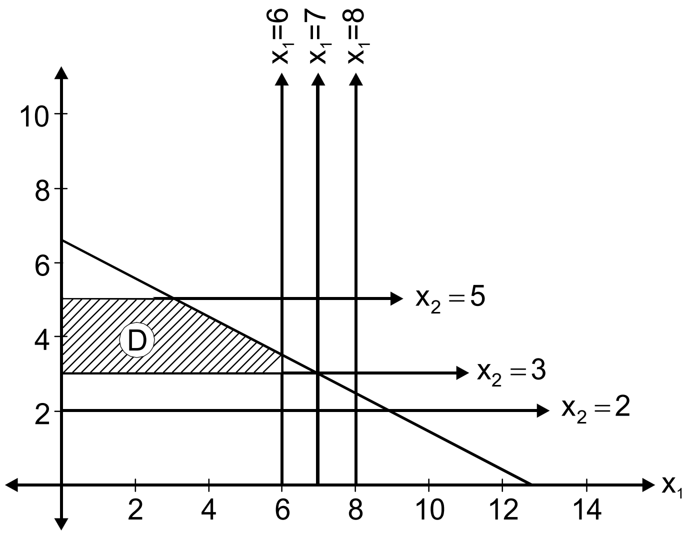
The solutions are

LP - D : $x_1 = 6, x_2 = 3.25, \text{Max. } z = z_4 = 34.25$

LP - E : No feasible solution exists because constraints

$x_1 \ge 7$ and $x_2 \ge 3$ do not satisfy $2x_1 + 4x_2 \le 25$.

So this branch is terminated.



In problem - D solution $x_2 = 3.25$ is not an integer solution. Create new sub problems F and G from sub problem D with two new constraints $x_2 \le 3$ and $x_2 \ge 4$.

LP - F	LP - G
Max. $z = 3x_1 + 5x_2$	Max. $z = 3x_1 + 5x_2$
Subject to,	Subject to,
$2x_1 + 4x_2 \le 25$	$2x_1 + 4x_2 \le 25$
$x_1 \le 8$	$x_1 \le 8$
$2x_2 \le 10$	$2x_2 \le 10$
$x_2 \ge 3$	$x_2 \ge 3$
$x_1 \le 6$	$x_1 \le 6$
$x_2 \le 3$	$x_2 \ge 4$

and $x_1, x_2 \geq 0$ and integers. and $x_1, x_2 \geq 0$ and integers.

The graphical solution of sub - problems F and G gives

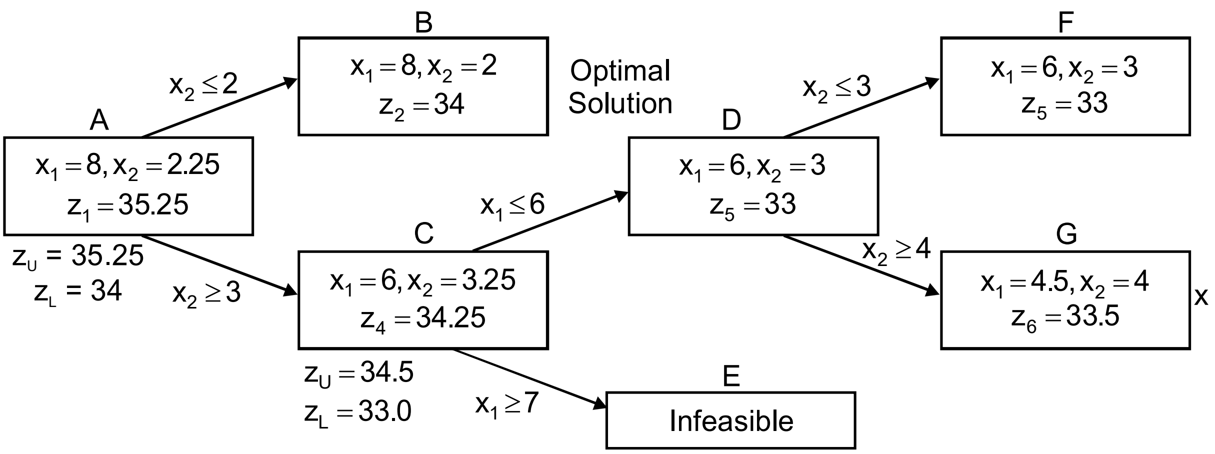
sub - problems F : $x_1 = 6, x_2 = 3$ and Max. $z = z_5 = 33$

sub - problems G : $x_1 = 4.25, x_2 = 4$ and Max. $z = z_6 = 33.5$

The branching process is terminated when new upper bound is less than or equal to the lower bounds of previous solutions or no further branching is possible.

Although the solution at node G is non - integer, no additional branching is required from this node because $z_6 < z_4$. The branch and bound algorithm is terminated and the optimal integer solution is $x_1 = 8, x_2 = 2$ and z = 34 yielded at node B.

The branch and bound procedure for the above problem is given below.



2) Use branch and bound technique and solve the following integer programming problem.

Max. $z = 7x_1 + 9x_2$

Subject to,

$-x_1 + 3x_2 \leq 6$

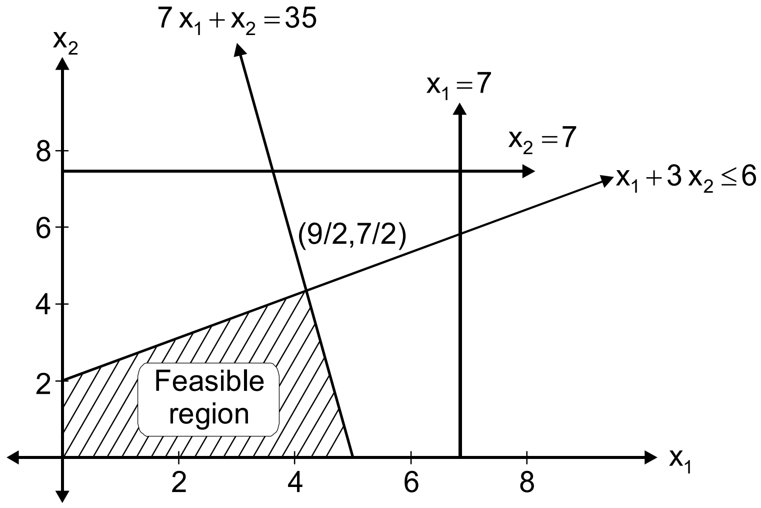
$7x_1 + x_2 \leq 35$

$0 \leq x_1, x_2 \leq 7$

and x_1, x_2 are integers.

Answer

Relaxing the integers requirement the optimal non - integer solution obtained by graphical method is as follows.



$$x_1 = \frac{9}{2}, x_2 = \frac{7}{2}$$

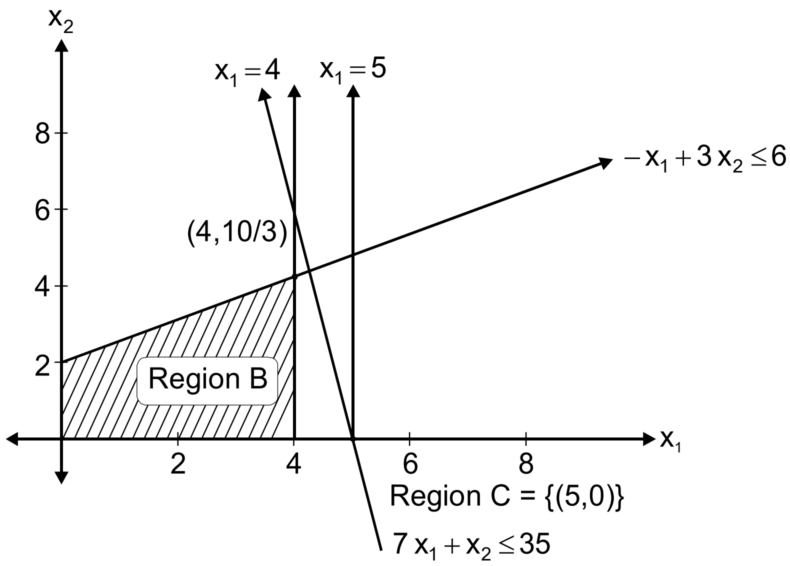
and $z_1 = 7\left(\frac{9}{2}\right) + 9\left(\frac{7}{2}\right) = 63$

Thus $z_u = 63$ and $z_L = 7(4) + 9(3) = 55$

Both x_1 and x_2 are non - integer solution values. Choose $x_1 = \frac{9}{2}$ for dividing the given problem into two sub problems LP - B and LP - C. Two new constraints $x_1 \le 4$ and $x_1 \ge 5$ are added to LP - B and LP - C respectively.

LP - B	LP - C
Max. $z = 7x_1 + 9x_2$	Max. $z = 7x_1 + 9x_2$
Subject to,	Subject to,
$-x_1 + 3x_2 \le 6$	$-x_1 + 3x_2 \le 6$
$7x_1 + x_2 \le 35$	$7x_1 + x_2 \le 35$
$0 \le x_1, x_2 \le 7$	$0 \le x_1, x_2 \le 7$
$x_1 \le 4$	$x_1 \ge 5$
and x_1, x_2 are integers.	and x_1, x_2 are integers.

The solution to sub problem LP - B and LP - C are obtained by graphical method.

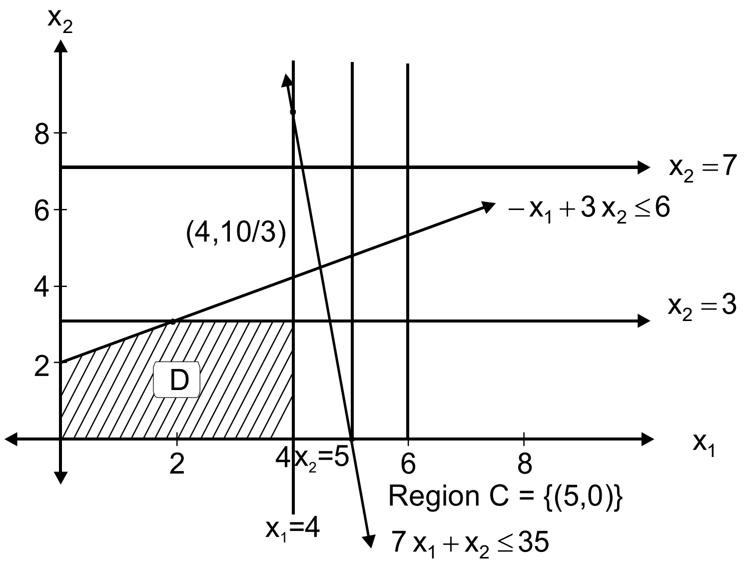


The solution of sub problem LP - B is $x_1 = 4, x_2 = \frac{10}{3}, z_2 = 58$. The feasible region for subproblem LP - C is {(5, 0)}. Therefore the solution of subproblem LP - C is $x_1 = 5, x_2 = 0, z_3 = 35$. Since all the variables have integer values, we stop the search for this subproblem i. e. no further branching is required from node C. The value z = 35 becomes the new lower bounds on the IP problems optimal solution. A non - integer solution of subproblem B and $z_2 > z_3$, both indicate that further branching is necessary from node B.

The sub - problemB is now branched into two new subproblem D and E, and are obtained by additing the constraints $x_2 \le 3$ and $x_2 \ge 4$ (as for problem B, $x_2 = 10/3$).

LP - D	LP - E
Max $Z = 7x_1 + 9x_2$	Max. $Z = 7x_1 + 9x_2$
Subject to,	Subject to,
$-x_1 + 3x_2 \le 6$	$-x_1 + 3x_2 \le 6$
$7x_1 + x_2 \le 35$	$7x_1 + x_2 \le 35$
$0 \le x_1, x_2 \le 7$	$0 \le x_1, x_2 \le 7$
$x_1 \le 4$	$x_1 \le 4$
$x_2 \le 3$	$x_2 \ge 4$

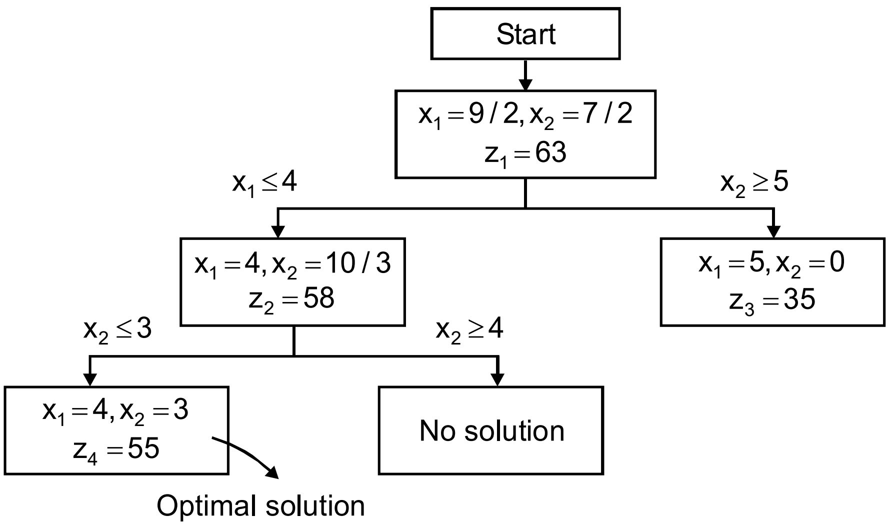
The graphical solutions to LP - D and LP - E are as follows.



There is no feasible region for LP-E, Since $x_1 \le 4$ and $x_2 \ge 4$ do not satisfy $-x_1 + 3\,y_2 \le 6$ as such there is no feasible solution for problem LP - E. The solution of subproblem LP - D is $x_1 = 4, x_2 = 3$ and $z_4 = 55$. Since there is no solution for subproblem LP - E no further branching is required for this subproblem. Since solution to LP - D is an integer solution, no further branching is required for LP - D asa.

Thus finally, we get the optimal solution to the given integer LP problem as z = 55, $x_1 = 4, x_2 = 3$.

The tree - diagram corresponding to this problem is shown in the following figure.



Remark

If the number of variables are more than 2 then exclude the redendent constraints and solve these problems by simplex method and obtain solutions corresponding to each sub - problem.

~ ~ ~ ~ ~ EXERCISE ~ ~ ~ ~ ~

Use branch and bound technique and solve the following integer programming problems.

1) Max. $z = 3x_1 + 3x_2 + 13x_3$

Subject to,

$$-3x_1 + 6x_2 + 7x_3 \le 8$$

$$5x_1 - 3x_2 + 7x_3 \le 8$$

$$0 \le x_j \le 5$$

and all x_j are integer.

2) Max. $z = 3x_1 + x_2$

Subject to,

$$3x_1 - x_2 + x_3 = 12$$

$$3x_1 + 11x_2 + x_4 = 66$$

$$x_j \ge 0, j = 1, 2, 3, 4$$

3) Max. $z = x_1 + x_2$

Subject to,

$$4x_1 - x_2 \le 10$$

$$2x_1 + 5x_2 \le 10$$

$$x_1, x_2 = 0, 1, 2, 3$$

4) Min. $z = 3x_1 + 2.5x_2$

Subject to,

$$x_1 + 2x_2 \ge 20$$

$$3x_1 + 2x_2 \ge 50$$

$x_1, x_2 \ge 0$ and integers.

$\left(\text{Ans.}: x_1 = 14, x_2 = 4, z = 52\right)$

5) Max. $z = 2x_1 + 3x_2$

Subject to,

$$x_1 + 3x_2 \le 9$$

$$3x_1 + x_2 \le 7$$

$$x_1 - x_2 \le 1$$

$x_1, x_2 \ge 0$ and integers.

$\left(\text{Ans.}: x_1 = 0, x_2 = 3, z = 9\right)$

6) Max. $z = 7x_1 + 6x_2$

Subject to,

$$2x_1 + 3x_2 \le 12$$

$$6x_1 + 5x_2 \le 30$$

$x_1, x_2 \ge 0$ and integers.

$\left(\text{Ans.}: x_1 = 5, x_2 = 0, z = 35\right)$

7) Max. $z = 5x_1 + 4x_2$

Subject to,

$$x_1 + x_2 \ge 2$$

$$5x_1 + 3x_2 \le 15$$

$$3x_1 + 5x_2 \le 15$$

and $x_1, x_2 \ge 0$ and integers.

$\left(\text{Ans.}: x_1 = 3, x_2 = 0, z = 15\right)$

8) Max. $z = -3x_1 + x_2 + 3x_3$

Subject to,

$$-x_1 + 2x_2 + x_3 \le 4$$

$$2x_2 - 1.5x_3 \le 1$$

$$x_1 - 3x_2 + 2x_3 \le 3$$

$x_1, x_2 \geq 0$

x_3 - non - negative integers.

$\left(\text{Ans.}: x_1 = 0, x_2 = \frac{8}{7}, x_3 = 1, z = \frac{29}{7}\right)$

9) Max. $z = x_1 + x_2$

Subject to,

$2x_1 + 5x_2 \geq 16$

$6x_1 + 5x_2 \leq 30$

$x_2 \geq 0$

x_1 - non - negative integer.

$\left(\text{Ans.}: x_1 = 4, x_2 = \frac{6}{5}, z = \frac{26}{5}\right)$

10) Max. $z = 110x_1 + 100x_2$

Subject to,

$6x_1 + 5x_2 \leq 29$

$4x_1 + 14x_2 \leq 48$

$x_1, x_2 \geq 0$ and integers. $(\text{Ans.}: x_1 = 4, x_2 = 1, z = 540)$

❑ ❑ ❑ ❑

UNIT 04

INFORMATION THEORY

AIM :

To quantitatively assess the quality of information contained in a piece of information.

OBJECTIVES :

After studying this unit we should be able to -

1. Know various communication processes and parts of communication system.
2. Understand measure of information.
3. Understand applications and axioms of entropy function.
4. Know the basic requirements to be satisfied by the logarithmic form of entropy function.
5. Measure channel capacity, efficiency and redundancy.
6. Apply Shannon-Fano encoding procedure to obtain decodable code of a message.

4.1 INTRODUCTION

Information theory is a branch of probability theory with a large number of applications to communication systems. Mathematical theory of communications was principally initiated by Claude Shannon in 1984.

If something is very likely to occur, the statement that it will occur does not give much information. On the other hand if something is unlikely to occur, the statement that it will occur gives a good deal of information.

The amount of information in the message should be measured by the extent of the change in probability produced by the message.

4.2 COMMUNICATION PROCESSES

Definition 4.1 : The communication process may be defined as the procedure by which one mind affects another is called communication process.

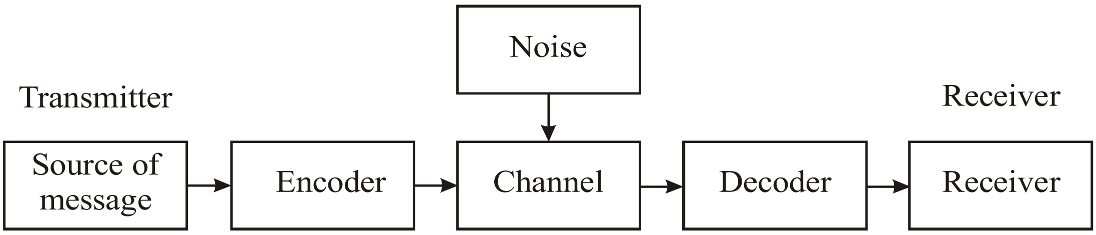
This may be any means by which the information is carried from a transmitter to receiver. There are three essential parts of a simplest communication system.

(i) Transmitter or source.

(ii) Channel or transmission network which carries the message from transmitter to receiver.

(iii) Receiver or sink.

Model for Communication System :



Each part of communication system is explained below.

Definition 4.2 : Transmitter : It is a person or machine which produces the information to be communicated.

Definition 4.3 : Encoder : The device which is used to improve the efficiency of the medium through which the message is transmitted is called encoder.

Encoder acts as step-up transformer.

Definition 4.4 : Channel : It is a medium over which the coded message is transmitted. It is the transmission network (or media) that carries the message from the source to receiver e.g. human voice, newspapers etc.

Definition 4.5 : Decoder : A device which transforms encoded message into the original form which is acceptable to the receiver. This is used to transform encoded message into the original form at the receivers end.

Definition 4.6 : Receiver : This is the destination to which the message is conveyed from the source (or transmitter) through a communication channel.

Definition 4.7 : Noise : This is the general term which creates interruptions or disturbances in the transmission of message (or information) from transmitter to receiver e.g. noise or disturbance in radio or television during the relay of programmes, errors in newspapers printing etc. Noise is anything which tends to produce errors in transmission.

Theorem 4.1 : Fundamental theorem of information theory :

It is possible to transmit information through a noisy channel at any rate less than the channel with an arbitrary small probability of error.

(In the next section the basic concepts in statistics are studied. This section is pre-requisit and is not a part of syllabus).

4.3 BASIC STATISTICS REQUIRED TO STUDY COMMUNICATION SYSTEM

Definition 4.8 :

The set $S = \{e_1, e_2, e_3, ..., e_n\}$ is called a sample space of an experiment satisfying the following two conditions.

1. Each element of the set S denotes one of the possible outcomes.
2. The outcome is one and only one element of the set S whenever the experiment is performed.

Definition 4.9 : If to each elementary event $e_i \in S$ we assign a real number $p(e_i)$ is called the probability of an elementary event e_i such that

(i) $p(e_i) \geq 0$, $\forall i$ and (ii) $\sum_{i=1}^{n} p(e_i) = 1$

Such an assignment is called acceptable assignment.

Definition 4.10 : Events :

1. Event : Every subset of the sample S of an experiment is called an event generally denoted by E.

2. Simple Event : Any event that contains only one element is called simple event.

3. Equally Likely Event : If there is no reason to expect any one in preference to any other i.e. probability of happenng two or more events is same.

4. Mutually Exclusive Events : The happening of one excludes the happening of the other.

5. Dependent and Independent Events : Two events are said to be independent when occurence of one has no effect on the probability of other.

Probability : With each even E_i in a finite sample space S, we associate a real number say $p(E_i)$ called the probability of an event E_i satisfying the following conditions.

(i) $0 \le P(E_i) \le 1$

(ii) $P(E_1 \cup E_2 \cup E_3 ..... \cup E_n) = P(E_1) + P(E_2) + P(E_3) + ... + P(E_n)$

where $E_1, E_2, ..., E_n$ are mutually exclusive.

(iii) $P(S) = 1$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Conditional Probability

Consider the two events E_1 and E_2 in a sample space S. Here E_1 represents the event that has occured m_1 number of time in n number of trails and E_2 represents event that has occured m_2 number of times out of these m_1 number of occurances of E_1. The probability of combine happening of E_1 and E_2 in the same trial is

$$P(E_1 \cap E_2) = \frac{m_1}{n} \cdot \frac{m_2}{m_1} = \frac{m_2}{n}$$

The relative frequency $\frac{m_2}{m_1}$ is the conditional probability of occurance of event E_2 given that E_1 has occured. Which is denoted by $P(E_2 / E_1)$, $(P(E_1) \neq 0)$.

$$P(E_1 \cap E_2) = P(E_1) P(E_2 / E_1)$$

Definition 4.11 : Conditional probability of E_2 given E_1 is

$$P(E_2 / E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)}$$

$P(E_2 / E_1)$ satisfies following properties.

(i) $0 \le P(E_2 / E_1) \le 1$

(ii) If E_2 is an event which cannot occur then $P(E_2 / E_1) = 0$.

(iii) If the event E_2 is entire space S then $P(E_2 / E_1) = P(S / E_1) = 1$.

(iv) If E_2 and E_3 are independent events then $P(E_2 \cap E_3 / E_1) = P(E_2 / E_1) + P(E_3 / E_1)$.

(v) If occurance of E_1 does not affect the occurance of E_2 then $P(E_2/E_1) = P(E_2)$.

E_1, E_2 are independent events if and only if $P(E_1 \cap E_2) = P(E_1) \cdot P(E_2)$.

Law of Probability :

$$P(E_1 \cap E_2 \cap E_3 \cap ....\cap E_n) = P(E_1)P(E_2/E_1)P(E_3/E_1E_2)....(E_n/E_1E_2...E_{n-1})$$

4.4 STATISTICAL NATURE OF COMMUNICATION SYSTEMS

In communication system the source selects and transmits of symbols from a given alphabet to the channel, based on some statistical rule. The channel also transmits this symbolic information to the receiver under some statistical rule.

Memoryless Channel :

A memoryless channel os described by an input alphabet $A = \{x_1, x_2, ..., x_m\}$ and output alphabet $B = \{y_1, y_2, y_3, ..., y_n\}$ and a set of conditional probabiloities $P(y_j/x_i) \forall i, j$, where $P(y_j/x_i)$ is the probability that the output symbol y_j will be received in, the input x_i is sent.

Definition 4.12 : Binary Symmetric Channel :

A binary symmetric channel has two input symbols $(x_1 = 0, x_2 = 1)$ and two output symbols $(y_1 = 0, y_2 = 1)$ and it is symmetric in the same that $P(y_1/x_1) = P(y_2/x_2) = \overline{p}$, $P(y_1/x_2) = P(y_2/x_1) = p$ where $\overline{p} = 1 - p \cdot p$ being the probability of error transmission.

The Channel Matrix :

The input to the channel, the output from the channel and conditional probabilities for a pair of symbols can be expressed in the form of a matrix called channel matrix.

Output

$$A = \text{Input} \begin{array}{c} \\ x_1 \\ x_2 \\ \vdots \\ x_m \end{array} \begin{array}{c} \begin{array}{cccc} y_1 & y_2 & \cdots & y_n \end{array} \\ \begin{bmatrix} p_{1/1} & p_{2/1} & \cdots & p_{n/1} \\ p_{1/2} & p_{2/2} & \cdots & p_{n/2} \\ \vdots & & & \\ p_{1/m} & p_{2/m} & \cdots & p_{n/m} \end{bmatrix} \end{array}$$

Where $p_{j/i} = p(y_j/x_i)$, i = 1, 2, ..., m, j = 1, 2, ..., n.

e.g. The channel matrix of the Binary symmetric channel is

$$\begin{bmatrix} \overline{p} & p \\ p & \overline{p} \end{bmatrix}$$

Probability Relation in a Channel

If $p_{i0} = p(x_i)$ denotes the probability that the symbol x_i is selected for transmission, $p_{0j} = p(y_j)$ the probability that the symbol y_j is received then the relation between the probabilities of various input symbols and output symbols is expressed as

$$\sum_{i=1}^{m} p_{i0} p_{j/i} = p_{0j}\text{ , } i = 1, 2, \ldots, n \qquad \ldots\ldots \text{(i)}$$

(i) The joint probabilities of sending a symbol x_i and receiving the symbol y_j is given by

$$p(x_i, y_j) = p_{j/i} p_{i0} \quad \forall i, j \qquad \ldots\ldots \text{(ii)}$$

(ii) The conditional backward input probabilities when it is known that the symbol y_j has been received is

$$p(x_i / y_j) = p_{j/i}(p_{i0} / p_{0j}) \quad \forall i, j \qquad \ldots\ldots \text{(iii)}$$

The relations (ii) give the joint probabilities of sending a symbol x_i and receiving a symbol y_j while relation (iii) give the backward channel probabilities given that an output y_j has been received.

Example 1 : Consider a binary channel with input symbol $X\{0,1\}$, output symbol $Y\{0,1\}$ and the channel matrix $\begin{bmatrix} 2/3 & 1/3 \\ 1/10 & 9/10 \end{bmatrix}$. Let us assume that the input probabilities $p_{10} = 3/4$ and $p_{20} = 1/4$. Then

$$p_{01} = \sum_{i=1}^{2} p_{io} p_{1/i} = \frac{3}{2} \cdot \frac{2}{3} + \frac{1}{4} \cdot \frac{1}{10} = \frac{21}{40}$$

$$p_{02} = \sum_{i=1}^{2} p_{io} p_{2/i} = \frac{3}{4} \cdot \frac{1}{3} + \frac{1}{4} \cdot \frac{9}{10} = \frac{19}{40}$$

The conditional backward input probabilities are obtained by using (iii),

$$p(0/0)=\frac{\frac{2}{3}\cdot\frac{3}{4}}{\frac{21}{40}}=\frac{20}{21}; \qquad p(0/1)=\frac{\frac{1}{3}\cdot\frac{3}{4}}{\frac{19}{40}}=\frac{10}{19}$$

$$p(1/0)=\frac{\frac{1}{10}\cdot\frac{1}{4}}{\frac{21}{40}}=\frac{1}{21}; \qquad p(1/1)=\frac{\frac{9}{10}\cdot\frac{1}{4}}{\frac{19}{40}}=\frac{9}{19}$$

The joint probabilities are obtained by relation (ii)

$$p(x_i, y_j)=p_{j/i}p_{i0}$$

$$p(0,0)=\frac{2}{3}\cdot\frac{3}{4}=\frac{1}{2} \qquad p(0,1)=\frac{1}{3}\cdot\frac{3}{4}=\frac{1}{4}$$

$$p(1,0)=\frac{1}{10}\cdot\frac{1}{4}=\frac{1}{40} \qquad p(1,1)=\frac{9}{10}\cdot\frac{1}{4}=\frac{9}{40}$$

Example 2 : Consider a binary channel with input symbol $X\{0,1\}$, output symbol $Y\{0,1\}$ and the channel matrix $\begin{bmatrix} 1/3 & 2/3 \\ 1/5 & 4/5 \end{bmatrix}$. Further assume the input probabilities $p_{10}=\frac{6}{7}$ and $p_{20}=\frac{1}{7}$.

$$p_{10}=\sum_{i=1}^{2}p_{i0}p_{1/i}=\frac{6}{7}\cdot\frac{1}{3}+\frac{1}{7}\cdot\frac{1}{5}=\frac{11}{35}$$

$$p_{02}=p_{10}p_{2/1}+p_{20}p_{2/2}=\frac{6}{7}\cdot\frac{2}{3}+\frac{1}{7}\cdot\frac{4}{5}=\frac{24}{35}$$

The joint probabilities are obtained by using (ii)

$$p(x_i, y_j)=p_{j/i}p_{i0}$$

$$p(0,0)=\frac{1}{3}\cdot\frac{6}{7}=\frac{2}{7}; \qquad p(0,1)=\frac{2}{3}\cdot\frac{6}{7}=\frac{4}{7}$$

$$p(1,0)=\frac{1}{5}\cdot\frac{1}{7}=\frac{1}{35}; \qquad p(1,1)=\frac{4}{5}\cdot\frac{1}{7}=\frac{4}{35}$$

The conditional backward input probabilities are obtained by using (iii)

$$p(0/0)=\frac{\frac{1}{3}\cdot\frac{1}{7}}{\frac{11}{35}}=\frac{10}{11} \qquad p(0/1)=\frac{\frac{2}{3}\cdot\frac{6}{7}}{\frac{24}{35}}=\frac{5}{6}$$

$$p(1/0)=\frac{\frac{1}{5}\cdot\frac{1}{7}}{\frac{11}{35}}=\frac{1}{11} \qquad p(1/1)=\frac{\frac{4}{5}\cdot\frac{1}{7}}{\frac{24}{35}}=\frac{1}{6}$$

Noisless Channel :

If the channel matrix contains only one non-zero element in each column, then such a channel is called a noiseless channel.

e.g. (i) Binary symmetric channel with p = 0 or 1.

(ii) $\begin{bmatrix} 1/2 & 0 & 1/6 & 0 \\ 0 & 5/7 & 0 & 0 \\ 0 & 0 & 0 & 2/3 \end{bmatrix}$ (iii) $\begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3/5 & 5/10 & 1/10 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

4.5 A MEASURE OF INFORMATION

Basic Assumptions :

(i) There is a finite set $X=\{x_1,x_2,....,x_n\}$ of events and the probability of occurence of event x_i is p_i, such that $p_1+p_2+....+p_n=1$. Consider the event x_k has occured. So according to the statement that the quantity of information received is inversely proportional to the likelihood of the event, if $I(x_k)$ denote the amount of information received from the occurence of the even x_k with probability p_k of occurence then $I(x_k)>I(x_r)$ for $p_k<p_r$.

(ii) If x_k has $p_k=1$ then $I(x_k)=0$ because no information is received provided occurence of a particular event is known in advance.

If we are interested in the probabilities of the occurence of an event in set X and not in their actual natures, the expected value of information received may be written as

$$\sum_{i=1}^{m} p_i I(p_i)$$

Where $I(p_i)=-\log_2 p_i$.

The $\log_2 p_i$ indicates the probability concerning the receiver before receiving the information, provided the fact that the communication system is noiseless.

Example 1 : Suppose a baby is born at a neighbour's house and the question is asked whether the baby is a boy or girl ? The answer 'It is a boy' gives specific amount as information.

Since the probability that it is a boy is $p = \frac{1}{2}$ therefore the amount of information is

$$-\log p = -\log \frac{1}{2} = \log_2 2 = 1 \text{ bit}$$

Example 2 : A large field is divided into 64 squares (8×8). In the dark night a cow has entered in this field. This cow is located by a member of searching party who sends back an information giving the location of the cow as the 43rd square. Calculate the amount of information obtained in the reception of this message.

Answer : Before the message was received, the probability that the cow was in 43rd square = $\frac{1}{64}$. The quantity of information received with a message is

$$-\log p = -\log \frac{1}{64} = \log_2 64 = 6 \log_2 2 = 6 \text{ bits}$$

Alternatively, if the message received was, the cow was in the square 6th row and 3rd column of the field, then before the information was received the cow was equally likely to have been in any of the different columns i.e. $p_c = \frac{1}{8}$.

Similarly the probability that it was in row 6 is $p_r = \frac{1}{8}$. A two symbol information will be sent in which the first symbol stands for row and second for column.

The first symbol gives $-\log p_c = -\log \frac{1}{8} = \log 8 = 3$ bits of information and the second symbol gives $-\log p_r = -\log \frac{1}{8} = \log_2 8 = 3$ bits of information.

Thus the total amount of information is 3 + 3 = 6 bits.

The name given to the unit of information is bit.

Choice of Measure :

The expected value of information can be interpreted as the expected amount of information needed to determine which event of set X has occured. It is the measure of uncertainty regarding which event of X has occured or will occur. The uncertainty is considered to be maximum when each event $x_1, x_2, ..., x_m$ of X are equally probable i.e.

$$p(x_1) = p(x_2) = ... = p(x_m) = \frac{1}{m}$$

Shannon and Wiener have suggested the following expression as the measure of expected amount of information.

$$H(p_1, p_2, ..., p_m) = -\sum_{i=1}^{m} p_i \log p_i$$

The function H is also known as entropy function.

4.7 PROPERTIES OF ENTROPY FUNCTION

1. Continuity : The entropy function $H(p_1, p_2, ..., p_n)$ is continuous for each independent variable p_i, $0 \le p_i \le 1$.

Proof : $-H(p_1, p_2, ..., p_0) = p_1 \log p_1 + p_2 \log p_2 + ... + p_{n-1} \log p_{n-1} + p_n \log p_n$

$$= p_1 \log p_1 + p_2 \log p_2 + ... + p_{n-1} \log p_{n-1} +$$

$$(1 - p_1 - p_2 ...... - p_{n-1}) \log (1 - p_1 - p_2 ..... - p_{n-1})$$

Since all p_i are continuous and independent variables i = 1, 2, ..., n–1 in the interval (0, 1), therefore the log of continuous function is also continuous.

2. Symmetry : The entropy function is symmetric function in all variables.

$$H(p_1, p_2, ..., p_n) = -\sum_{i=1}^{n} p_i \log p_i$$

$H(p_1, p_2) = H(p_2, p_1)$ for $p_1 + p_2 = 1$

3. Maximum Value Property

$$\text{Max}\, H(p_1, p_2, p_3, ..., p_n) = H\left(\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right)$$

4. Additivity : If a particular event x_n with probability p_n is divided into m mutually exclusive subsets say $e_1, e_2, e_3, ..., e_m$ with probabilities $q_1, q_2, ..., q_m$ respectively such that $p_n = q_1 + q_2 + q_3 + ... + q_m$ then

$$H(p_1, p_2, ..., p_{n-1}, q_1, q_2, q_3, ..., q_m) = H(p_1, p_2, p_3, ..., p_{n-1}, p_n) + p_n H\left(\frac{q_1}{p_n}, \frac{q_2}{p_n}, ..., \frac{q_m}{p_n}\right)$$

Proof : Since the event x_n with probability p_n is divided into disjoint subsets $e_1, e_2, e_3, ..., e_m$ with respective probabilities $q_1, q_2, q_3, ..., q_m$; $p_n = q_1 + q_2 + ... + q_m$.

$$H(p_1, p_2, ..., p_{n-1}, q_1, q_2, q_3, ..., q_n) = -\sum_{k=1}^{n-1} p_k \log p_k - \sum_{i=1}^{m} q_i \log q_i$$

$$= -\left(\sum_{k=1}^{n} p_k \log p_k - p_n \log p_n\right) - \sum_{i=1}^{m} q_i \log q_i$$

$$= H(p_1, p_2, ..., p_n) + p_n \log p_n - \sum q_i \log q_i \qquad .... \text{(i)}$$

$$p_n \log p_n - \sum_{i=1}^{m} q_i \log q_i = p_n\left(\frac{p_n}{p_n} \log p_n\right) - p_n \sum_{i=1}^{m} \frac{q_i}{p_n} \log q_i$$

$$= p_n\left[\frac{\sum_{i=1}^{m} q_i}{p_n} \log p_n\right] - p_n \sum_{i=1}^{m} \frac{q_i}{p_n} \log q_i$$

$$= p_n\left[\sum_{i=1}^{m} \frac{q_i}{p_n}(\log p_n - \log q_i)\right]$$

$$= p_n\left[\sum_{i=1}^{m} \frac{q_i}{p_n}\left(-\log\left(\frac{q_i}{p_n}\right)\right)\right]$$

$$= p_n\left[-\sum_{i=1}^{m} \frac{q_i}{p_n} \log\left(\frac{q_i}{p_n}\right)\right]$$

$$= p_n H\left(\frac{q_1}{p_n}, \frac{q_2}{p_n}, ..., \frac{q_n}{p_n}\right) \qquad .... \text{(iii)}$$

Theorem 4.2 : Let $p_1, p_2, ..., p_m$ and $q_1, q_2, q_3, ..., q_m$ be arbitrary non-negative numbers with $\sum_{i-1}^{m} p_i = \sum_{i-1}^{m} q_i$.

Then $-\sum_{i=1}^{m} p_i \log p_i \leq - = \sum_{i=1}^{m} p_i \log q_i$ with equality iff $q_i = p_i \quad \forall i$.

Proof : Since log is convex function $\log x \leq x - 1$ with equality iff *x* = 1.

For $x = \frac{q_i}{p_i}$, $\log\left(\frac{q_i}{p_i}\right) \leq \left(\frac{q_i}{p_i}\right) - 1$ with equality iff $\frac{q_i}{p_i} = 1$ i.e. $q_i = p_i$.

$$\therefore \sum_{i=1}^{m} p_i \log\left(\frac{q_i}{p_i}\right) \leq \sum_{i=1}^{m}\left(\frac{q_i}{p_i}\right) - 1 = \frac{\Sigma(q_i - p_i)}{\Sigma p_i} \qquad (\because \Sigma q_i = \Sigma p_i)$$

Thus $\sum_{i=1}^{m} p_i \log\left(\frac{q_i}{p_i}\right) \leq 0$

$$\sum_{i=1}^{m} p_i \left[\log q_i - \log p_i\right] \leq 0 \Rightarrow \sum_{i=1}^{m} p_i \log q_i \leq \sum_{i=1}^{m} p_i \log p_i$$

i.e. $-\sum_{i=1}^{m} p_i \log p_i \leq -\sum_{i=1}^{m} p_i \log q_i$

Example : Evaluate them average uncertainty associated with the sample of events A, B and C which are mutually exclusive with probability distribution.

Event	:	A	B	C
Probability	:	1/5	4/15	8/15

Solution : We have $p_1 = \frac{1}{5}$, $p_2 = \frac{4}{15}$, $p_3 = \frac{8}{15}$

$$H(p_1, p_2, p_3) = p_1 \log p_1 - p_2 \log p_2 - p_3 \log p_3$$

$$= -\frac{1}{5}\log\left(\frac{1}{5}\right) - \frac{4}{15}\log\left(\frac{4}{15}\right) - \frac{8}{15}\log\left(\frac{8}{15}\right)$$

$$= \frac{1}{15}\left[3\log\left(\frac{1}{5}\right) + 4\log\left(\frac{4}{15}\right) + 8\log\left(\frac{8}{15}\right)\right]$$

$$= \frac{1}{15}\left[3(\log 1 - \log 5) + 4(\log 4 - \log 15) + 8(\log 8 - \log 15)\right]$$

$$= -\frac{1}{15}\left[-3\log 5 + 4\log(2^2) - 4\log(3 \times 5) + 8\log(2^3) - 8\log(3 \times 5)\right]$$

$$= -\frac{1}{15}\left[-3\log 5 + 4\log 2^2 + 8\log 2^3 - 12\log 3 - 12\log 5\right]$$

$$= -\frac{1}{15}\left[-15\log 5 - 12\log 3 + 32\right]$$

$$= \log 5 + \frac{4}{5}\log 3 - \frac{32}{5}$$

Example : Evaluate the average uncertainty associated with the probability space of events A, B, C, D which are mutually exlcusive with probability distribution.

Event	:	A	B	C	D
Probability	:	1/2	1/4	1/8	1/8

Solution : We have $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{4}$, $p_3 = \frac{1}{8}$, $p_4 = \frac{1}{8}$.

$$H(p_1, p_3, p_3, p_4) = p_1 \log p_1 - p_2 \log p_2 - p_3 \log p_3 - p_4 \log p_4$$

$$\therefore H\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{4}\log\frac{1}{4} - \frac{1}{8}\log\frac{1}{8} - \frac{1}{8}\log\frac{1}{8}$$

$$= +\frac{1}{2}\log 2 + \frac{1}{4}\log 4 + \frac{1}{8}\log 8 + \frac{1}{8}\log 8$$

$$= \frac{1}{2}\log 2 + \frac{1}{4}\log 2^2 + \frac{1}{8}\log 2^3 + \frac{1}{8}\log 2^3$$

$$= \left(\frac{1}{2} + \frac{2}{4} + \frac{3}{8} + \frac{3}{8}\right)\log 2$$

$$= \frac{1}{8}(4 + 4 + 3 + 3)\log 2$$

$$= \frac{14}{8}\log 2 = \frac{7}{4} \text{ bits}$$

4.8 JOINT AND CONDITIONAL ENTROPIES

Consider two sets of messages

$X = \{x_1, x_2, x_3, ..., x_m\}$ and $Y = \{y_1, y_2, y_3, ..., y_n\}$

Where x_i's are the messages send (message input) and y_i's are the messages received (channel output).

Let $p_{ij} = P\left(X = x_i, Y = y_j\right)$, *i* = 1, 2, 3, ..., *m*; *j* = 1, 2, 3, ..., *n* denote the probability of the joint event that message x_i is sent and message y_j is received.

Define marginal probability distributions of X and Y by $p_{i0} = \sum_{j=1}^{n} p_{ij}$ and $p_{0j} = \sum_{i=1}^{m} p_{ij}$ $\forall i, j$ respectively

Definition 4.13 : The marginal entropies of the two marginal distribution are given by,

$$H(X) = -\sum_{i=1}^{m} p_{i0} \log p_{i0} \quad \text{and} \quad H(Y) = -\sum_{j=1}^{n} p_{0j} \log p_{0j}$$

The entropy H(X) measures the uncertainty of the message sent (irrespective of message received) and H(Y) performs the same role for the message received.

Definition 4.14 : The joint entropy is the entropy of the joint distributions of the messages sent and received and is given by,

$$H(X,Y) = -\sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} \log p_{ij}$$

Definition 4.3 : $H(X,Y) \le H(X) + H(Y)$ with equality off X and Y are independent.

Proof : $H(X) + H(Y) = -\sum_{i=1}^{m} p_{i0} \log p_{i0} - \sum_{j=1}^{n} p_{0j} \log p_{0j}$

$$= -\sum_{i=1}^{m} \left(\sum_{j=1}^{n} p_{ij} \right) \log p_{i0} - \sum_{j=1}^{n} \left(\sum_{i=1}^{m} p_{ij} \right) \log p_{0j}$$

$$= -\sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} \left(\log p_{i0} + \log p_{0j} \right)$$

$$= -\sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} \log \left(p_{i0} p_{0j} \right)$$

$$H(X) + H(Y) = -\sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} \log\, q_{ij} \qquad \text{where } q_{ij} = p_{i0} p_{0j} \qquad \text{.... (i)}$$

By definition $H(X,Y) = -\sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} \log p_{ij}$ .... (ii)

But $\sum_{i=1}^{m}\sum_{j=1}^{n} q_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{n} p_{i0}p_{0j} = \left(\sum_{i=1}^{m} p_{i0}\right)\left(\sum_{j=1}^{n} p_{0j}\right) = 1 = \sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}$

By thereom 4.2, since

$$\sum_{i=1}^{m}\sum_{j=1}^{n} q_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij} = 1, \quad -\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}\log p_{ij} \le -\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}\log q_{ij}$$

Thus from equation (i) and (ii) we have $H(X,Y) \le H(X) + H(Y)$.

Example : A transmitter has an alphabet consisting of five letters $\{x_1, x_2, x_3, x_4, x_5\}$ and the receiver has an alphabet consisting of four letters $\{y_1, y_2, y_3, y_4\}$. The joint probabilities for the communication are given by,

$$\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array}\begin{array}{c} \begin{array}{cccc} y_1 & y_2 & y_3 & y_4 \end{array} \\ \begin{bmatrix} 0.25 & 0.00 & 0.00 & 0.00 \\ 0.10 & 0.30 & 0.00 & 0.00 \\ 0.00 & 0.05 & 0.10 & 0.00 \\ 0.00 & 0.00 & 0.05 & 0.10 \\ 0.00 & 0.00 & 0.05 & 0.00 \end{bmatrix} \end{array}$$

Determine the marginal, conditional and joint entropies for this channel (assume that 0 log 0 = 0)

Solution : The channel is described here with joint probability p_{ij}, i = 1, 2, 3, 4, 5 and j = 1, 2, 3, 4,

By definition, marginal probability distributions of X and Y are

$$p_{i0} = \sum_{j=1}^{n} p_{ij} \quad \text{and} \quad p_{0j} = \sum_{i=1}^{m} p_{ij}$$

Here $p_{10} = \sum_{j=1}^{n} p_{0j} = p_{11} + p_{12} + p_{13} + p_{14}$

= 0.25 + 0.00 + 0.00 + 0.00 = 0.25

Similarly,

p_{20} = 0.10 + 0.30 + 0.00 + 0.00 = 0.40

p_{30} = 0.00 + 0.05 + 0.10 + 0.00 = 0.15

p_{40} = 0.00 + 0.00 + 0.05 + 0.10 = 0.15

p_{50} = 0.00 + 0.00 + 0.05 + 0.00 = 0.05

Since $p_{0j} = \sum_{i=1}^{m} p_{ij}$

$p_{01} = p_{11} + p_{21} + p_{31} + p_{41} + p_{51}$

= 0.25 + 0.10 + 0.00 + 0.00 + 0.00 = 0.35

Similarly,

p_{02} = 0.00 + 0.30 + 0.05 + 0.00 + 0.00 = 0.35

p_{03} = 0.00 + 0.00 + 0.10 + 0.05 + 0.05 = 0.2

p_{04} = 0.00 + 0.00 + 0.00 + 0.10 + 0.00 = 0.1

$$p_{j/i} = \frac{p_{ij}}{p_{i0}} \begin{bmatrix} \frac{0.25}{0.25} & \frac{0.00}{0.25} & \frac{0.00}{0.25} & \frac{0}{0.25} \\ \frac{0.1}{0.4} & \frac{0.3}{0.4} & \frac{0.00}{0.4} & \frac{0.00}{0.4} \\ \frac{0.0}{0.15} & \frac{0.05}{0.15} & \frac{0.1}{0.15} & \frac{0}{0.15} \\ \frac{0.00}{0.15} & \frac{0.00}{0.15} & \frac{0.05}{0.15} & \frac{0.10}{0.15} \\ \frac{0.00}{0.05} & \frac{0.00}{0.05} & \frac{0.05}{0.05} & \frac{0.00}{0.05} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.25 & 0.75 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Marginal Entropies :

$$H(X) = -\sum_{i=1}^{5} p_{i0} \log p_{i0}$$

$$= -\{0.25\log 0.25 + 0.4\log 0.4 + 0.15\log 0.15 + 0.15\log 0.15 + 0.05\log 0.05\}$$

$$= -\left\{\frac{1}{4}\log\frac{1}{4} + \frac{2}{5}\log\frac{2}{5} + \frac{3}{20}\log\frac{3}{20} + \frac{3}{20}\log\frac{3}{20} + \frac{1}{20}\log\frac{1}{20}\right\}$$

$$= \frac{1}{4}\log 4 + \frac{2}{5}\log\frac{5}{2} + \frac{3}{10}\log\frac{20}{3} + \frac{1}{20}\log 20$$

$$= \frac{1}{4}\log 2^2 + \frac{2}{5}\left(\log 5 - \log 2\right) + \frac{3}{10}\left(\log 2^2 + \log 5 - \log 3\right) + \frac{1}{20}\log\left(2^2 \times 5\right)$$

$$= \frac{1}{2}\log 2 + \frac{2}{5}\log 5 - \frac{2}{5}\log 2 + \frac{3}{5}\log 2 + \frac{3}{10}\log 5 - \frac{3}{10}\log 3 + \frac{1}{10}\log 2 + \frac{1}{20}\log 5$$

$$= \log 2\left[\frac{1}{2} - \frac{2}{5} + \frac{3}{5} + \frac{1}{10}\right] + \log 5\left[\frac{2}{5} + \frac{3}{10} + \frac{1}{20}\right] + \log 3\left[-\frac{3}{10}\right]$$

$$= \frac{1}{10}(5 - 4 + 6 + 1) + \frac{1}{20}(8 + 6 + 1)\log 5 - \frac{3}{10}\log 3$$

$$= \frac{4}{5}\log 2 + \frac{3}{4}\log 5 - \frac{3}{10}\log 3$$

$$= \frac{3}{4}\log 5 - \frac{3}{10}\log 3 + \frac{4}{5} \text{ bits}$$

$$H(Y) = -\sum_{j=1}^{4} p_{0j}\log p_{0j}$$

$$= -\left[0.35\log 0.35 + 0.35\log 0.35 + 0.2\log 0.2 + 0.1\log 0.1\right]$$

$$= -\left[0.7\log\frac{7}{20} + 0.2\log\frac{1}{5} + 0.1\log\frac{1}{10}\right]$$

$$= 0.7\log\frac{20}{7} + 0.2\log 5 + 0.1\log 10$$

$$= 0.7\log\left(\frac{2^2 \times 5}{7}\right) + 0.2\log 5 + 0.1\log(2 \times 5)$$

$$= 0.7\left[2\log 2 + \log 5 - \log 7\right] + 0.2\log 5 + 0.1\left(\log 2 + \log 5\right)$$

$$= \log 2(1.4 + 0.1) + \log 5(0. + 0.2 + 0.1) - 0.7\log 7$$

$$= 1.4\log 2 + \log 5 - 0.7\log 7$$

Conditional Entropy :

$$H(Y/X) = -\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij} \log p_{j/i}$$

$$= -p_{11}\log p_{1/1} - p_{21}\log p_{2/1} - p_{22}\log p_{2/2}$$

$$-p_{32}\log p_{3/2} - p_{33}\log p_{3/3} - p_{43}\log p_{4/3}$$

$$-p_{44}\log p_{4/4} - p_{53}\log p_{5/3}$$ (Since remaining $p_{ij} = 0$)

$$= 0.25\log 1 + 0.1\log 4 + 0.3\log\frac{4}{3} + 0.05\log 3 + 0.1\log\frac{3}{2} +$$

$$0.05\log 3 + 0.1\log\frac{3}{2} + 0.05\log 1$$

$$= 0.2\log 2 + 0.3\left(\log 2^2 - \log 3\right) + 0.05\log 3 + 0.1\left(\log 3 - \log 2\right)$$
$$+0.05\log 3 + 0.1\left(\log 3 - \log 2\right)$$

$$= \log 2(0.2 + 0.6 - 0.1 - 0.1) + \log 3(-0.3 + 0.05 + 0.1 + 0.05 + 0.1)$$

$$= 0.6\log 2 = 0.6 \text{ Bits}$$

Joint Entropy :

$$H(X,Y) = H(X) + H(Y/X)$$

$$= \frac{3}{4}\log 5 - \frac{3}{10}\log 3 + \frac{4}{5} + 0.6\log 2$$

$$= \frac{3}{4}\log 5 - \frac{3}{10}\log 3 + \frac{14}{10}$$

Set Axioms For Entropy Function :

Assume the following four conditions as axioms :

1. Given a finite complete probability scheme $\left(p_1, p_2, p_3, ..., p_n\right)$

$$\text{Max } H\left(p_1, p_2, p_3, ..., p_n\right) = H\left(\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right)$$

2. For a joint finite complete scheme, associated entropies should satisfy

$$H(X,Y) = H(X) + H(Y/X)$$

3. $H\left(p_1, p_2, p_3, ..., p_n, 0\right) = H\left(p_1, p_2, p_3, ..., p_n\right)$

4. The entropy function is continuous with respect to all its arguments.

These axioms essentially lead to a unique expression for entropy of finite scheme.

4.7 UNIQUENESS THEOREM

Theorem 4.4 : The only function which satisfies for axioms for entropy function is

$$H(p_1, p_2, ..., p_n) = \lambda \sum_{i=1}^{n} p_i \log p_i$$

where λ is an arbitrary positive number and the logarithm base is any number greater than 1.

Proof : Consider $H\left(\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right) = f(n)$

Step 1 : To show that $f(n) = \lambda \log n$

Since $f(n) = H\left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right) \le H\left(\frac{1}{n+1}, \frac{1}{n+1}, ..., \frac{1}{n+1}\right) = f(n+1)$

$f(n)$ is non-decreasing function of *n*.

According to axiom (2), for any complete probability scheme consisting of the sum of *m* mujtually exclusive schemes.

$$H(x_1, x_2, x_3, ..., x_m) = H(X_1) + H(X_2) + ... + H(X_m) = \sum_{i=1}^{m} H(X_i)$$

If each scheme consists of *r* equally likely events then

$$H(X_1, X_2, ..., X_m) = mf(r) = f(r^m)$$

where *m* and *r* are any arbitrary integers.

Now take two integers *t* and *n* such that

$$r^m < t^n < r^{m+1}$$

$$m \log r < n \log t < (m+1) \log r$$

$$\therefore \frac{m}{r} < \frac{\log t}{\log r} < \frac{m+1}{n} \qquad .... \text{(i)}$$

Since $f(n)$ is non decreasing function

$$f(r^m) \le f(t^n) \le f(r^{m+1})$$

i.e. $mf(r) \le nf(t) \le (m+1) f(r)$

$$\frac{m}{n} \le \frac{f(t)}{f(r)} \le \frac{m+1}{n} \qquad \text{.... (ii)}$$

Comparing (i) and (ii) we have

$$\left| \frac{f(t)}{f(r)} - \frac{\log t}{\log r} \right| \le \frac{1}{n}$$

As $n \to \infty$, for any positive integers *r* and *t*,

$$\frac{f(t)}{f(r)} - \frac{\log t}{\log r} = 0$$

Thus $\frac{f(t)}{f(r)} = \frac{\log t}{\log r} = \lambda$ (say)

$f(t) = \lambda \log t$, $f(r) = \lambda \log r$ (say)

Thus $f(n) = \lambda \log n$.

Since $f(t)$ is non decreasing λ must be positive.

This proves the uniqueness theorem for particular case when all events have equal probabilities.

Step 2 : All probabilities are rational number (but not necessarily all equal).

Let a be a common denominator for the different rational p_k and $p_k = \frac{\alpha_k}{\alpha}$, $\sum \alpha_i = \alpha$, $\alpha_k > \alpha$.

Consider a probability scheme X.

Let the scheme Z consists of a equally likely events $\{z_1, z_2, z_3, ..., z_\alpha\}$.

Decompose these events into groups containing events $\alpha_1, \alpha_2, \alpha_3, ..., \alpha_{n-1}$ and α_n.

Denote the decomposed scheme by Y. When the event X_k with probability p_k occurs, all events partioned in the k^{th} group occur with equal probability in scheme Y. Thus occurs, all events partitioned in the k^{th} group occur with equal probability in scheme Y. Thus

$$H\left(\frac{1}{\alpha_k}, \frac{1}{\alpha_k}, ..., \frac{1}{\alpha_k}\right) = \lambda \log \alpha_k$$

$$= \lambda\left(\log p_k + \log \alpha\right)$$

$$H(Y/X) = \sum p_k H\left(\frac{1}{\alpha_k}, \frac{1}{\alpha_k}, \ldots, \frac{1}{\alpha_k}\right)$$

$$= \lambda \sum p_k \log p_k + \lambda \log \alpha$$

The totality of events in Z forms the sum of two schemes.

$$H(Z) = H(X,Y) = f(\alpha) = \lambda \log \alpha$$

$$H(X,Y) = H(X) + H(Y/X)$$

$$\therefore H(X) = H(X,Y) - H(Y/X)$$

$$= \lambda \log \alpha - (\lambda \sum p_k \log p_k + \lambda \log \alpha)$$

$$= -\lambda \sum p_k \log p_k$$

Thus, the uniqueness theorem also helds when $p_1, p_2, \ldots, p_n$ are rational numbers.

Step 3 : The continuity axiom of the entropy function ensures that the uniqueness theorem is valid when $p_1, p_2, p_3, \ldots, p_n$ are incommensurable.

4.8 MUTUAL INFORMATION

Expected Mutual Information :

Consider a set of messages sent $X = \{x_1, x_2, \ldots, x_m\}$ and the set of messages received $Y = \{y_1, y_2, y_3, \ldots, y_n\}$. Then the quantity

$$h(x_i, y_j) = \log \frac{p_{ij}}{p_{i0} p_{0j}} \qquad i = 1, 2, \ldots, m; j = 1, \ldots m, n$$

is known as mutual information of the messages sent x_i and the message received y_j.

Definition 4.15 : Expected Mutual Information : Expected mutual information of X and Y is denoted by $I(X,Y)$.

$$I(X,Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} \log \frac{p_{ij}}{p_{i0} p_{0j}}$$

Theorem 4.5 : $I(X,Y) = H(X) - H(X/Y) = H(Y) - H(Y/X)$

Proof : We have

$$H(X)-H(X/Y)=-\sum_{i=1}^{m}p_{i0}\log p_{j0}+\sum_{i=1}^{m}\sum_{j=1}^{n}p_{ij}\log p_{i/j}$$

$$=-\sum_{i=1}^{m}\left(\sum_{j=1}^{n}p_{ij}\right)\log p_{i0}+\sum_{i=1}^{m}\sum_{j=1}^{n}p_{ij}\log p_{i/j}$$

$$=\sum_{i=1}^{m}\sum_{j=1}^{n}p_{ij}\log\left(\frac{p_{i/j}}{p_{i0}}\right)$$

Since $p_{ij}=p_{i/j}p_{0j}$

$$p_{i/j}=\frac{p_{ij}}{p_{0j}}$$

Thus, $H(X)-H(X/Y)=\sum_{i=1}^{m}\sum_{j=1}^{n}p_{ij}\log\left(\frac{p_{ij}}{p_{i0}p_{0j}}\right)$

$$=I(X,Y)$$

Similarly, $H(Y)-H(Y/X)=I(X,Y)$

From theorem 4.5 we observe that the information conveyed about X by Y is same as the information conveyed about Y by X. When X and Y are independent, $I(X,Y)=0$.

Theorem 4.6 :

$$I(X,Y)=H(X)+H(Y)-H(X,Y)$$

Proof : We know that

$$H(X,Y)=H(X)+H(Y/X)$$

$$=H(Y)+H(X/\mathrm{Y})$$

$$\therefore H(X/Y)=H(X,Y)-H(\mathrm{Y})$$

$$\therefore I(X,Y)=H(X)-H(\mathrm{X/Y})$$

$$=H(X)-H(X,Y)+H(Y)$$

$$=H(X)+H(Y)-H(X,Y)$$

4.9 CHANNEL CAPACITY, EFFICIENCY AND REDUNDANCY

Definition 4.16 : Channel Capacity

$I(X,Y)$ indicates the measure of the average information per symbol transmitted in the system. The channel capacity is the maximum of information transmitted.

$$C = \max I(X,Y)$$

For noiseless chanel

$$I(X,Y) = H(X) = H(Y) = H(X,Y)$$

Therefore, $C = \max I(X,Y)$

$$= \max H(X)$$

$$= \max\left\{-\sum_{i=1}^{m} p_{i0} \log p_{i0}\right\}$$

Since $\max H(X) = H\left(\frac{1}{m}, \frac{1}{m}, \frac{1}{m}, ..., \frac{1}{m}\right)$

$$\therefore C = \max H(X) = H\left(\frac{1}{m}, \frac{1}{m}, ..., \frac{1}{m}\right)$$

$$= m\left(-\frac{1}{m}\log\left(\frac{1}{m}\right)\right)$$

$$= -\log\frac{1}{m} = \log m$$

Thus $C = \log m$ bits/symbol.

Definition 4.17 : Capacity of Channel

The capacity of channel can be expressed in bits per seconds. If symbols have a common duration of *t* sec, then channel capacity C per/sec is given by

$$C_t = \frac{C}{t} \text{ bits/sec} = \frac{\log n}{t} \text{ bits/sec}$$

Definition 4.18 : Redundancy

The difference between the actual rate of transmission of information $I(X, Y)$ and its maximum possible value is defined as the redundancy of the communication system.

Absolute redundancy for noise free channel = $C - I(X, Y)$

$= \log m - H(X)$

Relative Redundancy : The ratio of absolute redundancy to the channel capacity is defined as relative redundancy.

Relative redundancy for noise free channel $= \frac{\log m - H(X)}{\log m}$

$$= \frac{C - I(X,Y)}{C}$$

Definition 4.19 : Efficiency

The efficiency of the system is defined as the ratio of actual rate of transmission of information to its maximum possible value.

Efficiency of the noise free channel $= \frac{I(X,Y)}{C} = \frac{I(X,Y)}{\log m}$

Example : Find the capacity of the memoryless channel specified by

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 1 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

Solution : The capacity of memory less channel is given by

$$C = \max I(X,Y)$$

$$= \max\left[H(X) + H(Y) - H(X,Y)\right]$$

$$= -\sum_{ji}^{4} p(x_i, y_j) \log p(x_i, y_j)$$

$$= -\left\{\frac{1}{2}\log\frac{1}{2} + 2\left(\frac{1}{4}\log\frac{1}{4}\right) + 4\left(\frac{1}{4}\log\frac{1}{4}\right) + \log 1 + 2\left(\frac{1}{2}\log\frac{1}{2}\right)\right\}$$

$$= \frac{1}{2}\log 2 + \frac{1}{2}\log 4 + \log 4 + \log 2$$

$$= \left(\frac{1}{2} + 1 + 2 + 1\right)\log 2$$

$$= \frac{9}{2} \log 2 = \frac{9}{2} \text{ bits/symbol}$$

4.9 ENCODING

Definition 4.20 : Encoding may be defined as a transmission procedure of a message from sources to receiver through a noiseless channel in come code language.

In calculating the long run efficiency of communication system, the average length of a code word is of considerable interest. It is a quantity which is chosen to be minimum.

Following are the elements of the noiseless coding problem.

(i) A random variable X taking value $m_1, m_2, m_3, ..., m_n$ with prescribed probabilities $p\{m_1\}, p\{m_2\}, ..., p\{m_N\}$ respectively. X is to be observed independently over and over again. Thus, generating a sequence whose components belong to the set $\{m_1, m_4, ..., m_N\}$ such a sequence is called a message.

(ii) A $Y = \{a_1, a_2, a_3, ..., a_D\}$ set is called a set of code character or the code alphabet. Each symbol m_i is to be assigned a finite sequency of code characters called the code word associated with m_i. The collection of all code words is called a code. Code words are assumed to be distinct.

(iii) The objective of noiseless coding is to minimize the average code word length. Number of letters in a word is called length of the word.

Objective :

If the code word associated with m_i is of lenghth n_i, i = 1, 2, 3, ..., N, then the problem is to determine code that minimize the average length of messages.

The following are some of subclasses of code :

(i) Block Code : A code tha establishes a relationship with each of the symbols f the set X to a fixed sequence of symbols of the set Y is called a block code. e.g. m_1 may correspond to a_1a_7 or m_2 may correspond to $a_7a_8a_4$ etc.

(ii) Binary Code : In particular if the set Y = {0, 1} then the block code is called binary code $a_1 \to 1$, $a_2 \to 101$.

(iii) Non-singular code : A block code is said to be non-singular if all the words of the code are distinct.

(iv) Unique Decodable Code : A code is said to be unuquely decodable (separable) code if every finite sequence of symbols of the said Y is associated atmost one symbol of the set X.

e.g. $m_1 \to 0$, $m_2 \to 10$, $m_3 \to 110$, $m_4 \to 111$. Here encoding procedure established a one to one correspondance between message and their code words without the necessity of having any space between successive messages.

If we have 0000<u>10</u>0<u>10</u>0<u>110</u>1000111 the we have

$$m_1m_1m_1m_1m_2m_1m_2m_1m_3m_2m_1m_1m_4.$$

4.10 SHANNON FANO ENCODING PROCEDURE

This method of encoding is directed towards constructing reasonably efficient separable binary of codes for sources without memory.

Let $X = \{x_1, x_2, x_3, ..., x_n\}$ be the list of the messages that are to be transmitted from some source and $P = \{p_1, p_2, ..., p_n\}$ be their corresponding probabilities.

Aim is to device an encoding procedure so that a sequence of binary numbers {0, 1} of unspecified length can be associated to each message x_i. The sequence so obtained must satisfy the following conditions.

(i) No sequency of binary numbers can be obtained from any other sequence by adding additional binary terms to the sequences of shorter length.

(ii) Binary numbers associated with each messages x_i to form a sequence occur independently with equal probability.

The Procedure to Construct Code :

Step 1 : Arrange the messages $x_1, x_2, ..., x_n$ indescending order in terms of their probabilities without loss of generality let $p_1 > p_2 > p_3 .... > p_m$.

$\therefore$					
Message	x_1	x_2	x_3	....	x_m
Probability	p_1	p_2	p_3	....	p_m

Step 2 : Divide the set of messages X into two subsets X_1 and X_2 of equal probabilities.

Set	Message	Probabilities
X_1	x_1, x_2	$P(X_1) = p_1 + p_2$
X_2	$x_3, ..., x_m$	$P(X_2) = p_3 + p_4 + ... + p_m$

Such that $P(X_1) \cong P(X_2)$.

Step 3 : Again divide both subsets X_1 and X_2 into two subsets say X_{11}, X_{12} and X_{21}, X_{22} with equal probabilities respectively.

Step 4 : Assign binary number 0 to the first position of the coded word in each message in subset X_1 and 1 to the first position of the coded word in each message in subset X_2. The similar procedure of assignment is repeated for subsets of X_1 and X_2.

Setp 5 : The division and assignment will continue till each subset contains only one message (word).

Example : Apply Shannon Fano encoding procedure to the following set of messages.

X	:	m_1	m_2	m_3	m_4
P	:	0.4	0.3	0.2	0.1

Determine entropy of the source, expected length, efficiency and redundancy of the code that you obtain.

Answer :

Character	Probabilities	Partioning	Code word	Code word length
m_1	0.1 } X_1	X_1	0	1
m_2	0.3	X_{21}	10	2
m_3	0.2 } X_2	X_{22} → X_{221}	110	3
m_4	0.1	X_{22} → X_{222}	111	3

$X_1 = \{m_1\}$; $\qquad X_2 = \{m_2, m_3, m_4\}$

$X_{21} = \{m_2\}$; $\qquad X_{22} = \{m_3, m_4\}$

$X_{221} = \{m_3\}$; $\qquad X_{222} = \{m_4\}$

(a) The entropy of the source is

$$H(X) = -\sum p\{x_i\} \log p\{x_i\}$$

$$= -[0.4\log 0.4 + 0.3\log 0.3 + 0.2\log 0.2 + 0.1\log 0.1]$$

$$= 0.4\log\frac{10}{4} + 0.3\log\frac{10}{3} + 0.2\log\frac{10}{2} + 0.1\log 10$$

$$= 0.4\left[\log 2 + \log 5 - \log 2^2\right] + 0.3[\log 2 + \log 5 - \log 3]$$

$$+0.2[\log 5] + 0.1[\log 2 + \log 5]$$

$$= \log 5(0.4 + 0.3 + 0.2 + 0.1) - 0.3\log 3 + \log 2(0.4 - 0.8 + 0.3 + 0.1)$$

$$= \log 5 - 0.3\log 3$$

(b) The expected length of code is

$$L = \sum p\{m_i\} \cdot n_i$$

= (0.4)(1) + (0.3)(2) + (0.2)(3) + (0.1)(3)

= 0.4 + 0.6 + 0.6 + 0.3 = 1.9 bits/symbol

(c) Efficiency of code = $\dfrac{H(X)}{L}$

$$\eta = \frac{H(X)}{L} = \frac{\log 5 - 0.3\log 3}{1.9}$$

(d) Redundancy of code $\beta = 1 - \eta = \dfrac{1.9 - \log 5 + 0.3\log 3}{1.9}$

Example : The source memory has six characters with the following probabilities of transmission.

A	B	C	D	E	F
$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{12}$

Devise the Shannon Fano encoding procedure to obtain a uniquely decodable code to the above message. What is average length, efficiency and redundancy of the code that you obtain ?

Answer : Probabilities are already arranged in decending order.

$X_1 = \{A, B\}$

$X_2 = \{C, D, E, F\}$

$P(X_1) = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$

$P(X_2) = \frac{1}{8} + \frac{1}{8} + \frac{1}{12} + \frac{1}{12} = \frac{5}{12}$

$X_{11} = \{A\}$, $X_{12} = \{B\}$, $X_{21} = \{C, D\}$, $X_{22} = \{E, F\}$, $P(X_{21}) = \frac{1}{4}$,

$P(X_{22}) = \frac{1}{6}$, $X_{211} = \{C\}$, $X_{212} = \{D\}$, $X_{221} = \{E\}$, $X_{222} = \{F\}$.

	Length of code
Code of $A \in \{X_{11}\}$ is 00	2
Code of is $B \in \{X_{12}\}$ 01	2
Code of $C \in \{X_{211}\}$ is 100	3
Code of $D \in \{X_{212}\}$ is 101	3
Code of $E \in \{X_{221}\}$ is 110	3
Code of $F \in \{X_{222}\}$ is 111	3

(a) The entropy of source is given by

$$H(X) = -\sum p(x_i) \log p(x_i)$$

$$= -\left\{\frac{1}{3}\log\frac{1}{3} + \frac{1}{4}\log\frac{1}{4} + 2\left(\frac{1}{8}\log\frac{1}{8}\right) + 2\left(\frac{1}{12}\log\frac{1}{12}\right)\right\}$$

$$= \frac{1}{3}\log 3 + \frac{1}{4}\log 2^2 + \frac{1}{4}\log 2^3 + \frac{1}{6}(\log 3 + 2\log 2)$$

$$= \frac{1}{3}\log 3 + \frac{1}{2}\log 2 + \frac{3}{4}\log 2 + \frac{1}{6}\log 3 + \frac{1}{3}\log 2$$

$$= \log 2\left(\frac{1}{2} + \frac{3}{4} + \frac{1}{3}\right) + \log 3\left(\frac{1}{3} + \frac{1}{6}\right)$$

$$= \frac{19}{12}\log 2 + \frac{1}{2}\log 3$$

(b) The average code length of the message is given by,

$$L = \sum \ell_i p(x_i)$$

$$= 2\left(\frac{1}{3}\right) + 2\left(\frac{1}{4}\right) + 3\left(\frac{1}{8}\right) + 3\left(\frac{1}{8}\right) + 3\left(\frac{1}{12}\right) + 3\left(\frac{1}{12}\right)$$

$$= \frac{2}{3} + \frac{1}{2} + \frac{3}{8} + \frac{3}{8} + \frac{3}{12} + \frac{3}{12}$$

$$= \frac{16+12+9+9+6+6}{24} = \frac{58}{24} = \frac{29}{12}$$

= $\frac{29}{12}$ bits/symbol

Efficiency of code $= \frac{H(X)}{L} = \left[\frac{19}{12} + \frac{1}{2}\log 3\right]\frac{12}{29}$

Redundancy of code $= 1 - \left[\frac{19}{12} + \frac{1}{2}\log 3\right] \cdot \frac{12}{29}$

EXERCISE :

1. Write a critical essay on Information Theory.
2. Define entropy function.
3. Show that entropy function is maximum when mutually exclusive events are equiprobable.

4. An alphabet consists of 8 consonants and 8 vowels. Suppose that all the letter of the alphabet are equally probable and that there is no inter symbol influence. If consonants are always understood correctly but vowels are understood correctly only half the time being mistaken for other vowels the other half of the time, all vowels being involved in errors the same percentage of the time, what is the average rate of information transmission ?

5. Evaluate the entropy associated with the following probability distribution.

A	B	C	D
0.4	0.3	0.1	0.2

6. A transmitter and receiver has an alphabet that consists of three letters each. The joint probabilities for communication are given below.

$p(x_1, y_1)$	y_1	y_2	y_3
x_1	0.45	0.45	0.01
x_2	0.02	0.02	0.01
x_3	0.01	0.02	0.01

Determine the different entropies for this channel.

7. Apply Shannon Fano encoding procedure to the following messasge.

[X]	:	A	B	C	D
[P]	:	0.4	0.3	0.2	0.1

8. Apply Shannon-fano encoding procedure to the following message ensemble.

X	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
P	0.49	0.14	0.14	0.07	0.07	0.04	0.02	0.03

9. Find the capacity of the memoryless channel specified by the following channel matrix.

$$P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

❑ ❑ ❑ ❑