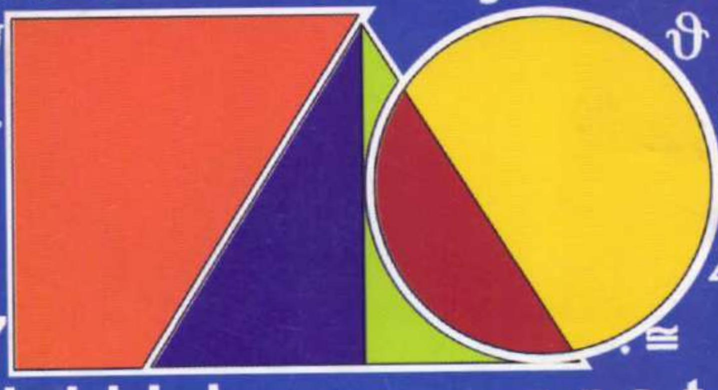
Prescribed by CHSE, Odisha



ELEMENTS OF MATHEMATICS

Class - XII

ODISHA STATE BUREAU OF TEXTBOOK PREPARATION AND PRODUCTION
PUSTAK BHAVAN, BHUBANESWAR

BUREAU'S

HIGHER SECONDARY

ELEMENTS OF MATHEMATICS

CLASS-XII

Prescribed by Council of Higher Secondary Education, Odisha, Bhubaneswar

Published by

ODISHA STATE BUREAU OF TEXT BOOK PREPARATION AND PRODUCTION

PUSTAK BHAVAN

A-11, SUKA VIHAR, BHUBANESWAR – 751 022

Prescribed by the Council of Higher Secondary Education, Odisha for the Higher Secondary Certificate Examination, 2018 and onwards.

ELEMENTS OF MATHEMATICS, Class-XII

[For Higher Secondary Class-XII Students]

Board of Writers :

Prof. Gokulananda Das, Ex-Vice Chancellor, Utkal University
Dr. Rabindra Nath Das, Ex-Reader, G.M. College, Sambalpur
Dr. Rajani Ballav Dash, Ex-Principal, S.C.S. College, Puri
Dr. Jagannath Prasad Debata, Ex-Principal, N.C. College, Jajpur
Dr. Akshaya Kumar Mishra, Director, Institute of Mathematics and Applications, Bhubaneswar
Dr. Maheswar Mishra, Ex-Reader, Ravenshaw College, Cuttack
Dr. Sashikanta Mishra, Ex-Reader, Ravenshaw College, Cuttack
Dr. Hemanta Kumar Mohanty, Ex-Reader, Ravenshaw College, Cuttack
Prof. Surendranath Mohapatra, Ex-Member, Service Selection Board, Bhubaneswar
Prof. Sribatsa Nanda, Ex-Professor, Regional Engineering College, Rourkela
Dr. Askhyaya Kumar Ojha, IIT, Bhubaneswar
Prof. Swadheenananda Pattnayak, Ex-Director Institute of Mathematics
Prof. (Miss) Dipti Rath, Professor, Berhampur University
Dr. Muralidhar Samal, Ex-Principal, N.C. College, Jajpur.
Dr. Prasanna Kumar Satapathy, Ex-Principal, S.C.S. College, Puri
Prof. Byasadev Pany, Ex-Principal, Govt. Auto. College, Phulbani
Dr. Hadibandhu Pattanayak, Ex-Principal, S.B. Women's College, Cuttack
Dr. Sanjay Kumar Mohanty, Ex-Principal, Gopabandhu Science College, Athagarh
Dr. K.K. Dash, Reader, Govt. College, Sundargarh
Dr. Minati Samal, Reader, RD(W) University, Bhubaneswar
Dr. Surendra Prasad Jena, Lecturer, S.B. (W) College, Cuttack
Dr. Manas Bhuyan, Lecturer, Rajdhani College, Bhubaneswar
Dr. Ranjan Ray, F.M. (Jr.) College, Balasore

Coordinator :

Prof. Gokulananda Das

Published by

The Odisha State Bureau of Textbook Preparation and Production
Pustak Bhavan, Bhubaneswar, Odisha, India
Latest Edition : 2017/ 40,000 *copies*
Publication No. : 203
ISBN - 978-81-8005-397-9

Type Setting & Design : M/s Jagannath Process Pvt. Ltd., Cuttack-2
Printing at : **M/s Jagannath Process Pvt. Ltd., Cuttack-2**

Price : Rs. 115/- (Rupees One hundred fifteen) only

FOREWORD

The Elements of Mathematics, Class-XII has been prepared as per the revised syllabus of Council of Higher Secondary Education, Odisha for the admission batch 2016.

The Bureau is extremely thankful to Prof. Gokulananda Das and his learned team of writers for extending tremendous effort in this regard.

The Bureau is confident that this book will be of immense help to the students and teachers of Mathematics. Constructive suggestions for further improvement of the book deserve appreciation.

Shri Umakanta Tripathy

Snana Purnima
09.06.2017

Director
Odisha State Bureau of Text Book
Preparation and Production,
Pustak Bhavan, Bhubaneswar

PREFACE

Today it is hardly necessary to talk about the importance of Mathematics in shaping all the successive human civilisations culminating in the present modern world, Whatever has been said or will be said about it are too little. Famous Indian Mathematician Bhaskaracharya, born in A.D. 1114 conceived Mathematics as one embodied in the following verse :

यथा शिखा मयुराणां
नागानाम् मणयौर्यथा ।
तद्दत् वेदाङ्ग शास्त्राणां
गणितं मूर्द्वनिस्थितम् ।।

- Vedanga Jyotish by Lagda (About 1100 B.C.)

(As crest in a peacock's feather, jewel in a Cobra's hood, Mathematics is the crest-jewel of all scientific knowledges.)

Mathematics, as a continuous human endeavour, seeks to capture the natural laws in the form of supreme abstract formulations and as such it has to depend upon infallible logic yielding the conclusions as eternal and absolute truth. It is a sublime discipline where falsehood or any inaccuracy is not entertained.

Since the study of Mathematics has become inescapable for the acquisition of any scientific knowledge, be it the farfetched subject like music or language, it is necessary to make the study of Mathematics more absorbing and interesting. The only way to do this is to encourage the students to pick up the pen and paper and start solving the problems themselves. Just as one learns swimming only after entering inside the water, one enjoys the taste of the sweets only after putting it inside the mouth, Mathematics is learnt only through problem solving and this is the shortest route. No amount of lecturing on 'swimming' can equip one to swim.

The authors of the book, working under diverse constraints, are not fully certain if they have lived upto the expectations and aspirations of the members of the Orissa Mathematical Society in particular and teachers, students and the public in general.

Any suggestions for the improvement of the book shall be gratefully acknowledged in bringing out the successive editions.

The authors are grateful to the authorities of the Council of Higher Secondary Education and the Text Book Bureau for the patience and care in bringing out the book in the present form.

Prof. G. DAS

CONTENTS

CHAPTER 1

Relation and Function

The rarest and most valuable of all intellectual traits is the capacity to doubt the obvious.

- Albert Einstein.

1.1 Introduction

We have studied earlier that a **relation R from a set A to a set B is a subset of the cartesian product A×B i.e. R⊆A×B**. If $a \in A$ is related to $b \in B$ by the relation R then we express this as aRb or equivalently as (a,b)∈R. We know about the *domain* and *range* of a relation and also about many-one, one-many, one-one relations and their diagramatic upresentations.

We mention that since $\phi \subseteq A\times B$, ϕ is a relation called **empty relation from A to B.** Here no element of A is related to any element of B. Similarly $A\times B \subseteq A\times B$ shows that A×B is a relation called the **universal relation** from A to B. Here every element of A is related to every element of B. These two relations are sometimes called trivial relations.

As we know earlier, relations occur abundantly in nature and in mathematics. More interesting and useful relations are those with suitable restrictions. Particularly some special type of relations which are defined on a set A, i.e. from A to itself, (in stead of one set to another), play much more significant role which we now proceed to study.

1.2 Types of Relations on A Set :

We begin with definitions.

Definitions : A relation R on a set A is called

(i) *a* **reflexive** relation if aRa, $\forall a \in A$

(i.e. if $(a,a) \in R$ for every $a \in A$)

(ii) *a* **symmetric** relation if $a\text{R}b \Rightarrow b\text{R}a$, $a,b \in A$.

(i.e. $(a,b) \in R \Rightarrow (b,a) \in R$ for $a,b \in A$)

(iii) *a* **transitive** relation if aRb and bRc $\Rightarrow a$Rc; $a,b,c \in A$.

(i.e. $(a,b) \in R$ and $(b,c) \in R \Rightarrow (a,c) \in R$; for a, b, $c \in A$)

Example 1

Let A={1,2,3}. Consider the following relations :

(i) R_1=(1,1), (2,2), (3,3)}

Since for each of the elements 1,2 and 3 of A ; (1,1), (2,2), (3,3) are in R_1, R_1 is reflexive. It is also symmetric and transitive. Particularly note that it satistics the condition of a transitive relation. For example $(1,1)\in R_1$ and $(1,1)\in R_1 \Rightarrow (1,1)\in R_1$ is trivially satisfied and similarly for other elements 2 and 3. It also trivially follows that R_1 is symmetric.

(ii) R_2=(1,1), (3,3)}

Here R_2 is not reflexive on A, since $(2,2)\notin R_2$.

Note that R_2 is symmetric and transitive.

(iii) R_3={(1,1), (2,2), (3,3), (1,2), (2,1), (1,3), (3,1), (2,3), (3,2)}

R_3 is reflexive, symmetric and trangitive.

Here $R_3 = A\times A$= The universal relation on A.

(iv) $R_4 = \{(a,b) : a - b>3\}$ on A.

No pair of elements a,b in A satisfy the condition that their difference is greater than 3. Hence no pair exists in R_4. It is an empty relation i.e. $R_4 = \phi$.

(v) R_5={(2,3), (3,2), (2,2), (3,3)}

In this case R_4 is not reflexive as $(1,1)\notin R_5$.

but if is symmetric and transitive.

Example 2

Let T = The set of all triangles on a plane. Define a relation ~ on T as follows :

A~B if and only if A is similar to B where A, B$\in$T.

From properties of triangles we can check that ~ is reflexive, symmetric and, transitive.

N.B. The symbol ~ is pronounced as 'wiggle' or 'tilde'.

Example 3

Let L = The set of all lines on a plane and R be a relation on L defined by

l_1Rl_2 if and only if l_1 is 'perpendicular to' l_2 where $l_1, l_2\in$L.

If is easy to check that $l_1Rl_2 \Rightarrow l_2Rl_1$.

So R is a symmetric selation. But it is neither reflexiue nor transitive.

The following relations on the set R of reals or its subsets are easy to check.

Example 4

(i) The relation 'is equal to' given by $\{(a,b) : a=b\}$

on R is reflexiue, symmetric and transitive.

(ii) The relation 'is less than' given by $\{(a,b) : a<b\}$ on R is not reflexive, not symmetric but transitive.

(iii) The relation 'is a factor of' or 'is a divisor of' or 'divides' given by $\{(a,b):a|b\}$ on Z is reflexive, transitive but not symmetric.

(iv) The relation $\{(x,y):y=2x\}$ on R is not reflexive, not symmetric and not transtive.

Of all the relations on a set we give special attention to those which are reflexive, symmetric and transitive. So we bring them under a definition separately. The symbol '~' is usually used to denote such a relation.

Definition :

Equivalence Relation

A relation $R \subseteq A \times A$ is called an equivalence relation on A if it is (i) reflexive, (ii) symmetric and (iii) transitive.

(for an equivalence relation the symbol ~ is very often used instead of R.)

'Equality' on any set, 'congruency' and 'similarity', on the set of triangles, 'parallelism or coincidence' on the set of lines, 'has the same age as' on a set of people are some of the examples of equivalence relation. We will now discuss an important equivalence relation on Z.

1.3 Congruence Modulo Relation on Integers

Definition :

Let $a, b \in Z$ and let m be a fixed positive integer. We say that ***a* is congruent to *b* modulo m** and write this as :

$a \equiv b \pmod{m}$ *iff* m divides $a-b$.

Note that in this case a is of the form $a = b + mk$, for some $k \in Z$.

For example, $10 \equiv 1 \pmod{3}$

$-20 \equiv 1 \pmod 7$

$-9 \equiv 3 \pmod 4$

Instead of writing $20 \equiv -1$ (mod7) we often write this as $20 \equiv 6 \pmod 7$

The following properties directly follow from definition.

Let $a, b, c, d \in Z$. and m be a fixed positive integen.

Then (i) If $a \equiv b$ (mod m) and x is any integer then $a + x \equiv b + x$ (mod m)

(ii) If $a \equiv b$ (mod m) and $c \equiv d$ (mod m) then

$a + c \equiv b + d$ (mod m), $ac \equiv bd$ (mod m), The proof is left to the reader.

For example (i) $127 \equiv 119 \pmod 4 \Leftrightarrow 8 \equiv 0 \pmod 4$

(ii) $10 \equiv 2$ (mod4) and $8 \equiv 12 \pmod 4$

$\Rightarrow 18 \equiv 14 \pmod 4$ and $80 \equiv 24 \pmod 4$

Example 5

Show that

(i) $ca \equiv cb \ (mod\ m) \Rightarrow a \equiv b \ (mod\ m')$ where $m = m' \times gcd\ (c,m)$

(ii) $ca \equiv cb \ (mod\ m) \Rightarrow a \equiv b \ (mod\ m)$ if $gcd\ (c,m) = 1$,

Solution :

(i) Let $gcd\ (c,m) = h$.

Then let $c = c'h$ and $m = m'h$ and $gcd\ (c',m') = 1$

Now $ca \equiv cb\ (mod\ m) \Rightarrow m \mid c(a-b)$

$\Rightarrow m'h \mid c'h(a-b)$

$\Rightarrow m' \mid (a-b)\ (\because gcd\ (c',m') = 1)$

$\therefore a \equiv b$ (mod m')

(ii) Taking $h = 1$ in (i) we get (ii).

Thus (i) $8 \equiv 12 \pmod 4 \Rightarrow 4 \equiv 6 \pmod 2$ ($\because$ gcd (2,4)=2)

and $-8 \equiv 10 \pmod 3 \Rightarrow -4 \equiv 5 \pmod 3$ ($\because$ gcd (2,3)=1)

The following example illustrates the fact that congruence modulo relation on Z is an equivalence relation.

Example 6

Show that the relation $\sim$ on Z given by $\sim = \{(a,b): a \equiv b\ (mod\ 3)$ (i.e. $a \sim b$ iff 3 divides $a - b$), is an equivalence relation.

Solution :

$\sim$ is reflexive :

3 divides $a - a$ for all $a \in Z \Rightarrow a \equiv a \pmod 3 \Rightarrow a \sim a$ for all $a \in Z$

$\sim$ is symmetric :

$a \sim b \Rightarrow 3|(a-b) \Rightarrow a-b = 3k$ for some $k \in Z$

$\Rightarrow b-a = 3(-k), -k \in Z$

$\Rightarrow 3|(b-a) \Rightarrow b \equiv a\ (mod\ 3) \Rightarrow b \sim a.$

$\sim$ is transitive :

$a \sim b$ and $b \sim c$; $a, b, c \in Z$

$\Rightarrow 3|(a-b)$ and $3|(b-c)$

$\Rightarrow 3|[(a-b)+(b-c)]$

$\Rightarrow 3|(a - c) \Rightarrow a \sim c.$

Hence $\sim$ is an equivalence relation.

Note : There is nothing special about the integen 3 in the above example. We can simply replace 3 by any positive integer in to assert that congruence modulo relation on Z is an equivalence relation.

The most significant feature of equivalence relation on a set X is that it enables the set to be partitioned or divided into a disjoint collection of subsets of X whose union is X.

Consider, for example,

S = The set of all students of a school imparting education from class I upto class XII. Let a relation R be defined on S as follows :

For $s_1, s_2 \in S$; $s_1 R s_2$ if s_1 and s_2 are in the same class.

You can easily check that R is an equivalence relation. Now suppose $s_1 \in S$ is in class I. Let the set of all students in S who are related to s_1 i.e. are in the same calss as s_1, be denoted by $[s_1]$.

Thus $[s_1] = \{s \in S : sRs_1\}$

Clearly $s_1 \in [s_1]$ .... (i)

and $[s_1]$ represents the students of class I.

Further if s_1' is another student of class I i.e. if $s_1' Rs_1$ then $[s_1'] = [s_1]$ .... (ii)

which you can verify

Now choose another student $s_2 \in S$ not related to s_1.

If s_2 is in class II, as before we get

$[s_2]$ = The set of all students of class II.

More over $[s_1] \cap [s_2] = \phi$ .... (iii)

For, if $s \in [s_1] \cap [s_2]$ then sRs_1 and sRs_2 together imply that s_1Rs and sRs_2 by symmetry of R, which again yields s_1Rs_2 by transitivity. But s_1Rs_2 is a contradiction.

Thus by choosing a representative student s_i , i=1, 2, 3, ..., 12 such that s_i is not related to s_j $(i \neq j)$ we partition the students of the school into mutually disjoint different subsets called class I, class II, ..., class XII. In symbols,

$S = [s_1] \cup [s_2] \cup .... \cup [s_{12}]$. ... ... (iv)

We formalise this discussion below.

1.4 Equivalence class

Definition :

Let X be a set with an equivalence relation ~ defined on it. The **equivalence class of $x \in X$, denoted by $[x]_\sim$ is defined by**

$[x]_\sim = \{y \in X : y \sim x\}$.

We simply write $[x]$ in place of $[x]_\sim$ when there is no confusion with regard to the equivalence relation with respect to (w.r.f) which the equivalence class is considered.

As we have observed in (i) to (iv) above, the following facts are easily deducible from definition of equivalence class.

(d) For a set X with an equivalence relation on it,

(i) $x \in [x]$ $(\because x \sim x$ by reflexive property)

(ii) $x \sim y \Leftrightarrow [x]=[y]$

(iii) Either $[x] = [y]$ or $[x] \cap [y] = \phi$

(iv) $\underset{x \in X}{\cup} [x] = X$.

Definition : A **Partition** of a set X is a collection of disjoint, non empty subsets of X that have X as their union.

For example if $X = \{1,2,3\}$, then $\{\{1,2\},\{3\}\}$ and $\{\{1\},\{2\},\{3\}\}$ are two partitions of X.

Hence we state that given any equivalence relation on a set X we can partition the set X into subsets X_i called equivalence classes, such that.

(i) All elements of X_i are related to each other

(ii) Any element of X_i is NOT related to any element of X_j for $i \neq j$.

(iii) $X_i \cap X_j = \phi$ for $i \neq j$.

(iv) $\underset{i}{\cup} X_i = X$.

Example 7

Congruence modulo 3 relation partitions (or decomposes) the set Z into three disjoint equivalence classes.

Find them.

Solution :

Any integer $a \in Z$ leaves remainder 0 or 1 or 2 when divided by 3. Hence either $a \equiv 0$ (*mod* 3) or $a \equiv 1$ (*mod* 3) or $a \equiv 2$ (*mod* 3). In other words $a = 3m$ or $a = 3m+1$ or $a = 3m+2$ for $m \in Z$. First we find all $a \in Z$ such that $a \equiv 0$.

These are given by $\{0, \pm 3, \pm 6, \pm 9, ...\} = \{..-6, -3, 0, 3, 6,...\}$

It is the equivalence class of 0 denoted by $[0]_3$. So

$[o]_3=\{..., -6, -3, 0, 3, 6, ...\}$

Similarity $[1]_3=\{..., -5, -2, 1, 4, 7, ...\}$

$[2]_3=\{..., -4, -1, 2, 5, 8, ...\}$

Hence $Z = [0]_3 \cup [1]_3 \cup [2]_3$.

Remark : For a given positive integer m, congruence modulo m relation partitions the set Z into in equivalence classes denoted by $[0]_m$, $[1]_m$, ..., $[m-1]_m$.

Equivalence relation induced by a portion :

Converse to the above discussion, given any partition of a set we can obtain an equivalence relation on it. Take for example X={1,2,3,4,5,6}.

Let $X_1 = \{1,2,5\}$, $X_2=\{3,6\}$, $X_3=\{4\}$ be a partition of X. Define a relation R on X as follows :

*a*R*b* iff a and b 'are in the same subset'.

Now (i) aRa for all $a \in X \Rightarrow$ R is reflexive.

(ii) If aRb then bRa follows $\Rightarrow$ R is symmetric

(iii) If aRb and bRc then a and b are is the same subset and along with the fact that b and c are also in the same subset if follows that *a* and c are in the same subset.

Hence *a*Rc, implying that R is transitive.

Thus R is an equivalence relation. More over R is given explicitly by

R={(1,1),(1,2),(1,5),(2,1),(2,2),(2,5),(5,1),(5,2),(5,5),(3,3),(3,6),(6,3),(6,6), (4,4)}

and the equivalence classes are

$[1]=[2]=[3]=X_1$, $[3]=[6]=X_2$, $[4]=X_3$.

Exercise-1(a)

1. If A = {*a*,b,c,d} mention the type of relations on A given below, which of them are equivalence relations ?

   (i) {(*a*,*a*), (b,b)}

   (ii) {(*a*,*a*), (b,b), (c,c), (d,d)}

   (iii) {(*a*,b), (b,*a*), (b,d), (d,b)}

   (iv) {(b,c), (b,d), (c,d)}

   (v) {(*a*,*a*), (b,b), (c,c), (d,d), (*a*,d), (*a*,c), (d,*a*), (c,*a*), (c,d), (d,c)}

2. Write the following relations in tabular form and determine their type.

   (i) R={(*x*,*y*) : 2*x*-*y*=0] on A = {1,2,3,..., 13}

(ii) R= $\{x,y) : x$ divides y} on A = {1,2,3,4,5,6}

(iii) R = $\{(x,y) : x$ divides $2 - y\}$ on A = {1,2,3,4,5}

(iv) R = $\{(x,y) : y \leq x \leq 4\}$ on A = {1,2,3,4,5}.

3. Test whether the following relations are reflexive, symmetric or transitive on the sets specified.

(i) R = {(m,n):m-n $\geq$ 7} on Z.

(ii) R = {(m,n):2|(m+n)} on Z.

(iii) R = {(m,n) : m+n is not divisible by 3} Z.

(iv) R= {(m,n): $\frac{m}{n}$ is a power of 5} on Z-{0}.

(v) R = {(m,n):mn is divisible by 2} on Z.

(vi) R = {(m,n) : 3 divides m-n}on {1,2,3...,10}.

4. List the members of the equivalence relation defined by the following partitions on X={1,2,3,4}. Also find the equivalence classes of 1,2,3 and 4.

(i) $\{\{1\},\{2\},\{3,4\}\}$

(ii) $\{\{1,2,3\},\{4\}\}$

(iii) $\{\{1,2,3,4\}\}$

5. Show that if R is an equivalence relation on X then *dom*R = *rng*R=X.

6. Give an example of a relation which is

(i) reflexive, symmetric but not transitive.

(ii) reflexive, transitive but not symmetric.

(iii) symmetric, transitive but not reflexive.

(iv) reflexive tbut neither symmetric nor transitive.

(v) transitive but neither reflexive nor symmetric.

(vi) an empty relation.

(vii) a universal relation.

7. Let R be a relation on X, If R is symmetric then $x\text{R}y \Rightarrow y\text{R}x$. If it is also transitive then $x\text{R}y$ and $y\text{R}x \Rightarrow x\text{R}x$. So whenever a relation is symmetric and transitive then if is also reflexive. What is wrong in this argument ?

8. Suppose a box contains a set of n balls ($n \geq 4$) (denoted by B) of four different colours (may have different sizes), viz. red, blue, green and yellow. Show that a relation R defined on B as $R=\{(b_1,b_2) :$ balls b_1 and b_2 have the same colour$\}$ is an equivalence relation on B. How many equivalence classes can you find with respect to R ?

[Note : On any set X a relation $R=\{(x,y)$: x and y satisfy the same property P$\}$ is an equivalence relation. As far as the property P is concerned, elements x and y are deemed equivalent. For different P we get different equivalence relations on X]

9. Find the number of equivalence relations on $X=\{1,2,3\}$. [Hind : Each partition of a set gives an equivalence relation.]

10. Let R be the relation on the set R of real numbers such that aRb iff a-b is an integer. Test whether R is an equivalence relation. If so find the equivalence class of 1 and $\frac{1}{2}$ w.r.t. this equivalence relation.

11. Find the least positive integer r such that

(i) $185 \in [r]_7$, (ii) $-375 \in [r]_{11}$, (iii) $-12 \in [r]_{13}$.

12. Find least non negative integer r such that

(i) $7 \times 13 \times 23 \times 413 \equiv r \pmod{11}$

(ii) $6 \times 18 \times 27 \times (-225) \equiv r \pmod{8}$

(iii) $1237 \pmod{4} + 985 \pmod{4} \equiv r \pmod{4}$

(iv) $1936 \times 8789 \equiv r \pmod{4}$

13. Find least positive integer x satisfying

$276x + 128 \equiv 4 \pmod{7}$

[Hint : $276 \equiv 3$, $128 \equiv 2 \pmod{7}$]

14. Find three positive integers x_i, $i = 1, 2, 3$ satisfying $3x \equiv 2 \pmod{7}$

[Hint : If x_1 is a solution then any member of $[x_1]$ is also a solution]

1.5 Functions : Types of Functions

As we already know, a function from a set X to a set Y is a special type of relation f from X to Y such that for each $x \in X$ these is one and only one element $y \in Y$ which is related to x by the relation f. We write $y=f(x)$ and call it the image of x under f and call x, the pre-image of y under f.

In symbols,

$f: X \rightarrow Y$ if $f \leq X \times Y$ such that $\forall x \in X$, $\exists!\, y \in y$ and $(x, y) \in f$. (The symbol $\exists!$ means that there exists an element uniquely, i.e. there exists only one element) We also know that $dom f = D_f = X$, *codomain* of $f = Y$ and $rng\, f = R_f = f(X)$, We studied examples of different functions on the set R of real number and about their graphs. We continue our study of functions in some greater detail as the concept of function is extremely important in mathematics. We begin with the following definitions.

Definition: A function $f: X \rightarrow Y$ is said to be an **onto** or **surjective function** if $rng\, f = f(X) = Y$. i.e. if every element of Y is the image of some element of X.

Definition : A function $f: X \rightarrow Y$ is said to the an **into function** if $rng\, f = f(X) \subsetneq Y$ (a proper subset of Y)

i.e. if there is at least one $y \in Y$ which has no pre-image in X.

Definition: A function $f: X \rightarrow Y$ is said to be a **one-one** or **injective function** if for every $x_1, x_2 \in x$ $f:(x_1) = f(x_2) \Rightarrow x_1 = x_2$.

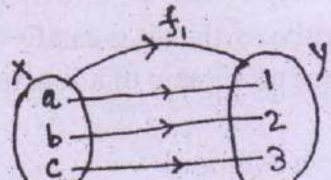
(Equivalently f is one-one if distinct elements. in X have distinct images in Y, i.e. $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$.

Definition: A function $f: X \rightarrow Y$ is said to be a many-one function if there exist $x_1, x_2 \in X$ with $x_1 \neq x_2$ such that $f(x_1) = f(x_2)$. i.e. more than one (many) elements of X have the same image under f.

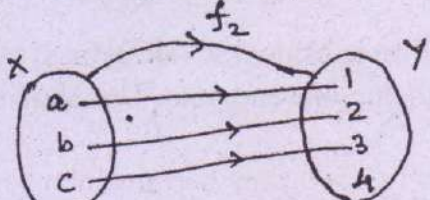
We illustrate these ideas in the following example.

Example 8

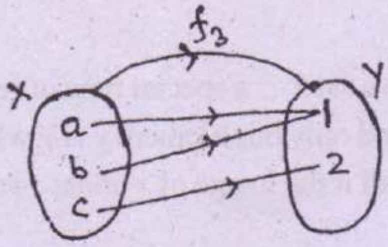
Consider the following functions represented diagramatically.



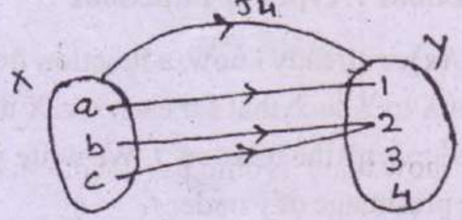
(i) One-one onto



(ii) One-one into



(iii) many-one onto



(iv) many-one into

In diagram (i) X={a,b,c}, Y={1,2,3} and f_1: X→Y such that $f_1(a)=1$, $f_1(b)=2$ and $f_1(c)=3$. Here distinct elements of x have distinct images in Y. Hence f_1 is a one-one map. More over $f_1(X)=rng\, f_1=\{1,2,3\}=Y$ implies that f_1 is onto. We say: f_1 is a **one-one onto function**.

In diagram (ii) f_2 is one-one. Moreover there is one element $4 \in$ Y which is not the image of any of the elements of X. Thus $rng\, f_2 = \{1,2,3\} \subsetneq Y = \{1,2,3,4]$. Hence f_2 is a **one-one into function.**

In diagram (iii) there are two distinct elements viz a and b, which have the same image 1, i.e. $a \neq$ b but $f_3(a) = f_3(b) = 1$. Hence f_3 is a many-one function. Further $f_3(X)=rng\, f_3=\{1,2\}=Y$ shows that f_3 is an onto function. Thus f_3 is a **many-one onto function**.

In diagram (iv) f_4 is a **many-one into function** as you can check.

Note that unlike a one-many relation, we do not have a one-many function. (why !)

Definition. A function f : X→Y is said to be a **bijective** (one-one onto) function which is both injective and surjective.

Definition : Two sets X and Y are said to be in one-one correspondence if there is a bijective function from X to Y. If there is a one-one correspondence between two sets X and Y then we say that these sets one equivalent/similar/equipollent or equipotent and we write X~Y.

Now we take a closer look at lijective functions which play a very significant role.

Example- 9

Consider f :N→N defined by $f(x)=2x$. $f(x_1)=f(x_2) \Rightarrow 2x_1=2x_2 \Rightarrow x_1=x_2$. Hence f is one-one. But f is not onto since I∈N has no preimage. for, if there is an $x \in$N s.t. $f(x)=1$ then we should get $2x=1 \Rightarrow x=\frac{1}{2} \notin$N.

But if we consider f:R→R defined by $f(x) = 2x$ as above, then it is easily verified that f is both one-one and onto. This shows the importance of domain of a function.

Example 10

Let f : R→R be defined by

$$f(x) = \begin{cases} x^2 - \dfrac{1}{x^2}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Show that f is onto but not one-one.

Solution :

In order to show that f is onto, given $y \in R$ we need to find $x \in R$ s.t. $y = f(x)$.

If $y = 0$, then we have $x = 0$ (by definition of f)

If $y \neq 0$, then $y = f(x) = x^2 - \dfrac{1}{x^2} = \dfrac{x^4 - 1}{x^2}$

$\Rightarrow x^4 - x^2y - 1 = 0$

$\Rightarrow x^2 = \dfrac{1}{2}\left(y + \sqrt{y^2 + 4}\right)$ (neglecting - sign before the radical as $x^2 > 0$.

$\Rightarrow x = \pm\dfrac{1}{\sqrt{2}}\sqrt{y + \sqrt{y^2 + 4}} \in R.$

So f is onto.

Clearly f is not one - one since $f(1) = f(-1) = f(0) = 0$.

Example 11

Let X and Y be **finite** sets and | X | = | Y |.

(Recall that |S| denotes the number of elements in a finite set S)

Show that if f: X→Y is onto then f must be one one and conversely if f is one-one then it must be onto.

Solution :

Let |X| = |Y| = m.

Let f be onto.

If f is not one-one suppose $x_1, x_2 \in X, x_1 \neq x_2$ and $f(x_1) = f(x_2)$. Since each element of x has only one image in Y there are atmost m-2 images of the elements $x_3, x_4 ... x_m$ of X in Y. So altogether we have (m-2)+1=m-1 images in Y. Hence $|rng\, f| \leq m-1 < m = |y|$ which contradicts the fact that f is onto. Hence f most be one-one.

Conversely let f be one-one since each of the m elements of X has a unique image in Y and f is one-one there must be m images in Y. i.e. $|rng\, f| = m = |Y|$

$\Rightarrow f$ is onto.

Note: This result is not true for infinite sets as we have already seen in Example-9 and Example-10.

Example 12

(i) The function $f: R \to R$ defined by $f(x) = x^2$ is not one-one as $f(-1)=f(1)=1$. It is not onto because any negative real number has no pre-image.

But $f: R \to R_+ = \{x \in R : x > 0\}$ is onto but not one-one However $f: R_+ \to R_+$ is bijective.

(ii) The function $f: R \to R$ defined by $f(x) = x^3$ is bijective. It is surjective : If $y \in R$ then we can find $y^{\frac{1}{3}} \in R$ such that $f\left(y^{\frac{1}{3}}\right) = y$. So *rng* $f = R$.

It is injective :

$$f(x_1) = f(x_2) \Rightarrow x_1^3 = x_2^3$$
$$\Rightarrow (x_1 - x_2)\left(x_1^2 + x_1x_2 + x_2^2\right) = 0$$
$$\Rightarrow (x_1 - x_2)\left(x_1^2 + x_1x_2 + x_2^2\right) = 0$$
$$\Rightarrow x_1 - x_2.$$

$\left(\because x_1^2 + x_1x_2 + x_2^2 \neq 0 \text{ for any } x_1, x_2 \in R \text{ both not zero}\right).$

(iii) The function $f: R \to R$ defined by $f(x) = 3x+5$ is bijective.

It is surjective : If $y \in R$ then we can find $\frac{y-5}{3} \in R$

Such that $f\left(\frac{y-5}{3}\right) = 3\left(\frac{y-5}{3}\right) + 5 = y$.

It is injective : $f(x_1) = f(x_2)$

$\Rightarrow 3x_1 + 5 = 3x_2 + 5 \Rightarrow x_1 = x_2 \,;\, x_1, x_2 \in R$.

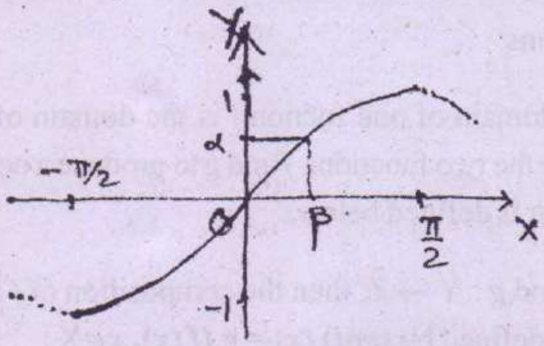
(iv) The function $f: R \to R$ defined by $f(x) = |x|$ is neither injective nor surjective since

(i) $-2 \neq 2$ and $|-2| = 2 = |2|$

and (ii) $rng\, f = R_+ \cup \{0\} \subsetneq R$.

(v) The function $f: R-\{0\} \to R-\{0\}$ defined by $f(x) = \frac{1}{x}$ is one-one and onto.

(vi) The function $f: \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1, 1]$ defined by $f(x) = \sin x$ is one-one and onto. It is clear from the graph of $\sin x$. otherwise,



graph of sin x

$\sin x_1 = \sin x_2 \Rightarrow 2\cos\frac{x_1+x_2}{2}\, \sin\frac{x_1-x_2}{2} = 0$

$\Rightarrow \sin\frac{x_1-x_2}{2} = 0, \Rightarrow x_1 - x_2 = 0 \Rightarrow x_1 = x_2.$

$\left(\text{If } \cos\frac{x_1+x_2}{2} = 0,\ x_1, x_2 = \pi \text{ and } x_1 + x_2 \in \left[\frac{-\pi}{2}, \frac{\pi}{2}\right] \Rightarrow x_1 = x_2 = \frac{\pi}{2}\right)$

Again if $\alpha \in [-1,1]$ we can find $\beta\ (=\sin^{-1}\alpha) \in \left[\frac{-\pi}{2}, \frac{\pi}{2}\right]$

(see figure) such that $\sin\beta = \alpha$. You will learn about inverse trigonometric functions later.

But f: R- [-1,1] defined by $f(x) = \sin x$ is onto but not one-one. You know from periodicity of sine that $\sin x = \sin(x+2\pi)$ for $x \in$ R. Thus by restricting *domain* of sin e function from from R to $\left[\frac{-\pi}{2}, \frac{\pi}{2}\right]$ we get a bijective function from one which is not bijective.

Similar observation can be had relating to other trigonometric functions. See article on trigonometric functions for detailed discussion.

(vii) The function f: R→R defined by $f(x) = \frac{x}{x^2+1}$ is neither one-one nor onto.

It is not onto; for if $\frac{x}{x^2+1} = 1 \in$ R then $x^2 - x + 1 = 0$, whose roots are not real. Hence 1 has no pre-image. Also f is not one-one since $f\left(\frac{1}{2}\right) = \frac{2}{5} = f(2)$.

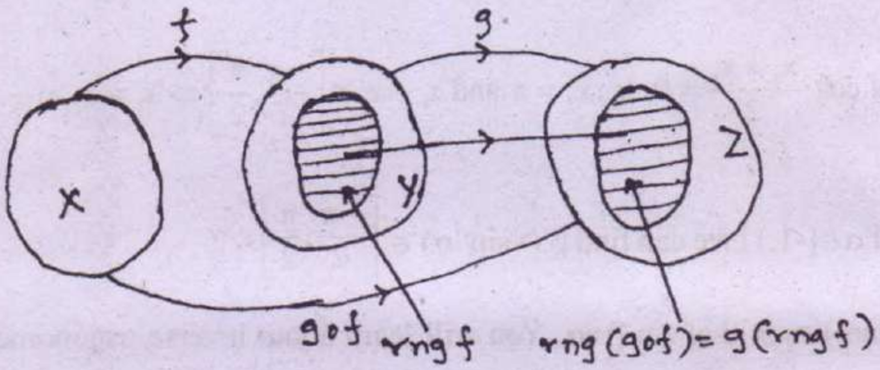
1.6 Composition of Functions

It is possible that the codomain of one fuction f is the domain of another function g. There is *a* way to combine the two functions f and g to produce a new function called the composition of f and g that is defined below.

Definition. If $f: X \to Y$ and $g: Y \to Z$, then the composition of f and g *denoted by gof* (read 'g composite f') is defined by $(gof)(x) = g(f(x), x \in X$.

Note the order in which f and g appear in the definition of composition of f and g. Also note that *gof* is defined only when $rng\, f \subseteq Y = dom\, g$. clearly $gof: X \to Z$ so that

$dom\,(gof) = X$ and $rng\,(gof) = g(rng\,f) \subseteq Z$. The idea of composition is illustrated in the adjoining figure.



Example 13.

Let X = {1,2,3}, Y={a,b,c,d} and Z={red, green, blue}. Let $f: X \to Y$ s.t. f = {(1,a), (2,c), (3,d)} and $g: Y \to Z$ s.t. g = {(a,green), (b, green), (c, blue), (d, red)} *gof* is obtained by observing the diagram.

$$\begin{array}{ccccc} & f & & g & \\ 1 & \longrightarrow & a & \longrightarrow & \text{green} \\ 2 & \longrightarrow & c & \longrightarrow & \text{blue} \\ 3 & \longrightarrow & d & \longrightarrow & \text{red} \end{array}$$

So $gof: X \to Z$ s.t. gof = {(1, green), (2, blue), (3, red)}

Dispensing with diagramatical approach, we have $(gof)(1) = g(f(1) = g(a)$ = green and similarly $(gof)(2)$=blue and $(gof)(3)$ = red, which defines $gof: X \to Z$.

Example 14

Consider the following real functions.

(i) $f_1(x) = \sin x^2$

(ii) $f_2(x) = \cos(\sin x)$

(iii) $f_3(x) = (2x+1)^2 + 3(2x+1)+2$

Hence f_1, f_2 and f_3 are some examples of functions obtained by composition of f and g in the following ways (Assumming f: R→R and g : R→R):

(i) If $g(x) = \sin x$ and $f(x)= x^2$, then $(gof)(x)=g((f(x))$

$= g(y)$ (writing $y = f(x)$

$= \sin y = \sin x^2 = f_1(x)$.

(ii) $g(x) = \cos x$ and $f(x) = \sin x \Rightarrow (gof)(x)=g((f(x))$

$= g(y) = \cos y = \cos(\sin x)$

(writing $y = f(x) = f_2(x)$.

(iii) $g(x) = x^2+3x+2$ and $f(x) = 2x+1$

$\Rightarrow (gof)(x)$

$= g((f(x)) = g(y) = y^2+3y+2 = (2x+1)^2+3(2x+1)+2 = f_3(x)$

It is very important to note that **composition of functions is not commutative** in general. If *gof* is defined then fog may not be defined and even if *fog* is defined, it may so happen that $gof \neq fog$ as the following example shows.

Example 15

(i) Let f: R→R be defined by $f(x) = x+1$ and g:R_+→R be defined by $g(x) = \sqrt{x}$.

In this case $(fog)(x)=(f(g(x)) = f(\sqrt{x}) = \sqrt{x}+1$.

But *gof* is not defined as $rngf \not\subset dom\ g$.

(ii) Let us take f:R_+→R and g:R_+→R_+ in (i) above. Then both *fog* and gof are defined

But $(fog)(x) = f(g(x) = f(\sqrt{x}) = \sqrt{x}+1$ and

$(gof)(x) = g(f(x) = g(x+1) = \sqrt{x+1}$. clearly $fog \neq gof$.

(iii) Let f:R$-\{0\}\to$R-$\{0\}$ and g:R$-\{0\}\to$R-$\{0\}$ be defined by

$$f(x) = x^2 \text{ and } g(x) = \frac{1}{x}.$$

Then both *gof* and *fog* are defined. Further

$$(gof)(x) = g(f(x) = g(x^2) = \frac{1}{x^2} \text{ and}$$

$$(fog)(x) = f(g(x) = f\left(\frac{1}{x}\right) = \frac{1}{x^2}$$

Here $fog = gof$ on R $-\{0\}$.

Some important results on composition of functions :

Theorem 1.

If f:X = Y, g:Y$\to$Z and h:Z $\to$ V then

ho(gof) = (hog) of i.e. composition is assocaitive.

Proof.

By definition of composition,

gof : X$\to$Z and hog:Y$\to$V. Hence

ho(gof): X $\to$ V and (hog)of : X $\to$ V. Moreover for $x \in$ X,

$$(\text{ho(gof)})(x) = h(gof)(x) = h(g(f(x))$$

and $((\text{hog})\text{of})(x) = (hog)(f(x)) = h(g(f(x)).$

Hence proved.

Theorem 2.

If f : X $\to$ Y and g : Y $\to$ Z then

(i) *gof* is onto if both f and g are onto.

(ii) *gof* is one-one if both f and g are one-one.

(iii) If *gof* is onto then g is onto.

(iv) If *gof* is one-one then f is one-one.

Proof.

(i) Let f and g be onto. Then $f(X) = Y$ and $g(Y)=Z$.

Hence $(gof)(X)=g(f(X))=g(Y)=Z$.

$\Rightarrow rng(gof)=Z \Rightarrow gof$ is onto.

(ii) Let f and g both be one-one.

Suppose $(gof)(x_1) = (gof)(x_2)$; $x_1, x_2 \in X$.

$\Rightarrow g(f(x_1)) = g(f(x_2))$

$\Rightarrow f(x_1) = f(x_2) \quad (\because g$ is one-one)

$\Rightarrow x_1 = x_2 \quad (\because f$ is one-one)

Hence gof is one-one.

(iii) Let $gof : X \to Z$ be onto.

Hence if $z \in Z$, then there is at least one $x \in X$ s.t. $(gof)(x) = z$,

i.e. $g(f(x)) = z$. Writing $y = f(x)$ we have $y \in Y$ s.t. $g(y)= z$.

Hence g is onto.

(iv) Let gof be one-one.

Suppose f is not one-one.

Then for some $x_1, x_2 \in X$ with $x_1 \neq x_2$

$f(x_1) = f(x_2)$

$\Rightarrow g(f(x_1)) = g(f(x_2))$

$\Rightarrow (gof)(x_1) = (gof)(x_2)$

This is contradiction to the fact that gof is one-one.

So gof is one-one $\Rightarrow f$ is one-one.

1.7 Inverse of a Function

We begin with functions f_1, f_2, f_3 and f_4 as given in Example 8. We have

$f_1 : \{a,b,c\} \rightarrow \{1,2,3\}$ and $f_1 = \{(a,1), (b,2), (c,3)\}$

$f_2 : \{a,b,c\} \rightarrow \{1,2,3,4\}$ and $f_2 = \{(a,1), (b,2), (c,3)\}$

$f_3 : \{a,b,c\} \rightarrow \{1,2\}$ and $f_3 = \{(a,1), (b,1), (c,2)\}$

$f_4 : \{a,b,c\} \rightarrow \{1,2,3,4\}$ and $f_4 = \{(a,1), (b,2), (c,2)\}$

For the function f_1 consider the inverse relation $g_1 = \{(1,a), (2,b), (3,c)\}$. Does it represent a function ? Here $g_1 = \{1,2,3\} \rightarrow (a,b,c\}$ such that $g_1(1)=a$, $g_1(2)=b$ and $g_1(3)=c$ and consequently g_1 is a function which we call to be the **inverse function** of f_1 as the *domain* and *condomain* of f_1 are interchanged with these of g_1.

On the otherhand, for the function f_2 , consider the inverse relation $g_2=(1,a), (2,b), (3,c)$. We see that g_2 is not a function from $\{1,2,3,4\}$ to $\{a,b,c\}$ since $g_2(4)$ is not defined. In this case inverse function of f_2 does not exist.

Similarly in case of f_3, the inverse function of f_3 cannot be defined by $g_3=\{(1,a),(1,b),(2,c)\}$ as $I \in dom\ g_3$ would have two different images. Also f_4 has no inverse function as the elements 3 and 4 of $dom g_4$ would be without images in the codomain of g_4.

Hence we observe that out of f_1, f_2, f_3 and f_4 only f_1, which is one-one onto, has the inverce functions g_1. We usually write f_1^{-1} in place of g_1.

This observation is true in general that all functions which are both one-one and onto (bijective) have inverse functions. We prove it below.

Theorem 3.

Let $f: X \rightarrow Y$. Then the relation f^{-1} is a function from Y to X if f is bijective.

Proof.

We show that

(i) $dom f^{-1} = Y$ and

(ii) every $y \in Y$ has a unique image in X under f^{-1}.

(i) By hypothesis, f is surjective so that for every $y \in Y$ there is $x \in X$ s.t. $(x,y) \in f$.

$\Rightarrow (y, x) \in f^{-1} \Rightarrow y \in dom\, f^{-1}$.

$\therefore y \subseteq dom f^{-1}$ and since $dom f^{-1} \subseteq Y$, we have $dom f^{-1} = Y$.

(ii) Now suppose $y \in Y$ has two distinct images x_1 and x_2 under f^{-1}.

So $(y, x_1) \in f^{-1}$ and $(y, x_2) \in f^{-1}$.

$\Rightarrow (x_1, y) \in f$ and $(x_2, y) \in f$.

Since f is injective, $x_1 = x_2$ a contradiction to assumption.

Therefore $f^{-1}: Y \to X$ is a function.

Now we have the following definition.

Definition.

A bijective function f: $X \to Y$ has an **inverse function** f^{-1}: $Y \to X$ given by

$f^{-1} = \{(y,x) : (x,y) \in f\}$

Thus if f is bijective, $y = f(x) \Leftrightarrow x = f^{-1}(y)$

A function. f is said to be **invertible** if f^{-1} exists. The next theorem helps us to test invertibility of f.

Theorem 4.

If f: $X \to Y$ is bijective then

(i) $f^{-1} o f = id_X$ and $f o f^{-1} = id_Y$ where id_X and id_Y are the identity functions on X and Y respectively.

(ii) If in addition g: $y \to Z$ is bijective then $(gof)^{-1} = f^{-1} o g^{-1}$.

Proof.

(i) To show $f^{-1} o f = id_X$, by definition of equality of funcitons, we must show that

$\left(f^{-1} of\right)(x) = id_X(x)$ for every $x \in X$.

Now $y = \left(f^{-1} of\right)(x) = f^{-1}\left(f(x)\right)$

$\Rightarrow \left(f(x), y\right) \in f^{-1}$

$$\Rightarrow (y, f(x)) \in f \Rightarrow f(y) = f(x) \Rightarrow y = x$$

($\therefore f$ is one-one)

$$\Rightarrow y = \left(f^{-1}of\right)(x) = x = id_x(x),\ x \in X.$$

$$\Rightarrow f^{-1}of = id_X.$$

Similarly we can prove that $fof^{-1} = id_Y$.

(ii) By hypothesis, f^{-1}: Y→X, and g^{-1}: Z→Y. By Theorem 2, *gof* is bijective and $(gof)^{-1}$: Z→X.

In order to show that $(gof)^{-1} = f^{-1}og^{-1}$ we must show, by definition of equality of functions, that for every z∈Z, $(gof)^{-1}(z) = (f^{-1}og^{-1})(z)$.

Now, let z∈Z.

Then we have x∈X such that $(gof)^{-1}(z) = x$ ... ... (A)

$$\Rightarrow z = (gof)(x) = g(f(x)) = g(y)$$

$$(\text{writing } y = f(x))$$

$$\Rightarrow y = g^{-1}(z).$$

But $y = f(x) \Leftrightarrow x = f^{-1}(y) = f^{-1}\left(g^{-1}(z)\right) = \left(f^{-1}og^{-1}\right)(z)$. ... (B)

From (A) and (B) if follows that

$$(gof)^{-1}(z) = \left(f^{-1}og^{-1}\right)(z) \text{ for any } z \in Z.$$

Hence $(gof)^{-1} = f^{-1}og^{-1}$.

Remark.

The converse of Theorem 4(i) is true, i.e. if we can find g: Y→X such that $gof = id_X$ and $fog = id_Y$ then f is bijective and $g = f^{-1}$. (see Q.14, exercise-1(b))

We use this fact to test invertibility of a bijective function.

Example-16

Let f: R→R be defined by $f(x) = 3x+5$

Show that f is bijective. Find $f^{-1}(1)$ and $f^{-1}(0)$.

Solution :

To show that f is bijective refer Example 12 (iii).

Now $y = f(x) = 3x+5 \Rightarrow x = \dfrac{y-5}{3}$.

So for any $y \in R$ we can find a unique $\dfrac{y-5}{3} \in R$

Such that $f\left(\dfrac{y-5}{3}\right) = 3.\dfrac{y-5}{3} + 5 = y$. Hence f^{-1} is given by $f^{-1}(x) = \dfrac{x-5}{3}, x \in R$.

Then $f^{-1}(1) = -\dfrac{4}{3}$ and $f^{-1}(0) = -\dfrac{5}{3}$.

Example- 17

Show that f:R→R defined by $f(x) = x^2-1$ is not invertible in general. Find the *domain* and *codomain* where f is invertible. Also find f^{-1}.

Solution

For $f(x)=x^2-1, f(-1)=0=f(1)$. So f is not one-one. Hence f is not bijective and therefore not invertible.

Now, let $y = x^2-1 \in R \Rightarrow x = \pm\sqrt{y+1}$. For x to be real we must have $y \geq -1$ ... (A)

Moreever f is many-one since both $-\sqrt{y+1}$ and $+\sqrt{y+1}$ map onto y under f.

To make f one-one we restrict the *domain* to $R_+ \cup \{0\}$ i.e. $\{x \in R: x \geq 0\}$ ... (B)
Using (A) and (B) we consider f: $[0,\infty) \rightarrow [0,\infty)$.

Now let us define g:$[-1,\infty) \rightarrow [0,\infty)$ by $g(y)=\sqrt{y+1}$.

Then $(gof)(x) = g(f(x)=g(x^2-1)=\sqrt{(x^2-1)+1} = x = id_x(x)$

and $(fog)(y) = f(g(y)) = f\left(\sqrt{y+1}\right) = \left(\sqrt{y+1}\right)^2 - 1 = y = id_Y(y)$

since $gof = id_x$ and $fog=idy$, $g=f^{-1}$.

$\therefore f^{-1}$ is defind by $f^{-1}(x) = \sqrt{x+1}$, $x \in R$.

Exercises 1 (b)

1. Let X=$\{x,y\}$ and Y=$\{u,v\}$. Write down all the functions that can be defined from X to Y. How many of these are (i) one-one (ii) onto and (ii) one-one and onto ?

2. Let X and Y be sets containing m and n elements respectively.

   (i) What is the total member of functions from X to Y.

   (ii) How many functions from X to Y are one-one according as m<n, m>n and m=n ?

3. Examine each of the following functions if it is

   (i) injective (ii) surjective, (iii) bijective and (iv) none of the three

   (a) $f : R \to R, f(x)=x^2$

   (b) $f : R \to [-1,1], f(x)= \boldsymbol{sin}\ x$

   (c) $f : R_+ \to R+, f(x)=x+\frac{1}{x}$ where $R_+ = \{x \in R : x > 0\}$

   (d) $f : R \to R, f(x) = x^3+1$

   (e) $f: (-1,1) \to R, f(x) = \frac{x}{1-x^2}$

   (f) $f: R \to R, f(x) = [x]$ = the greatest integer $\le x$.

   (g) $f:R \to R, f(x) = |x|$

   (h) $f:R \to R, f(x) = \boldsymbol{sgn}\ x$

   (i) $f:R \to R, f = id_R =$ the identity function on R.

4. Show that the following functions are injective.

   (i) $f(x) = \sin x$ on $\left[0, \frac{\pi}{2}\right]$

   (ii) $f(x) = \cos x\ [0,\pi]$

   (iii) $f(x) = \log_a x$ on $(0,\infty)$, $(a > 0$ and $a \neq 1)$

   (iv) $f(x) = a^x$ on R. $(a > 0$ and $a \neq 1)$

5. Show that functions f and g defined by $f(x)=2 \log x$ and $g(x) = \log x^2$ are not equal even though $\log x^2 = 2 \log x$.

6. Give an example of a function which is

(i) Surjective but not injective.

(ii) injective but not surjective.

(iii) neither injective nor surjective.

(iv) bijective

7. Prove that the following sets are equivalent :

$\{1, 2, 3, 4, 5, 6,...\}$

$\{2, 4, 6, 8, 10,...\}$

$\{1, 7, 5, 7, 9,...\}$

$\{1, 4, 9, 16, 25,...\}$

8. Let $f = \{(1,a), (2,b), (3,c), (4,d)\}$ and

$g = \{(a, x), (b, x), (c, y), (d, x)\}$

Determine *gof* and *fog* if possible. Test whether *fog=gof*.

9. Let $f = \{(1,3), (2,4), (3,7)\}$

and $g = \{(3,2), (4,3), (7,1)\}$

Determine *gof* and *fog* if possible. Test whether *fog* = *gof*.

10. Let $f(x) = \sqrt{x}$ and $g(x) = 1-x^2$.

(i) Find natural domains of *f* and *g*.

(ii) Compute *fog* and *gof* and find their natural domains.

(iii) Find natural domain of $h(x) = 1-x$.

(iv) Show that h= *gof* only on $R_0 = \{x \in R : x \geq 0\}$ and not on R.

11. Find the composition *fog* and *gof* and test whether *fog* = *gof* when *f* and *g* are functions on R given by the following :

(i) $f(x) = x^{3+1}$, $g(x) = x^2 - 2$

(ii) $f(x) = \sin x$, $g(x) = x^5$

(iii) $f(x) = \cos x$, $g(x) = \sin x^2$

(iv) $f(x) = g(x) = \left(1-x^3\right)^{\frac{1}{3}}$

12. (a) Let f be a real function. Show that

$h(x) = f(x) + f(-x)$ is always an even function

and $g(x) = f(x) - f(-x)$ is always an odd function.

(b) Express each of the following function

as the sum of an even function and an odd function:

(i) $1+x+x^2$, (ii) x^2, (iii) e^x, (iv) $e^x + \sin x$

13. Let X = {1, 2, 3, 4} Determine whether f: X→X

defined as given below have inverses. Find f^{-1} if it exists :

(i) f = {(1,4), (2,3), (3,2), (4,1)}

(ii) f = {(1,3), (2,1), (3,1), (4,2)}

(iii) f = {(1,2), (2,3), (3,4), (4,1)}

(iv) f = {(1,1), (2,2), (2,3), (4,4)}

(v) f = {(1,2), (2,2), (3,2), (4,2)}

14. Let f: X→Y.

If there exists a map g:Y→X such that $gof = id_X$ and $fog = id_Y$, then show that

(i) f is bijective and (ii) $g = f^{-1}$.

[Hint- since id_X is a bijective function, $gof = id_X$ is bijective. By Theorem 2(iv) f is injective. Similarly fog is bijective ⇒ f is surjective by Theorem 2(iii)]

15. Construct an example to show that $f(A \cap B) \neq f(A) \cap f(B)$ where $A \cap B \neq \phi$

16. Prove that for any f: X→Y, $f \circ id_X = f = id_Y \circ f$.

17*. Prove that f: X→Y is surjective iff for all $B \subseteq Y$, $f(f^{-1}(B)) = B$.

18*. Prove that f: X→Y is injective iff $f^{-1}(f(A)) = A$ for all $A \subseteq X$.

19*. Prove that f: X→Y is injective iff for all subsets A, B of X, $f(A \cap B) = f(A) \cap f(B)$.

20*. Prove that f: X→Y is surjective iff for all $A \subseteq X$, $(f(A))' \subseteq f(A')$, where A' denotes the complement of A in X.

1.8 Binary Operations

Addition, subtraction, multiplication and division of real members are some of the arithmetic operations with which we are all familiar. A closer look at these operations makes it clear that we take two real numbers and apply these operations to obtain another real number. We can think of this operation as applying a function on R×R to R. Since these operations involve basically 'two' elements, they are called binary operations, we note that the operations of addition and multiplication in R denoted by symbals '+' and '×' respectively, satisfy a property given by a+b = b+a and a×b=b×a. In the following we discuss about binary operations in general and study some of the properties they satisfy. We begin with

Definition.

Let A be a nonempty set. A binary operation on A is a function from A×A to A.

In otherwords, given any ordered pair (a,b) of elements a and b of A the **binary operation** '∗' associates to this ordered pair a unique element c of A. We write a∗b=c instead of writing $*\left((a,b)\right)$=c.

Example 18

Addition, subtration and multiplication are binary operations on R because given any two real members a and b, a+b, a-b and a×b (or ab) are also real numbers.

Subtraction is not a binary operation or N because for $a<$ b, $a-$ b$\notin$N.

Division is not a binary operation on R becomes a ÷ b is not defined when b = 0.

But it is a binary operation on R-{0}.

Multiplicaiton is a binary operation on R as a×b ∈R for any a,b∈R. But multiplication of a real number by a constant is not a binary operation because it (i.e. $x \to kx$, k is constant, $x \in$ R) is not a function or R×R (why ?).

Definition :

Let B ⊂ A and ∗ be a binary operation in A. If for every pair of elements (x,y)∈B, x ∗ y is also in B, we say that B is **closed** under '∗'.

If there is at least one pair (x,y) in B such that $x * y$ is not in B then we say that B is not closed under '∗'. Thus as shown in Example-18, N is not closed under subtraction but closed under addition. We also say that B satisfies **closure law** w.r.t. '∗ ' if B is closed under '∗'.

Example 19.

Let X be any set and P be the set of all subsets of X.

Let U:P×P→P be given by (A,B) →A ∪ B and ∩ : P×P→P be given by (A,B) →A ∩ B.

Because union (intersection) of a pair of sets (A,B) is a unique set denoted by $A\cup B$ ($A\cap B$) in P, $\cup$ and $\cap$ are binary operations in P.

(Note that binary operation ' in P ' and 'on P ' shall be used synonimously)

Example 20

$(x, y) \rightarrow x^y$ is a binary operation or N because for every pair of positive integers (x,y) there is associated a unique integer x^y. But this is not a binary operation on Z or Q or R (We cannot always associate (0,y) to a member).

If the order of a set is small we can represent a binary operation on it by a table called composition table/operation table/multiplication table, like the multiplication table which is known to us from our school days.

(Note that the coinage 'multiplication table' does not necessarily mean that the binary operation involved is our well known multiplication operation. In fact, the symbols used for binary operations have flexible meanings specific to their definitions in the context. Just see the table that follows.)

Consider the following example.

Let $A=\{\alpha,\beta,\gamma\}$. define '+' on A as follows. $\alpha+\alpha=\alpha$, $\alpha+\beta=\beta$, $\alpha+\gamma=\gamma$, $\beta+\alpha=\beta$, $\beta+\beta=\gamma$, $\beta+\gamma=\alpha$, $\gamma+\alpha=\gamma$, $\gamma+\beta=\alpha$ and $\gamma+\gamma=\beta$. We can write the above binary operation '+' by the following table.

+	α	β	γ
α	α	β	γ
β	β	γ	α
γ	γ	α	β

We shall now consider some important type of binary operations.

Definition A binary operation '$*$' on a set A is said to be

(i) **commutative** if $a * b = b * a$ for all $a, b \in A$.

(ii) **associative** if $(a*b)*c = a*(b*c)$ for all $a,b,c\in A$.

For example, '+' is commutative but '-' is not commutative in R and '÷' is not associative in R-{0}.

Definition.

Let $*$ be a binary operation on a set A.

(i) The set A is said to have an **identity element** if there exists a unique element $e\in A$ such that $e*a = a*e$ for all $a\in A$. In that case e is called the identity element w.r.t. $*$.

(ii) If the set A has an identity element e w.r.t a binary operation, then an element $a\in A$ is said to have an **inverse element** w.r.t. $*$, denoted by a^{-1}, if $a*a^{-1}=a^{-1}*a=e$.

In that case a^{-1} is called the inverse of a w.r.t. $*$.

(Also note that the meaning of a^{-1} is specific to the definition given here. It does not necessarily mean the exponential a^{-1})

Example 21

The integer $0\in Z$ is the identity element of Z w.r.t. '+' since $a+0=0+a=a$ for every $a\in Z$.

For any integer $n\in Z$, -m is the inverse of m w.r.t '+'.

The integer $1\in Z$ is the identity element of Z w.r.t 'x' since $1\times a=a\times 1=a$ for all $a\in Z$. But any integer $m\in Z$, $m\neq 1$, does not have an inverse w.r.t. '×'.

Again $1\in Q$ (or R) is the identity element of Q (or R) w.r.t. ×, the multiplication operation and for any $x\in Q$ (or R), $x\neq 0$, $\frac{1}{x}$ is the inverse of x w.r.t.'×'.

These concepts have far reaching consequences in higher mathematics.

Example 22

Show that '+' and '×' defined on R×R by the rules

(i) $(x_1,y_1)+(x_2,y_2)=(x_1+x_2, y_1+y_2)$ and

(ii) $(x_1,y_1)\times(x_2,y_2)=(x_1x_2-y_1y_2, x_1y_2+x_2y_1)$

are both commutative and assaciative

where $(a,b)=(c,d)\Leftrightarrow a=c$ and $b=d$.

Solution :

(i) $(x_1,y_1)+(x_2,y_2)=(x_1+x_2, y_1+y_2)$

$=(x_2+x_1, y_2+y_1)$ ($\because$ addition on R is commutative)

$=(x_2,y_2)+(x_1,y_1)$ (by defination of '+' on R×R)

Hence '+' is commutative.

Since addition of real numbers obey associative property we can similarly show that '+' satisfies this property on R×R.

(ii) We only test the commutative property for '×'.

$(x_1,y_1)\times(x_2,y_2)=(x_1x_2-y_1y_2, x_1y_2+x_2y_1)$ (by definition of '×') and

$(x_2,y_2)\times(x_1,y_1)=(x_2x_1-y_2y_1, x_2y_1+y_2x_1)$

since $x_1x_2-y_1y_2=x_2x_1-y_2y_1$

and $x_1y_2+x_2y_1=x_2y_1+y_2x_1$ (by commutative property of multiplication in R), by defination of equality in R×R, we see that $(x_1,y_1)\times(x_2,y_2)=(x_2,y_2)\times(x_1,y_1)$.

Testing of associativity is left as exercise.

Exercises- 1 (c)

1. Show that the operation $*$ given by $x*y=x+y-xy$ is a binary operation on Z, Q and R but not on N.

2. Determine whether the following operations as defined by $*$ are binary operations on the sets specified in each case. Give reasons if it is not a binary operation.

   (i) $a*b=2a+3b$ on Z.

   (ii) $a*b=ma-nb$ on Q+ where m and $n \in N$.

   (iii) $a*b=a+b \ (mod\ 7)$ on $\{0,1,2,3,4,5,6\}$

   (iv) $a*b=\min\{a,b\}$ on N.

   (v) $a*b=GCD\{a,b\}$ on N.

   (vi) $a*b=LCM\{a,b\}$ on N.

   (vii) $a*b=LCM\{a,b\}$ on $\{0,1,2,3,4......,10\}$

   (viii) $a*b=\sqrt{a^2+b^2}$ on Q_+

   (ix) $a*b=a\times b \ (mod\ 5)$ on $\{0,1,2,3,4\}$.

   (x) $a*b=a^2+b^2$ on N.

   (xi) $a*b=a+b-ab$ on $R-\{1\}$.

3. In case $*$ is a binary operation in Q2 above, test whether it is (i) associative (ii) commutative, Test further if the iedentity element exists and the inverse element for any element of the respective set exists.

4. Constract the composition table /multiplication table for the binary operation $*$ defined on $\{0,1,2,3,4\}$ by $a*b=a\times b \ (mod\ 5)$. Find the identity element if any. Also find the inverse elements of 2 and 4.

   [This operation is called multiplication moduls 5 and denoted by $\times_5$. In general, on a finite subset of N, $\times_m$ denotes the operation of multiplication modulo m where m is a fixed positive integer].

●

Inverse Trigonometric Functions

Mathematics possesses not only truth, but supreme beauty, cold and austere, like that of sculpture and capable of stern perfection such as only the greatest art can show.

- Bertrand Russell

2.0 Introduction

In earlier classes you have studied the concepts of trigonometric functions, such as sine, cosine, tangent, cotangent, secant and cosecant; their domain, range and some properties. Also you know that trigonometric functions are not one-one and onto e.g. $\text{Sin}\frac{\pi}{6}=\frac{1}{2}=\text{Sin}\frac{5\pi}{6}$. Hence in general we cannot define their inverse functions. But by suitable restriction of their domains we can make bijective functions out of them. In this chapter we shall consider the inverse trigonometric functions. These functions have extensive use in mathematics especially in integral calculus and also in engineering and technology.

2.1 Definitions :

Let us first consider the function

$\text{Sin} : R \to [-1, 1]$

Let $y = \text{Sin}\, x, x \in R$. Look at the graph of $\sin x$. For $y \in [-1, 1]$, there is a unique number x in each of the intervals..., $\left[-\frac{3\pi}{2}, -\frac{\pi}{2}\right], \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \left[\frac{\pi}{2}, \frac{3\pi}{2}\right], \cdots$ such that $y = \sin x$. Hence any one of these intervals can be chosen to make sine function bijective. We usually choose $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ as the *domain* of sine function. Thus $\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1, 1]$ is bijective and hence admits of an inverse function with range $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ denoted by $\mathbf{sin^{-1}}$ or **arcsin** (see foot note).

Each of the above mentioned intervals as range gives rise to different **branches** of $\sin^{-1}$ function. The **function $\sin^{-1}$ with the range $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ is called the principal branch** which is defined below.

$\text{Sin}^{-1} : [-1, 1] \to \left[-\frac{\pi}{2}, -\frac{\pi}{2}\right]$ defined by $y = \sin^{-1}x \Leftrightarrow x = \sin y$

The values of $y(=\sin^{-1}x)$ in $\left[-\frac{\pi}{2}, -\frac{\pi}{2}\right]$ are called **principal values** of $\sin^{-1}$.

---

* The prefix 'arc' in 'arcsin x' stands for 'the trigonometric argument whose sine is x'. In a unit circle it is same as an arc of measure of θ radians.

Similar considerations for other trigonometric functions give rise to respecive inverse functions. We define below the **principal branches** of $\cos^{-1}$, $\tan^{-1}$ and $\cot^{-1}$.

$\cos^{-1} : [-1, 1] \to [0, \pi]$ defined by $y=\cos^{-1}x \Leftrightarrow x = \cos y$

$\tan^{-1} : R\to (-\frac{\pi}{2},\frac{\pi}{2})$ defined by $y=\tan^{-1}x \Leftrightarrow x = \tan y$

$\cot^{-1} : R\to (0, \pi)$ defined by $y=\cot^{-1}x \Leftrightarrow x = \cot y$

Observe that in the above definition for $\tan^{-1}$, since $x = \tan y$ is undefined for $y = \pm\frac{\pi}{2}$, $\left\{-\frac{\pi}{2}, \frac{\pi}{2}\right\}$ is excluded from the range of $\tan^{-1}$. For similar reason $\{0, \pi\}$ is excluded from the range of $\cot^{-1}$.

To define $\sec^{-1}$, we see that $y = \sec^{-1}x \Leftrightarrow x = \sec y$. Since $|\sec y| \geq 1$ i.e. $\sec y \leq -1$ or $\sec y \geq 1$, the domain of $\sec^{-1}$ is all of R excluding the numbers in the interval $(-1, 1)$. Further $x = \sec y = \frac{1}{\cos y}$ shows that no real value of x exists when $y = \frac{\pi}{2}$. So $y \neq \frac{\pi}{2}$ and hence $\left\{\frac{\pi}{2}\right\}$ is to be excluded from the range of $\sec^{-1}$. Similar considerations lead us to exclude $(-1, 1)$ from the domain and $\{0\}$ from the range of cosec^{-1}. Thus we define $\sec^{-1}$ and cosec^{-1} as follows :

$\sec^{-1} : R - (-1, 1) \to [0, \pi] - \left\{\frac{\pi}{2}\right\} = [0,\frac{\pi}{2})\cup(\frac{\pi}{2},\pi]$ defined by $y = \sec^{-1}x \Leftrightarrow x = \sec y$.

$\mathrm{cosec}^{-1} : R - [-1, 1]\to\left[-\frac{\pi}{2},\frac{\pi}{2}\right] - \{0\}=[-\frac{\pi}{2},0)\cup(0,\frac{\pi}{2}]$, defined by $y = \mathrm{cosec}^{-1}x \Leftrightarrow x = \mathrm{cosec}\, y$.

Like $\sin^{-1}$, the inverse trigonometric funtions $\cos^{-1}$, $\tan^{-1}$, $\cot^{-1}$, $\sec^{-1}$ and cosec^{-1} are also denoted by arccos, arctan, arccot, arcsec and arccosec respectively.

Remark :

1. The notation $\sin^{-1}x$ should not be confused with $\frac{1}{\sin x}$, the reciprocal of $\sin x$, which we denote by $(\sin x)^{-1}$. Remember that $\sin^{-1}$ is a single or non-composite notation for the inverse function of sine rather than the (-1)th power of some number.

2. Unless otherwise mentioned, by $\sin^{-1}x$ we shall always mean the principal value of $\sin^{-1}x$ and similarly for other inverse trigonometric functions.

2.2 Graphs :

To obtain the graph of $y=f^{-1}(x)$ we can take sample points $x_1, x_2, x_3.........x_n$ on x-axis and evaluate the function f^{-1} at these points giving $y_1, y_2, \ldots\ldots\ldots\, y_n$ and then join the points (x_1, y_1), $(x_2, y_2), \ldots(x_n, y_n)$ by smooth hand.

But the following general technique can be adopted for obtaining the graph of the inverse function f^{-1} when the graph of f is already known.

Let us first consider the $\sin^{-1}$ function. Since $y = \sin^{-1}x \Leftrightarrow x = \sin y$, the graph of $y=\sin^{-1}x$ can be obtained from that of $x=\sin y$ by interchanging x- and y- axes. Look at the graph of $y=\sin x$ given in Fig. (*a*). The bold part of the graph is to be taken into consideration for the principal branch of $\sin^{-1}x$, where $x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

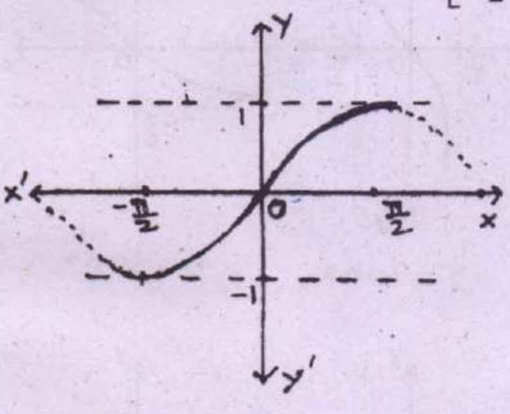


Fig.(a) y = sin*x*

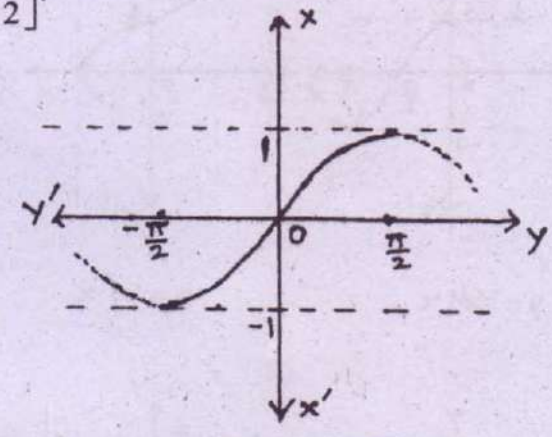


Fig. (b) *x* = siny

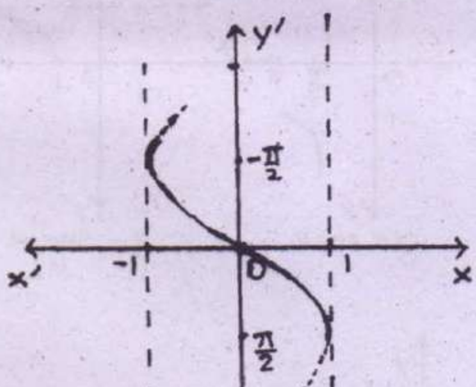


Fig. (c)

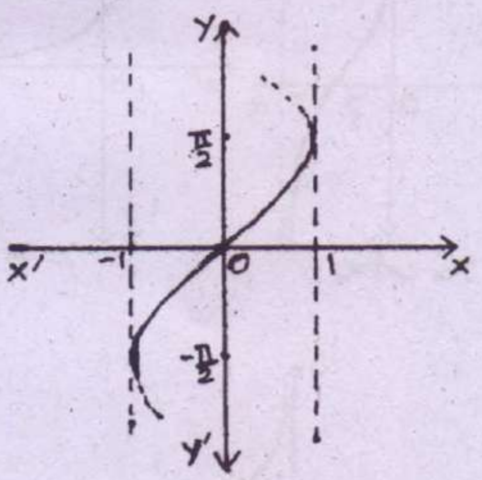


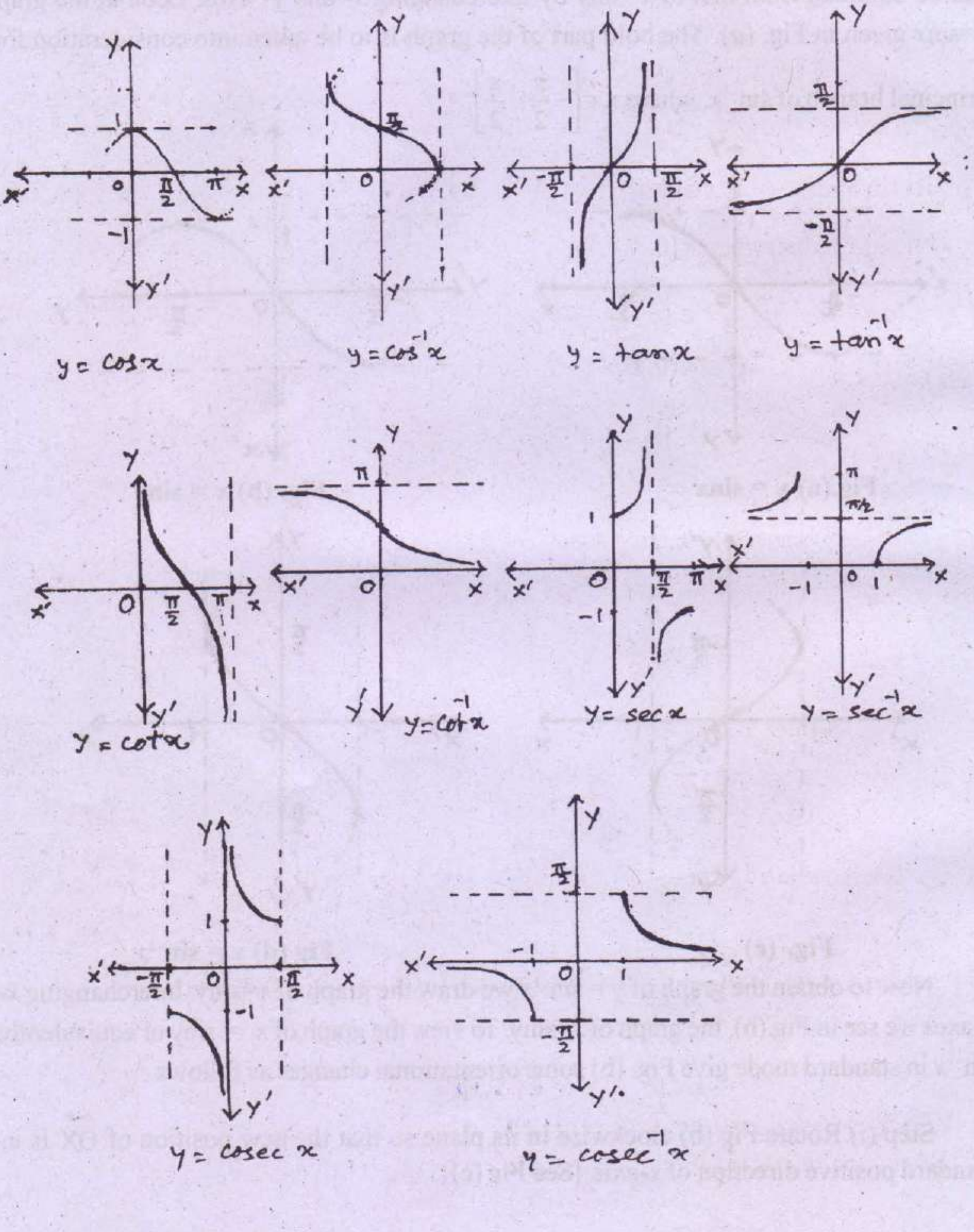
Fig (d) y = sin^{-1}*x*

Now to obtain the graph of $y = \sin^{-1}x$ we draw the graph of $x=\sin y$. Interchanging x-and y-axes we see in Fig.(b), the graph of $x=\sin y$. To view the graph of $x = \sin y$ or equivalently $y = \sin^{-1}x$ in standard mode give Fig. (b) some orientational changes as follows :

Step (i) Rotate Fig (b) clockwise in its plane so that the new position of $\overrightarrow{OX}$ is in the standard positive direction of x-axis {See Fig (c)}.

Step (ii) Rotate Fig. (c) about $\overrightarrow{X'X}$ so that $\overrightarrow{OY}$ is now in the standard positive direction of y=axis {See Fig. (d)}. Fig. (d) gives the graph of $y=\sin^{-1}x$.

The above procedure can be followed to obtain graphs of $\cos^{-1}x$, $\tan^{-1}x$, $\cot^{-1}x$, $\sec^{-1}x$ and $\operatorname{cosec}^{-1}x$. The principal branches of these inverse functions are shown below. Along with the graphs of inverse functions the graphs of corresponding trigonometric functions are also given for understanding the above procedure.



2.3 Important Properties

Property-I.

We know that when $f: X \to Y$ is invertible then $fof^{-1}=I_Y$ and $f^{-1}of=I_X$.

Applying this we have.

(i) $\sin(\sin^{-1}x) = x,\ x \in [-1,1]$

$\cos(\cos^{-1}x) = x,\ x \in [-1,1]$

$\tan(\tan^{-1}x) = x,\ x \in \mathbb{R}$

$\cot(\cot^{-1}x) = x,\ x \in \mathbb{R}$

$\sec(\sec^{-1}x) = x,\ x \in \mathbb{R} - (-1,1)$

$\text{cosec}(\text{cosec}^{-1}x) = x,\ x \in \mathbb{R} - (-1,1)$

(ii) $\sin^{-1}(\sin x) = x,\ x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

$\cos^{-1}(\cos x) = x,\ x \in [0, \pi]$

$\tan^{-1}(\tan x) = x,\ x \in (-\frac{\pi}{2}, \frac{\pi}{2})$

$\cot^{-1}(\cot x) = x,\ x \in (0, \pi)$

$\sec^{-1}(\sec x) = x,\ x \in [0, \pi] - \left\{\frac{\pi}{2}\right\}$

$\text{cosec}^{-1}(\text{cosec}\, x) = x,\ x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] - \{0\}$.

Illustration

$$sin\left(sin^{-1}\left(-\frac{1}{2}\right)\right) = sin\left(-\frac{\pi}{6}\right) = -sin\frac{\pi}{6} = -\frac{1}{2}$$

$$sin^{-1}\left(sin\frac{5\pi}{6}\right) \neq \frac{5\pi}{6}\ \left(\because \frac{5\pi}{6} \notin \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]\right)$$

We have $sin^{-1}\left(sin\frac{5\pi}{6}\right) = sin^{-1}\left(sin\left(\pi - \frac{\pi}{6}\right)\right) = sin^{-1}\left(sin\frac{\pi}{6}\right) = \frac{\pi}{6}$.

Similarly $cos^{-1}\left(cos\frac{4\pi}{3}\right) = cos^{-1}\left(cos\left(2\pi - \frac{2\pi}{3}\right)\right) = cos^{-1}\left(cos\frac{2\pi}{3}\right) = \frac{2\pi}{3}$.

Property-II :

(i) $\sin^{-1}(-x) = -\sin^{-1}x,\quad x \in [-1,1]$

(ii) $\text{cosec}^{-1}(-x) = -\text{cosec}^{-1}x,\quad |x| \geq 1$

(iii) $\tan^{-1}(-x) = -\tan^{-1}x,\quad x \in \mathbb{R}$

Proof .

(i) Let $y = sin^{-1}(-x), y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

$\Leftrightarrow \sin y = -x$

$\Leftrightarrow x = -\sin y = \sin(-y)$

$\Leftrightarrow -y = sin^{-1} x \ (\because -y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

$\Leftrightarrow y = -\sin^{-1}x$

Hence $sin^{-1}(-x) = -sin^{-1}x$.

Similarly properties (ii) and (iii) can be proved.

(iv) $\mathbf{\cos^{-1}(-x) = \pi - \cos^{-1}x, x \in [-1, 1]}$

(v) $\mathbf{\sec^{-1}(-x) = \pi - \sec^{-1}x, |x| \geq 1}$

(vi) $\mathbf{\cot^{-1}(-x) = \pi - \cot^{-1}x, x \in R}$

Proof. (iv) Let $y = \cos^{-1}(-x), y \in [0,\pi]$

$\Leftrightarrow \cos y = -x$

$\Leftrightarrow x = -\cos y = \cos(\pi - y)$

$\Leftrightarrow \pi - y = cos^{-1} x \ (\because y \in [0,\pi] \Leftrightarrow \pi - y \in [0,\pi]$

$\Leftrightarrow y = \pi - \cos^{-1}x$

Hence $\cos^{-1}(-x) = \pi - cos^{-1}x$.

Similarly (v) and (vi) can be proved.

Property - III

(i) $\mathbf{sin^{-1}\left(\frac{1}{x}\right) = cosec^{-1}x, \ |x| \geq 1}$

(ii) $\mathbf{cos^{-1}\left(\frac{1}{x}\right) = sec^{-1}x, \ |x| \geq 1}$

(iii) $\mathbf{tan^{-1}\left(\frac{1}{x}\right) = \begin{cases} \cot^{-1}x, & x > 0 \\ -\pi + \cot^{-1}x, & x < 0 \end{cases}}$

Proof.

(i) Let $y = sin^{-1}\left(\frac{1}{x}\right), |x| \geq 1$

Then $y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] - \{0\}$ since no real value x exists for which $y = 0$.

Further $\frac{1}{x} = \sin y \Leftrightarrow x = \text{cosec}\, y$

Since $y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] - \{0\} = rng\ \text{cosec}^{-1}$, it follows that $y = \text{cosec}^{-1}x$.

Hence $\sin^{-1}\left(\frac{1}{x}\right) = \text{cosec}^{-1}x \quad |x| \geq 1.$

Result (ii) can be proved similarly.

(iii) Let $y = \tan^{-1}\left(\frac{1}{x}\right)$.

Consider first $x > 0$.

Hence $y \in (0, \frac{\pi}{2})$ and $\frac{1}{x} = \tan y \Leftrightarrow x = \cot y$

Since $y = (0, \frac{\pi}{2}) \subset (0, \pi) = rng\ \cot^{-1}$, it follows that $y = \cot^{-1}x$

Hence $\tan^{-1}\frac{1}{x} = \cot^{-1} x,\ x > 0.$

Now Let $x < 0$.

Since $x < 0$, by definition of $\tan^{-1}$, $y \in (\frac{-\pi}{2}, 0)$

Now $y = \tan^{-1}\frac{1}{x} \Leftrightarrow \frac{1}{x} = \tan y \Leftrightarrow x = \cot y = \cot(\pi + y)$

Since $y \in (\frac{-\pi}{2}, 0)$, $\pi + y \in (\frac{\pi}{2}, \pi) \subset (0, \pi) = rng\ \cot^{-1}$

Hence $x = \cot(\pi + y) \Rightarrow \pi + y = \cot^{-1}x$

$$\Rightarrow y = -\pi + \cot^{-1}x$$

$$\Rightarrow \tan^{-1}\frac{1}{x} = -\pi + \cot^{-1}x,\ x < 0.$$

Illustration.

$\tan^{-1}\left(\frac{1}{\sqrt{3}}\right) = \frac{\pi}{6} = \cot^{-1}\sqrt{3}$ where as

$\tan^{-1}\left(-\frac{1}{\sqrt{3}}\right) = -\frac{\pi}{6} \neq \frac{5\pi}{6} = \cot^{-1}\left(-\sqrt{3}\right)$. In fact we have

$$tan^{-1}\left(-\frac{1}{\sqrt{3}}\right)=-\frac{\pi}{6}=\frac{5\pi}{6}-\pi=cot^{-1}\left(-\sqrt{3}\right)-\pi.$$

Property - IV

(i) $\sin^{-1}x+\cos^{-1}x=\frac{\pi}{2},\quad |x|\le 1$

(ii) $\tan^{-1}x+\cot^{-1}x=\frac{\pi}{2},\quad x\in R$

(iii) $\sec^{-1}x+\text{cosec}^{-1}x=\frac{\pi}{2},\quad |x|\ge 1$

Proof.

(i) Let $y=\sin^{-1}x$.

$$\therefore\ y\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$$

$$\Leftrightarrow x=sin\,y=cos\left(\frac{\pi}{2}-y\right)$$

$$\Leftrightarrow \frac{\pi}{2}-y=cos^{-1}x\quad\left(\because \frac{\pi}{2}-y\in[0,\pi]\right)$$

$$\Leftrightarrow \frac{\pi}{2}-sin^{-1}x=cos^{-1}x$$

$$\Leftrightarrow cos^{-1}x+sin^{-1}x=\frac{\pi}{2}.$$

Results (ii) and (iii) can be similarly proved.

Example 1. Prove that

(i) For $0\le x\le 1$.

$$\sin^{-1}x=\cos^{-1}\sqrt{1-x^2}=tan^{-1}\frac{x}{\sqrt{1-x^2}},(x\ne 1)$$

$$=cot^{-1}\frac{\sqrt{1-x^2}}{x},\qquad (x\ne 0)$$

$$=sec^{-1}\frac{1}{\sqrt{1-x^2}},\qquad (x\ne 1)$$

$$=\text{cosec}^{-1}\frac{1}{x}.,\qquad (x\ne 0)$$

(ii) For $-1 \le x < 0$

$$\sin^{-1}x = -\cos^{-1}\sqrt{1-x^2} = tan^{-1}\frac{x}{\sqrt{1-x^2}}, \qquad (x \neq -1)$$

$$= -\pi + cot^{-1}\frac{\sqrt{1-x^2}}{x}$$

$$= -sec^{-1}\frac{1}{\sqrt{1-x^2}}, \qquad (x \neq -1)$$

$$= cosec^{-1}\frac{1}{x}.$$

Solution.

(i) Let $y = \sin^{-1}x$, $0 \le x \le 1$.

(a) Then $y \in \left[0, \frac{\pi}{2}\right]$

and $\sin y = x$ so that $\cos y = \sqrt{1-x^2}$.

Since $y \in \left[0, \frac{\pi}{2}\right] \subset [0,\pi] = rng \cos^{-1}$, $y = \cos^{-1}\sqrt{1-x^2}$.

(b) Now for $x \neq 1$ i.e. $y \neq \frac{\pi}{2}$, $\tan y = \frac{sin\, y}{cos\, y} = \frac{x}{\sqrt{1-x^2}}$.

$\Rightarrow y = \tan^{-1}\frac{x}{\sqrt{1-x^2}}$, $x \neq 1$ since $y \in [0, \frac{\pi}{2}) \subset (-\frac{\pi}{2}, \frac{\pi}{2})$.

(c) Again for $x \neq 0$ i.e. $y \neq 0$, $\cot y = \frac{cos\, y}{sin\, y} = \frac{\sqrt{1-x^2}}{x}$

$\Rightarrow y = \cot^{-1}\frac{\sqrt{1-x^2}}{x}$, $x \neq 0$, since $y \in (0, \frac{\pi}{2}) \subset (0,\pi)$

Similarly it can be proved that

$\sin^{-1}x = \sec^{-1}\frac{1}{\sqrt{1-x^2}}$, $x \neq 1$ and $\sin^{-1}x = \text{cosec}^{-1}\frac{1}{x}$, $x \neq 0$.

(ii) Let $y = \sin^{-1}x$, $-1 \le x < 0$

Let $x = -w$, $w>0$ so that $x^2 = w^2$

$\sin^{-1}x = \sin^{-1}(-w) = -\sin^{-1}w = \text{A (say)}$

(a) $A = -\sin^{-1} w \quad = -\cos^{-1}\sqrt{1-w^2}$ by (i) above.

$$= -\cos^{-1}\sqrt{1-x^2}$$

(b) $A = -\sin^{-1} w \quad = \tan^{-1}\frac{w}{\sqrt{1-w^2}}, \quad$ for $w \neq 1$

$$= \tan^{-1}\frac{(-w)}{\sqrt{1-w^2}}, \quad \text{for } w \neq 1$$

$$= \tan^{-1}\frac{x}{\sqrt{1-x^2}}, \quad \text{for } x \neq -1$$

(c) $A = -\sin^{-1} w \quad = -\cot^{-1}\frac{\sqrt{1-w^2}}{w}, \quad$ for $w \neq 0$

$$= -\left(\pi - \cot^{-1}\frac{\sqrt{1-w^2}}{-w}\right), \ w \neq 0 \ (\sin \cot^{-1}(-w) = -\pi - \cot^{-1} w)$$

$$= -\pi + \cot^{-1}\frac{\sqrt{1-x^2}}{x}, \quad x \neq 0.$$

Similarly other two results can be proved.

Illustration

$$\sin^{-1}\left(\frac{1}{2}\right) = \frac{\pi}{6} = \cos^{-1}\sqrt{1-\frac{1}{4}} = \cos^{-1}\left(\frac{\sqrt{3}}{2}\right) = \tan^{-1}\left(\frac{1}{\sqrt{3}}\right) \text{ etc.}$$

$$\sin^{-1}\left(-\frac{1}{2}\right) = -\frac{\pi}{6} = -\cos^{-1}\left(\frac{\sqrt{3}}{2}\right) = \tan^{-1}\left(\frac{-1}{\sqrt{3}}\right) = -\pi + \cot^{-1}\left(-\sqrt{3}\right) = -\pi + \frac{5\pi}{6} \text{ etc.}$$

Property V

(i) $\mathbf{\tan^{-1}x + \tan^{-1}y = \tan^{-1}\frac{x+y}{1-xy}}, \quad$ **for $xy < 1$**

(ii) $\mathbf{\tan^{-1}x + \tan^{-1}y = \pm\pi + \tan^{-1}\frac{x+y}{1-xy}}, \quad$ **for $xy > 1$**

according as $x > 0$ or $x < 0$.

Before proving (i) above we note the following.

Let $\alpha = \tan^{-1} x$ and $\beta = \tan^{-1} y$ and $xy < 1$.

Then $\alpha, \beta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ and $x = \tan\alpha$, $y = \tan\beta$

Now $tan(\alpha+\beta) = \frac{tan\,\alpha + tan\,\beta}{1 - \tan\alpha\, tan\,\beta} = \frac{x+y}{1-xy}$

From the above equality we cannot conclude that $\alpha+\beta = tan^{-1}\frac{x+y}{1-xy}$ unless it is proved

that $\alpha+\beta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ when $\alpha, \beta \in (-\frac{\pi}{2}, \frac{\pi}{2})$. So we proceed as follows.

Proof.

(i) Let $\alpha = \tan^{-1}x$, $\beta = \tan^{-1}y$, *and* $xy < 1$.
The case when xy=0 is trival.

case (I). Let $x > 0$ and $y > 0$. Then $\alpha, \beta \in (0, \frac{\pi}{2})$.

Now $xy < 1 \Rightarrow x < \frac{1}{y} \Rightarrow tan\,\alpha < cot\,\beta = tan(\frac{\pi}{2} - \beta)$.

Since tan is an increasing function,

$\alpha < \frac{\pi}{2} - \beta \Rightarrow \alpha + \beta < \frac{\pi}{2}$.

Thus $\alpha+\beta \in (0, \frac{\pi}{2})$ and

$$tan(\alpha+\beta) = \frac{tan\,\alpha + tan\,\beta}{1 - tan\,\alpha\, tan\,\beta} = \frac{x+y}{1-xy}$$

$$\Rightarrow \alpha+\beta = tan^{-1}\frac{x+y}{1-xy}$$

$$\Rightarrow tan^{-1}x + tan^{-1}y = tan^{-1}\frac{x+y}{1-xy}.$$

Case (II) Let $x < 0$, $y < 0$. Then $-x > 0$ and $-y > 0$

We have $(-x) > 0$, $(-y) > 0$ and $(-x)(-y) = xy < 1$

By case (I) we have

$$tan^{-1}(-x) + tan^{-1}(-y) = tan^{-1}\frac{(-x)+(-y)}{1-(-x)(-y)}$$

$$\Rightarrow tan^{-1} x + tan^{-1} y = tan^{-1}\frac{x+y}{1-y}, \text{ since } tan^{-1}(-\theta) = -tan^{-1}\theta$$

Case (III) Let $x<0$ and $y>0$. Then $\alpha \in (-\frac{\pi}{2}, 0)$ and $\beta \in (0, \frac{\pi}{2})$

So $\alpha+\beta \in (-\frac{\pi}{2}, \frac{\pi}{2})$

Similarly if $x>0$ and $y<0$ then also $\alpha+\beta \in (-\frac{\pi}{2}, \frac{\pi}{2})$

Now, as in case (I),

$$\tan(\alpha+\beta) = \frac{tan\,\alpha + tan\,\beta}{1 - tan\,\alpha\, tan\,\beta} = \frac{x+y}{1-xy}$$

$$\Rightarrow \tan^{-1}x + \tan^{-1}y = \tan^{-1}\frac{x+y}{1-xy}.$$

This completes proof of (i).

(ii) Let $xy>1$

First suppose $x>0$ obviously $y>0$.

Hence $\alpha, \beta \in (0, \frac{\pi}{2})$.

$$\text{Now } xy>1 \Rightarrow x > \frac{1}{y} \Rightarrow \tan\alpha > \cot\beta = \tan\left(\frac{\pi}{2} - \beta\right)$$

$$\Rightarrow \alpha > \frac{\pi}{2} - \beta \quad \text{(since tan is increasing function)}$$

$$\Rightarrow \alpha+\beta > \frac{\pi}{2} \quad \text{But } 0<\alpha<\frac{\pi}{2} \text{ and } 0<\beta<\frac{\pi}{2}.$$

$$\text{Hence } \frac{\pi}{2} < \alpha+\beta < \pi \Rightarrow \frac{-\pi}{2} < \alpha+\beta-\pi < 0.$$

$$\text{Now } \tan(\alpha+\beta-\pi) = -\tan(\pi-(\alpha+\beta)) = \tan(\alpha+\beta)$$

$$= \frac{tan\,\alpha + tan\,\beta}{1 - tan\,\alpha\, tan\,\beta} = \frac{x+y}{1-xy}$$

$$\text{Hence } \alpha+\beta-\pi = \tan^{-1}\frac{x+y}{1-xy} \text{(since } \alpha+\beta-\pi) = \tan^{-1}\frac{x+y}{1-xy}\left(\text{since } \alpha+\beta-\pi \in (\frac{-\pi}{2}, 0)\right)$$

$$\Rightarrow tan^{-1}x + tan^{-1}y = \pi + tan^{-1}\frac{x+y}{1-xy}.$$

On the other hand if $x < 0$ then $xy > 1 \Rightarrow y < 0$.
Then $-x > 0$, $-y > 0$ and $(-x)(-y) = xy > 1$ give us

$$\tan^{-1}(-x) + \tan^{-1}(-y) = \pi + \tan^{-1}\frac{(-x)+(-y)}{1-(-x)(-y)}$$, as shown previously

$$\Rightarrow \tan^{-1}x + \tan^{-1}y = -\pi + \tan^{-1}\frac{x+y}{1-xy} \text{ (since } \tan(-\theta) = -\tan\theta)$$

Hence proved.

By writing -y in place of y in the above property V (i) and (ii) we get the following properties. However we need to remember only (i) and (ii)./

(iii) $\tan^{-1}x + \tan^{-1}y = \tan^{-1}\frac{x-y}{1+xy}$, for $xy > -1$

(iv) $\tan^{-1}x - \tan^{-1}y = \pm\pi + \tan^{-1}\frac{x-y}{1+xy}$, for $xy < -1$ according as $x>0$ or $x<0$.

We only derive (iv) from (ii) and (iii) is left as exercise.

Proof of (iv).

Let $xy < -1$. Let $x > 0$.
Hence by property V (ii), since $x(-y) = -xy > 1$ and $x > 0$

$$\tan^{-1}x + \tan^{-1}(-y) = \pi + \tan^{-1}\frac{x+(-y)}{1-x(-y)}, \qquad \text{for } x(-y)>1$$

$$\text{i.e. } \tan^{-1}x - \tan^{-1}y = \pi + \tan^{-1}\frac{x-y}{1+xy}, \qquad \text{for } xy < -1.$$

If $x < 0$ then again by Property V (ii),

$$\tan^{-1}x + \tan^{-1}(-y) = -\pi + \tan^{-1}\frac{x+(-y)}{1-x(-y)}, \qquad \text{for } x(-y)>1.$$

$$\text{i.e. } \tan^{-1}x - \tan^{-1}y = -\pi + \tan^{-1}\frac{x-y}{1+xy}, \qquad \text{for } xy < -1.$$

By taking $x = y$ in Property V(i) and (ii) we get the following formulae.

(v) $$2\tan^{-1}x = \begin{cases} tan^{-1}\frac{2x}{1-x^2}, & |x|<1 \\ \pm\pi + tan^{-1}\frac{2x}{1-x^2}, & \text{according as } x>1 \text{ or } x<-1 \end{cases}$$

Illustration

(i) $tan^{-1}\frac{1}{3}+tan^{-1}\frac{1}{2}=tan^{-1}\frac{5/6}{1-1/6}=tan^{-1}1=\frac{\pi}{4}, \quad \left(\frac{1}{3}.\frac{1}{2}<1\right)$

(ii) $tan^{-1}\left(-\frac{1}{3}\right)+tan^{-1}\frac{1}{2}=tan^{-1}\frac{-\frac{1}{3}+\frac{1}{2}}{1-\left(-\frac{1}{3}\right)\cdot\frac{1}{2}}=tan^{-1}\frac{1/6}{7/6}=tan^{-1}\frac{1}{7} \quad \left(-\frac{1}{3}.\frac{1}{2}<1\right)$

(iii) $\tan^{-1}3+\tan^{-1}2=\pi+\tan^{-1}2=\frac{3+2}{1-3.2}=\pi+tan^{-1}(-1)=\pi-\frac{\pi}{4}=\frac{3\pi}{4}$

(iv) $tan^{-1}(-3)+tan^{-1}(-2)=-\pi+tan^{-1}\frac{-3-2}{1-3\cdot2}=\pi+tan^{-1}1=-\pi+\frac{\pi}{4}=-\frac{3\pi}{4}$

(v) $tan^{-1}3-tan^{-1}2=tan^{-1}3+tan^{-1}(-2)$

$=tan^{-1}\frac{3+(-2)}{1-3\cdot(-2)}$, applying Property V (i)

$=tan^{-1}\frac{1}{7}$

Alternatively applying property V (iii)

$\tan^{-1}3-\tan^{-1}2=\tan^{-1}\frac{3-2}{1+3.2}$ (since $3.2=6>-1$)

$=\tan^{-1}\frac{1}{7}$

Property-VI

(i) For $x^2+y^2\le 1$

$$\sin^{-1}x+\sin^{-1}y=\sin^{-1}\left(x\sqrt{1-y^2}+y\sqrt{1-x^2}\right)$$

(ii) For $x^2+y^2>1$ with $x,y\in[-1,1]$

$$\sin^{-1}x+\sin^{-1}y=\begin{cases} sin^{-1}\left(x\sqrt{1-y^2}+y\sqrt{1-x^2}\right) & \text{if } x \text{ and } y \text{ have opposite signs.} \\ \pm\pi-sin^{-1}\left(x\sqrt{1-y^2}+y\sqrt{1-x^2}\right) & \text{according as } x,y\ge 0 \text{ or } x,y\le 0 \end{cases}$$

Proof.

(i) Let $x^2+y^2\le 1$.

$\Rightarrow x^2\le 1$ and $y^2\le 1\Rightarrow x,y\in[-1,1]$

Now let $\alpha = \sin^{-1}x$ and $\beta = \sin^{-1}y$.

$\Rightarrow x = \sin\alpha,\ y = \sin\beta$ and $\alpha, \beta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

$\therefore \cos\alpha = \sqrt{1-x^2}, \cos\beta = \sqrt{1-y^2}$ since $\cos\alpha$ and $\cos\beta$ are both positive in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

Now $\sin(\alpha+\beta) = \sin\alpha\sin\beta + \cos\alpha\sin\beta = x\sqrt{1-y^2} + y\sqrt{1-x^2}$ ... (1)

Case. 1

$x \geq 0$ and $y \geq 0$. $\Rightarrow \alpha, \beta \in \left[0, \frac{\pi}{2}\right]$

Now $x^2+y^2 \leq 1 \Rightarrow x^2 \leq 1-y^2 \Rightarrow x \leq \sqrt{1-y^2} \Rightarrow \sin\alpha \leq \cos\beta$

$\Rightarrow \sin\alpha \leq \sin\left(\frac{\pi}{2}-\beta\right) \Rightarrow \alpha \leq \frac{\pi}{2}-\beta$ (since sin is an increasing function in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$)

$\therefore \alpha+\beta \leq \frac{\pi}{2}$. Thus $0 \leq \alpha+\beta \leq \frac{\pi}{2}$.

Hence from (i) we get $\alpha+\beta = \sin^{-1}\left(x\sqrt{1-y^2} + y\sqrt{1-x^2}\right)$

i.e. $\sin^{-1}x + \sin^{-1}y = \sin^{-1}\left(x\sqrt{1-y^2} + y\sqrt{1-x^2}\right)$.

Case 2

$x \leq 0$ and $y \leq 0$

$\Rightarrow -x \geq 0$ and $-y \geq 0$.

Hence by case (1)

$\sin^{-1}(-x) + \sin^{-1}(-y) = \sin^{-1}\left((-x)\sqrt{1-y^2} + (-y)\sqrt{1-x^2}\right)$

$\Rightarrow -\sin^{-1}x - \sin^{-1}y = -\sin^{-1}\left(x\sqrt{1-y^2} + (y)\sqrt{1-x^2}\right)$

$\Rightarrow \sin^{-1}x + \sin^{-1}y = \sin^{-1}\left(x\sqrt{1-y^2} + y\sqrt{1-x^2}\right)$

***Case* 3.**

Let $x \geq 0$ and $y \leq 0$

Then $\alpha \in \left[0, \frac{\pi}{2}\right]$ and $\beta \in \left[-\frac{\pi}{2}, 0\right]$

Hence $\alpha+\beta \in \left[-\frac{\pi}{2},\frac{\pi}{2}\right]$ and from (i) we have

$$\alpha+\beta = sin^{-1}\left(x\sqrt{1-x^2}+y\sqrt{1-x^2}\right)$$

$$\Rightarrow sin^{-1}x+sin^{-1}y = sin^{-1}\left(x\sqrt{1-y^2}+y\sqrt{1-x^2}\right)$$

The case when $x \le 0$ and $y \ge 0$ is similarly dealt with.
This proves (i).

(ii) Let $x^2+y^2 \ge 1$ while $x,y \in [-1,1]$.

Let $\alpha = \sin^{-1}x$ and $\beta = \sin^{-1}y$

$$\Rightarrow x = sin\,\alpha,\ y = sin\,\beta \text{ and } \alpha,\beta \in \left[-\frac{\pi}{2},\frac{\pi}{2}\right].$$

As shown earlier

$$sin(\alpha+\beta) = sin\,\alpha\,cos\,\beta + cos\,\alpha\,sin\,\beta = x\sqrt{1-y^2}+y\sqrt{1-x^2} \quad .... \quad (2)$$

Case 1.

First let $x \ge 0$ and $y \le 0$.

Then $\alpha \in \left[0,\frac{\pi}{2}\right]$ and $\beta \in \left(-\frac{\pi}{2},0\right) \Rightarrow \alpha+\beta \in \left[-\frac{\pi}{2},\frac{\pi}{2}\right]$

Similarly if $x \le 0$ and $y \ge 0$ then $\alpha+\beta \in \left[-\frac{\pi}{2},\frac{\pi}{2}\right]$

Hence from (2) we get in either case,

$$\alpha+\beta = sin^{-1}\left(x\sqrt{1-y^2}+y\sqrt{1-x^2}\right)$$

i.e. $sin^{-1}x+sin^{-1}y = sin^{-1}\left(x\sqrt{1-y^2}+y\sqrt{1-x^2}\right)$

***Case* 2.** First suppose that $x \ge 0$ and $y \ge 0$.

Then $\alpha,\beta \in \left[0,\frac{\pi}{2}\right] \Rightarrow \alpha+\beta \in [0,\pi]$.

Now $x^2+y^2 \ge 1 \Rightarrow x^2 \ge 1-y^2 \Rightarrow x \ge \sqrt{1-y^2}$

$$\Rightarrow sin\,\alpha \ge cos\,\beta = sin\left(\frac{\pi}{2}-\beta\right)$$

$$\Rightarrow \alpha \geq \frac{\pi}{2} - \beta \text{ since sin is increasing in } \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

$$\Rightarrow \alpha + \beta \geq \frac{\pi}{2}$$

Thus $\frac{\pi}{2} \leq \alpha + \beta \leq \pi \Rightarrow -\frac{\pi}{2}, \alpha + \beta - \pi \leq 0$

$$\Rightarrow 0 \leq \pi - (\alpha + \beta) \leq \frac{\pi}{2}.$$

Now $\sin\left(\pi - (\alpha + \beta)\right) = \sin(\alpha + \beta) = x\sqrt{1-y^2} + y\sqrt{1-x^2}$ from (2)

$$\Rightarrow \pi - (\alpha + \beta) = \sin^{-1}\left(x\sqrt{1-y^2} + y\sqrt{1-x^2}\right)$$

$$\Rightarrow (\alpha + \beta) = \pi - \sin^{-1}\left(x\sqrt{1-y^2} + y\sqrt{1-x^2}\right)$$

or $\sin^{-1}x + \sin^{-1}y = \pi - \sin^{-1}\left(x\sqrt{1-y^2} + y\sqrt{1-x^2}\right)$. ... (A)

Next suppose that $x, y \leq 0$.

Then $-x \geq 0$ and $-y \geq 0$. Hence by the previous case.

$$\sin^{-1}(-x) + \sin^{-1}(-y) = \pi - \sin^{-1}\left((-x)\sqrt{1-y^2} + (-y)\sqrt{1-x^2}\right)$$

$$\Rightarrow -(\sin^{-1}x + \sin^{-1}y) = \pi + \sin^{-1}\left(x\sqrt{1-y^2} + y\sqrt{1-x^2}\right)$$

$\Rightarrow \sin^{-1}x + \sin^{-1}y = -\pi - \sin^{-1}\left(x\sqrt{1-y^2} + y\sqrt{1-x^2}\right)$. ... (B)

combining A and B we get (ii).

To evaluate $\sin^{-1}x - \sin^{-1}y$ we write $-y$ in place of y in the formula for $\sin^{-1}x + \sin^{-1}y$.

Illustration

(i) $\Rightarrow sin^{-1}\left(\frac{3}{5}\right)+sin^{-1}\left(\frac{7}{25}\right)=sin^{-1}\left(\frac{3}{5}\sqrt{1-\left(\frac{7}{25}\right)^2}+\frac{7}{25}\sqrt{1-\left(\frac{3}{5}\right)^2}\right)$

$$\left(\because\left(\frac{3}{5}\right)^2+\left(\frac{7}{25}\right)^2=\frac{274}{625}<1\right)$$

$$=sin^{-1}\left(\frac{3}{5}\cdot\frac{24}{25}+\frac{7}{25}\cdot\frac{4}{5}\right)=sin^{-1}\left(\frac{4}{5}\right).$$

(ii) $sin^{-1}\left(\frac{3}{5}\right)-sin^{-1}\left(\frac{7}{25}\right)=sin^{-1}\left(\frac{3}{5}\right)+sin^{-1}\left(\frac{-7}{25}\right)$

$$=sin^{-1}\left(\frac{3}{5}\cdot\frac{24}{25}-\frac{7}{25}\cdot\frac{4}{5}\right)=sin^{-1}\left(\frac{44}{125}\right).$$

(iii) $sin^{-1}\left(\frac{4}{5}\right)+sin^{-1}\left(\frac{24}{25}\right)=\pi-sin^{-1}\left(\frac{4}{5}\cdot\frac{7}{25}+\frac{24}{25}\cdot\frac{3}{5}\right)$

$$\left(\therefore\left(\frac{4}{5}\right)^2+\left(\frac{25}{25}\right)^2=\frac{976}{625}>1\right)$$

$$=\pi-sin^{-1}\left(\frac{4}{5}\right).$$

Property- VII

(i) For $x+y\geq 0$ with $-1\leq x, y\leq 1$

$\cos^{-1}x+\cos^{-1}y=\cos^{-1}\left(xy-\sqrt{1-x^2}\sqrt{1-y^2}\right)$

(ii) For $x+y\leq 0$ with $-1\leq x, y\leq 1$

$\cos^{-1}x+\cos^{-1}y=2\pi-cos^{-1}\left(xy-\sqrt{1-x^2}\sqrt{1-y^2}\right)$

Proof

Let $\alpha=cos^{-1}x, \beta=cos^{-1}y, -1\leq x, y\leq 1.$

$\Rightarrow \alpha,\beta\in[0,\pi]$ and $x=cos\,\alpha, y=cos\,\beta.$

Hence $\alpha+\beta\in[0,2\pi]$ and $\sin\alpha=\sqrt{1-x^2}, sin\,\beta=\sqrt{1-y^2}.$

Case (i)

Let $x + y \geq 0$.

$\Rightarrow \cos\alpha + \cos\beta \geq 0$

$\Rightarrow \cos\alpha > -\cos\beta = \cos(\pi - \beta)$

$\Rightarrow \alpha \leq \pi - \beta$ since cos is decreasing in $[0,\pi]$

$\Rightarrow \alpha + \beta \leq \pi$

$\Rightarrow 0 \leq \alpha + \beta \leq \pi$ .... (1)

Now $\cos(\alpha+\beta) = \cos\alpha \cos\beta - \sin\alpha \sin\beta = xy - \sqrt{1-x^2}\sqrt{1-y^2}$

$\Rightarrow \alpha + \beta = \cos^{-1}\left(xy - \sqrt{1-x^2}\sqrt{1-y^2}\right)$ using (1)

i.e. $\cos^{-1}x + \cos^{-1}y = \cos^{-1}\left(xy - \sqrt{1-x^2}\sqrt{1-y^2}\right)$

Case (ii)

Let $x + y \leq 0$

$\Rightarrow \alpha + \beta \geq \pi$, reversing the inequalities in case (i)

$\Rightarrow \pi \leq \alpha + \beta \leq 2\pi$.

$\Rightarrow -\pi \geq -(\alpha + \beta) \geq -2\pi$

$\Rightarrow 2\pi - \pi \geq 2\pi - (\alpha + \beta) \geq 0$

$\Rightarrow 0 \leq 2\pi - (\alpha + \beta) \leq \pi$

$\cos(2\pi - (\alpha + \beta)) = \cos(\alpha + \beta) = xy - \sqrt{1-x^2}\sqrt{1-y^2}$ as in cosec (i)

$\Rightarrow 2\pi - (\alpha + \beta) = \cos^{-1}\left(xy - \sqrt{1-x^2}\sqrt{1-y^2}\right)$

$\Rightarrow (\alpha + \beta) = 2\pi - \cos^{-1}\left(xy - \sqrt{1-x^2}\sqrt{1-y^2}\right)$

i.e. $\cos^{-1}x + \cos^{-1}y = 2\pi - \cos^{-1}\left(xy - \sqrt{1-x^2}\sqrt{1-y^2}\right)$

Hence proved.

Now to evaluate $\cos^{-1}x - \cos^{-1}y$ we write $\cos^{-1}x - \cos^{-1}y = \{\cos^{-1}x + \cos^{-1}(-y)\} - \pi$, (since $\cos^{-1}(-y) = \pi - \cos^{-1}y$) and apply the formula used for $\cos^{-1}x + \cos^{-1}y$. However we state the formula for $\cos^{-1}x - \cos^{-1}y$ below for ready use.

(iii) $$\cos^{-1}x - \cos^{-1}y = \begin{cases} \cos^{-1}\left(xy + \sqrt{1-x^2}\sqrt{1-y^2}\right) & \text{if } x \leq y \\ -\cos^{-1}\left(xy + \sqrt{1-x^2}\sqrt{1-y^2}\right) & \text{if } x \geq y \end{cases}$$

when $-1 \le x, y \le 1$

Proof is left as exercise.

Illustration

(i) $$\cos^{-1}\left(-\frac{7}{25}\right)+\cos^{-1}\left(\frac{3}{5}\right)=\cos^{-1}\left(\frac{-7}{25}\cdot\frac{3}{5}-\sqrt{1-\left(\frac{7}{25}\right)^2}\sqrt{1-\left(\frac{3}{5}\right)^2}\right)$$

$$\left(\because \frac{-7}{25}+\frac{3}{5}>0\right)$$

$$=\cos^{-1}\left(-\frac{117}{125}\right)=\pi-\cos^{-1}\left(\frac{117}{125}\right).$$

(ii) $$\cos^{-1}\left(\frac{3}{5}\right)-\cos^{-1}\left(\frac{7}{25}\right)=\cos^{-1}\left(\frac{3}{5}\right)+\cos^{-1}\left(\frac{-7}{25}\right)-\pi$$

$$=\cos^{-1}\left(-\frac{117}{125}\right)-\pi \quad \text{(as in (i))}$$

$$=-\cos^{-1}\left(\frac{117}{125}\right)$$

or alternatively

$$=\cos^{-1}\left(\frac{3}{5}\right)-\cos^{-1}\left(\frac{7}{25}\right)=\cos^{-1}\left(\frac{3}{5}\cdot\frac{7}{25}+\sqrt{1-\left(\frac{3}{5}\right)^2}\sqrt{1-\left(\frac{7}{25}\right)^2}\right) \text{ by property VII (iii)}$$

$$=-\cos^{-1}\left(\frac{117}{125}\right)$$

(iii) $$\cos^{-1}\left(\frac{7}{25}\right)+\cos^{-1}\left(-\frac{3}{5}\right)=2\pi-\cos^{-1}\left(\frac{7}{25}\cdot\left(\frac{-3}{5}\right)-\sqrt{1-\left(\frac{7}{25}\right)^2}\sqrt{1-\left(\frac{3}{5}\right)^2}\right)$$

$$\left(\because \frac{7}{25}-\frac{3}{5}<0\right)$$

$$=2\pi-\cos^{-1}\left(-\frac{117}{125}\right)$$

$$=2\pi-\left(\pi-\cos^{-1}\left(\frac{117}{125}\right)\right)$$

$$= \pi + cos^{-1}\left(\frac{117}{125}\right).$$

Property VIII.

$$2\sin^{-1}x = \begin{cases} sin^{-1}\left(2x\sqrt{1-x^2}\right), & -\frac{1}{\sqrt{2}} \le x \le \frac{1}{\sqrt{2}} \\ \pi - sin^{-1}\left(2x\sqrt{1-x^2}\right), & \frac{1}{\sqrt{2}} \le x \le 1 \\ -\pi - sin^{-1}\left(2x\sqrt{1-x^2}\right), & -1 \le x \le -\frac{1}{\sqrt{2}} \end{cases}$$

Proof.

Let $\alpha = \sin^{-1}x$

Hence x = sin α and $\alpha \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$

$\Rightarrow cos\,\alpha = \sqrt{1-x^2}$ $\left(\because \cos\alpha > 0 \text{ for } \alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]\right)$

$\therefore\ sin\,2\alpha = 2\,sin\,\alpha\,cos\,\alpha = 2x\sqrt{1-x^2}$ .... (1)

Case **(i)** Let $-\frac{1}{\sqrt{2}} \le x \le \frac{1}{\sqrt{2}}$

$\Rightarrow sin\left(\frac{-\pi}{4}\right) \le sin\,\alpha \le sin\frac{\pi}{4}$

$\Rightarrow -\frac{\pi}{4} \le \alpha \le \frac{\pi}{4}$ $\left(\text{sin is increasing in } \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]\right)$

$\Rightarrow -\frac{\pi}{2} \le 2\alpha \le \frac{\pi}{2}.$

Now $sin\,2\alpha = 2\,sin\,\alpha\,cos\,\alpha = 2x\sqrt{1-x^2}$

$\Rightarrow 2\alpha = sin^{-1}\left(2x\sqrt{1-x^2}\right)$

i.e. $2sin^{-1}x = \sin^{-1}(2x\sqrt{1-x^2})$

Case (ii) Let $\frac{1}{\sqrt{2}} \le x \le 1.$

Case (i). First suppose that $0 \le x \le 1$.

$\Rightarrow 0 \le \alpha \le \frac{\pi}{2} \Rightarrow 0 \le 2\alpha \le \pi.$

From (1) if follows that

$2\alpha = \cos^{-1}(2x^2-1)$ i.e. $2\cos^{-1}x = \cos^{-1}(2x^2-1)$.

***Case* (ii).** Again suppose that $-1 \le x \le 0$.

Then $-1 \le \cos\alpha \le 0 \Rightarrow \frac{\pi}{2} \le \alpha \le \pi$, since cos is a decreasing function on $[\frac{\pi}{2}, \pi]$

$\therefore \pi \le 2\alpha \le 2\pi.$

From the first incquality in (2) we get

$-\pi \ge -2\alpha \Rightarrow \pi \ge 2\pi - 2\alpha$ (adding 2π on both sides)

and from the second inequality in (2) we get

$2\pi - 2\alpha \ge 0.$

Thus $0 \le 2\pi - 2\alpha \le \pi$.

Now $\cos(2\pi - 2\alpha) = \cos 2\alpha = 2x^2 - 1$ (using (1))

$\Rightarrow 2\pi - 2\alpha = \cos^{-1}(2x^2-1)$

$\Rightarrow 2\alpha = 2\pi - \cos^{-1}(2x^2-1)$

i.e. $2\cos^{-1}x = 2\pi - \cos^{-1}(2x^2-1)$.

Hence proved.

Property X

$$3\sin^{-1}x = \begin{cases} \sin^{-1}(3x-4x^3) & \text{if } -\frac{1}{2} \le x \le \frac{1}{2} \\ \pi - \sin^{-1}(3x-4x^3) & \text{if } \frac{1}{2} \le x \le 1 \\ -\pi - \sin^{-1}(3x-4x^3) & \text{if } -1 \le x \le -\frac{1}{2} \end{cases}$$

Proof .

Let $\alpha = \sin^{-1}x$

Then $x = \sin\alpha$ and $\alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

Then $\sin 3\alpha = 3\sin\alpha - 4\sin^3\alpha = 3x - 4x^3$ ... ... (1)

Case (i) Let $-\frac{1}{2} \le x \le \frac{1}{2}$

$\Rightarrow \sin\left(\frac{-\pi}{6}\right) \le \sin\alpha \le \sin\left(\frac{\pi}{6}\right)$

$\Rightarrow -\frac{\pi}{6} \le \alpha \le \frac{\pi}{6}$ (Since sine increases in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$)

$\Rightarrow -\frac{\pi}{2} \le 3\alpha \le \frac{\pi}{2}$

Now $\sin 3\alpha = 3x - 4x^3$ (from (1))

$\Rightarrow 3\alpha = \sin^{-1}(3x-4x^3)$ (using (2))

i.e. $3\sin^{-1}x = \sin^{-1}(3x - 4x^3)$

Case - (ii)

Let $\frac{1}{2} \le x \le 1$

$\Rightarrow \frac{\pi}{6} \le \alpha \le \frac{\pi}{2}$

$\Rightarrow \frac{\pi}{2} \le 3\alpha \le 3\frac{\pi}{2}$

$\Rightarrow -\frac{\pi}{2} \le 3\alpha - \pi \le \frac{\pi}{2}$

Now $\sin(\pi - 3\alpha) = \sin 3\alpha = 3x - 4x^3$ by (1)

$\Rightarrow \pi - 3\alpha = sin^{-1}(3x - 4x^3)$

$\Rightarrow 3\alpha = \pi - \sin^{-1}(3x - 4x^3)$

i.e. $3sin^{-1}x = \pi - sin^{-1}(3x - 4x^3)$.

Proof of case (iii) is left as exercise.

We now give below similar properties for inverse cosine function which can be proved using similar argument.

Property XI

$$3\cos^{-1}x = \begin{cases} cos^{-1}(4x^3 - 3x) & \text{if } \frac{1}{2} \le x \le 1 \\ 2\pi - \cos^{-1}(4x^3 - 3x) & \text{if } -\frac{1}{2} \le x \le \frac{1}{2} \\ 2\pi + cos^{-1}(4x^3 - 3x) & \text{if } -1 \le x \le -\frac{1}{2} \end{cases}$$

Proof.

Let $\alpha = \cos^{-1}x$.

Then $x = \cos\alpha$ and $\alpha \in [0, \pi]$.

$\cos 3\alpha = 4\cos^3\alpha - 3\cos\alpha = 4x^3 - 3x$ .... (1)

Case (i)

Let $\frac{1}{2} \le x \le 1$

$\therefore \cos\frac{\pi}{3} \le \cos\alpha \le \cos 0$

$\Rightarrow 0 \le \alpha \le \frac{\pi}{3}$, (since cos is decreasing in $[0,\pi]$)

$\Rightarrow 0 \le 3\alpha \le \pi$

Now $\cos 3\alpha = 4x^3 - 3x$ .... (1)

$\Rightarrow 3\alpha = \cos^{-1}(4x^3 - 3x)$

i.e. $3\cos^{-1}x = \cos^{-1}(4x^3 - 3x)$ when $\frac{1}{2} \le x \le 1$.

Case **(ii)** When $x \in \left[-\frac{1}{2}, \frac{1}{2}\right]$ and **case (iii)** when $x \in \left[-1, \frac{-1}{2}\right]$ are left as exercises.

Alternatively using property X for $3\sin^{-1}x$ we can prove the Property XI for $3\cos^{-1}x$.

For example to prove the property when $x \in -\left[-1, -\frac{1}{2}\right]$ we proceed as follows.

Alternative proof for case (iii) :

We know that for $x \in \left[-1, -\frac{1}{2}\right]$

$3\sin^{-1}x = -\pi - \sin^{-1}(3x - 4x^3)$. ... ... (A)

Also $\sin^{-1}x + \cos^{-1}x = \frac{\pi}{2}$

Hence $3\cos^{-1}x = 3(\frac{\pi}{2} - \sin^{-1}x)$

$$= \frac{3\pi}{2} + -3\sin^{-1}x$$

$$= \frac{3\pi}{2} + \pi + \sin^{-1}(3x - 4x^3) \quad \text{by (A)}$$

$$= \frac{5\pi}{2} + \frac{\pi}{2} - \cos^{-1}(3x - 4x^3)$$

$$= 3\pi - \cos^{-1}(3x - 4x^3)$$

$$= 2\pi + \pi - \cos^{-1}(3x - 4x^3)$$

$$= 2\pi + \cos^{-1}(4x^3 - 3x).$$

Property XII

$$3\tan^{-1}x=\begin{cases}\tan^{-1}\dfrac{3x-x^3}{1-3x^2}, & -\dfrac{1}{\sqrt{3}}<x<\dfrac{1}{\sqrt{3}}\\ \pi+\tan^{-1}\dfrac{3x-x^3}{1-3x^2}, & x>\dfrac{1}{\sqrt{3}}\\ -\pi+\tan^{-1}\dfrac{3x-x^3}{1-3x^2}, & x<-\dfrac{1}{\sqrt{3}}\end{cases}$$

Proof. Let $\alpha=\tan^{-1}x$

Then $x=\tan\alpha$ and $\alpha\in(-\frac{\pi}{2},\frac{\pi}{2})$

$$\tan 3\alpha=\frac{3\tan\alpha-\tan^3\alpha}{1-3\tan^2\alpha}=\frac{3x-x^3}{1-3x^2} \quad \dots \quad \dots \quad (1)$$

Case (i) Let $-\frac{1}{\sqrt{3}}<x<\frac{1}{\sqrt{3}}$

$\Rightarrow \tan\left(-\frac{\pi}{6}\right)<\tan\alpha<\tan\left(\frac{\pi}{6}\right)$

$\Rightarrow -\frac{\pi}{6}<\alpha<\frac{\pi}{6}$ [since tan is an increasing function on $(-\frac{\pi}{2},\frac{\pi}{2})$]

$\therefore -\frac{\pi}{2}<3\alpha<\frac{\pi}{2}$.

Hence from (1) $3\alpha=\tan^{-1}\left(\frac{3x-x^3}{1-3x^2}\right)$

i.e. $3\tan^{-1}x=\tan^{-1}\left(\frac{3x-x^3}{1-3x^2}\right)$

Case (ii) Let $x>\frac{1}{\sqrt{3}}$

Then $\tan\alpha>\tan\frac{\pi}{6}\Rightarrow\alpha>\frac{\pi}{6}$.

Thus $\frac{\pi}{6}<\alpha<\frac{\pi}{2}\Rightarrow\frac{\pi}{2}<3\alpha<\frac{3\pi}{2}$.

$$\Rightarrow \frac{-\pi}{2} < 3\alpha - \pi < \frac{3\pi}{2} - \pi = \frac{\pi}{2}$$

Now $\tan(3\alpha - \pi) = -\tan(\pi - 3\alpha)$

$$= \tan 3\alpha$$

$$= \frac{3x - x^3}{1 - 3x^2} \quad \text{by (1)}$$

$$\Rightarrow 3\alpha - \pi = \tan^{-1}\left(\frac{3x - x^3}{1 - 3x^2}\right)$$

i.e. $3\tan^{-1} x = \pi + \tan^{-1}\frac{3x - x^3}{1 - 3x^2}$

case (iii) is left as an exercise.

Example 2 :

Find the value of $\cos \tan^{-1} \cot \cos^{-1}\frac{\sqrt{3}}{2}$

Solution :

$$\cos^{-1}\frac{\sqrt{3}}{2} = \theta \Rightarrow \cos\theta = \frac{\sqrt{3}}{2}$$

$$\Rightarrow \theta = \frac{\pi}{6} \Rightarrow \cos^{-1}\frac{\sqrt{3}}{2} = \frac{\pi}{6}.$$

$$\therefore \cos \tan^{-1} \cot \cos^{-1}\frac{\sqrt{3}}{2} = \cos \tan^{-1} \cot \frac{\pi}{6} = \cos \tan^{-1} \sqrt{3}$$

$$= \cos\frac{\pi}{3} = \frac{1}{2}. \qquad \left(\because \tan^{-1}\sqrt{3} = \frac{\pi}{3}\right)$$

Example 3.

Express the following in the simplest form.

(i) $\sin^{-1}\left(\frac{\sqrt{1+x} + \sqrt{1-x}}{2}\right), \qquad 0 < x < 1$

(ii) $\tan^{-1}\left(\sqrt{1+x^2} - x\right)$

Solution (i)

Let $x = \cos 2\theta$

$$\Rightarrow \sqrt{1+x} = \sqrt{1+\cos 2\theta} = \pm\sqrt{2}\cos\theta \quad \text{and}$$

$$\sqrt{1-x} = \sqrt{1-\cos 2\theta} = \pm\sqrt{2}\sin\theta$$

Given $0 < x < 1, 2\theta \in \left[0, \frac{\pi}{2}\right] \Rightarrow \theta \in \left(0, \frac{\pi}{4}\right)$.

So $\cos\theta > 0$ and $\sin\theta > 0 \Rightarrow \sqrt{1+x} = \sqrt{2}\cos\theta$ and $\sqrt{1-x} = \sqrt{2}\sin\theta$.

$$\therefore \frac{\sqrt{1+x}+\sqrt{1-x}}{2} = \frac{\sqrt{2}(\cos\theta+\sin\theta)}{2} = \frac{1}{\sqrt{2}}\cos\theta + \frac{1}{\sqrt{2}}\sin\theta = \sin\left(\frac{\pi}{4}+\theta\right)$$

Hence $\sin^{-1}\left(\frac{\sqrt{1+x}+\sqrt{1-x}}{2}\right) = \sin^{-1}\left(\sin\left(\frac{\pi}{4}+\theta\right)\right)$

$$= \frac{\pi}{4}+\theta = \frac{\pi}{4}+\frac{1}{2}\cos^{-1}x.$$

(ii) Let $x = \cot\theta$

$\Rightarrow \sqrt{1+x^2} - x = \operatorname{cosec}\theta - \cot\theta$

$$= \frac{1-\cos\theta}{\sin\theta} = \frac{2\sin^2\frac{\theta}{2}}{2\sin\frac{\theta}{2}\cos\frac{\theta}{2}} = \tan\frac{\theta}{2}.$$

$$\therefore \tan^{-1}\left(\sqrt{1+x^2}-x\right) = \tan^{-1}\left(\tan\frac{\theta}{2}\right) = \frac{\theta}{2} = \frac{1}{2}\cot^{-1}x.$$

Example 4

Prove that

$$2\tan^{-1}x = \begin{cases} \sin^{-1}\frac{2x}{1+x^2}, & |x| \le 1 \\ \cos^{-1}\frac{1-x^2}{1+x^2}, & x \ge 0 \end{cases}$$

Solution.

(i) Let $\tan^{-1}x = \theta \Rightarrow x = \tan\theta$ and $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$

Now $\sin 2\theta = \frac{2\tan\theta}{1+\tan^2\theta} = \frac{2x}{1+x^2}$.

Given $-1 \le x \le 1$, $\theta \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right] \Rightarrow 2\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

$$\therefore 2\theta = \sin^{-1}\frac{2x}{1+x^2}$$

i.e. $2\tan^{-1}x = \sin^{-1}\frac{2x}{1+x^2}, \qquad |x| \le 1.$

(ii) Let $x \ge 0$ so that $\theta \in [0, \frac{\pi}{2})$

$$\cos 2\theta = \frac{1-\tan^2\theta}{1+\tan^2\theta} = \frac{1-x^2}{1+x^2}$$

Now $\theta \in [0, \frac{\pi}{2}) \Rightarrow 2\theta \in [0, \pi)$

$$\therefore 2\theta = \cos^{-1}\frac{1-x^2}{1+x^2}$$

i.e. $2\tan^{-1}x = \cos^{-1}\frac{1-x^2}{1+x^2}, \quad x \ge 0.$

It is left as exercise to show that

(i) $2\tan^{-1}x = \pm\pi - \sin^{-1}\frac{2x}{1+x^2}$, according as $x > 1$ or $x < 1$

(ii) $2\tan^{-1}x = -\cos^{-1}\frac{1-x^2}{1+x^2}, \qquad if\ x \le 0.$

Example 5.

If $\theta = \tan^{-1}\sqrt{\frac{a(a+b+c)}{bc}} + \tan^{-1}\sqrt{\frac{b(a+b+c)}{ca}} + \tan^{-1}\sqrt{\frac{c(a+b+c)}{ab}}$, find the value of $\tan\theta$.

Solution,

Let $k = \sqrt{\frac{a+b+c}{abc}}$. Then $\theta = \tan^{-1}ak + \tan^{-1}bk + \tan^{-1}ck$

Since $\tan(A+B+C) = \frac{\tan A + \tan B + \tan C - \tan A \tan B \tan C}{1 - \tan A \tan B - \tan B \tan C - \tan C \tan A}$

writing $A = \tan^{-1}ak$, $B = \tan^{-1}bk$, $C = \tan^{-1}ck$ we get $\theta = A+B+C$ and

$$\tan\theta = \frac{ak+bk+ck-abck^3}{1-abk^2-bck^2-cak^2}$$

$$= \frac{k(a+b+c-abck^2)}{1-abk^2-bck^2-cak^2} = 0 \qquad (\text{since } k^2abc = a+b+c).$$

Example-6

Prove that $2\tan^{-1}\frac{1}{5} - \tan^{-1}\frac{1}{4} = \tan^{-1}\frac{8}{53}$

Proof :

$$2\tan^{-1}\frac{1}{5} = \tan^{-1}\frac{2.\frac{1}{5}}{1-\frac{1}{25}} = \tan^{-1}\frac{5}{12} \quad \text{[Property V(v)]}$$

$$\therefore \text{L.H.S.} = \tan^{-1}\frac{5}{12} - \tan^{-1}\frac{1}{4}$$

$$= \tan^{-1}\frac{\frac{5}{12}-\frac{1}{4}}{1+\frac{5}{12}.\frac{1}{4}} = \tan^{-1}\left(\frac{2}{12}\times\frac{48}{53}\right) = \tan^{-1}\frac{8}{53} = \text{R.H.S.}$$

Example-7

Prove that $\sin^{-1}\frac{4}{5} + \sin^{-1}\frac{5}{13} + \sin^{-1}\frac{16}{65} = \frac{\pi}{2}$

Proof :

We shall prove that $\sin^{-1}\frac{4}{5} + \sin^{-1}\frac{5}{13} = \frac{\pi}{2} - \sin^{-1}\frac{16}{65}$

Now $\sin^{-1}\frac{4}{5} + \sin^{-1}\frac{5}{13}$

$$= \sin^{-1}\left\{\frac{4}{5}\sqrt{1-\frac{5^2}{13^2}} + \frac{5}{13}\sqrt{1-\frac{4^2}{5^2}}\right\} \qquad \left(\because \left(\tfrac{4}{5}\right)^2 + \left(\tfrac{5}{13}\right)^2 = \frac{3329}{4225} < 1\right)$$

$$= \sin^{-1}\left\{\tfrac{4}{5}\times\tfrac{12}{13} + \tfrac{5}{13}\times\tfrac{3}{5}\right\} = \sin^{-1}\left\{\tfrac{63}{65}\right\}$$

$$= \cos^{-1}\sqrt{1-\frac{63^2}{65^2}} = \cos^{-1}\frac{16}{65}$$

$$= \frac{\pi}{2} - \sin^{-1}\frac{16}{65} \qquad \text{(Using property IV)}$$

Example-8

If $\cos^{-1}x + \cos^{-1}y + \cos^{-1}z = \pi$, prove that $x^2 + y^2 + z^2 + 2xyz = 1$

Proof :

$\cos^{-1}x + \cos^{-1}y + \cos^{-1}z = \pi$,

$\Rightarrow \cos^{-1}x + \cos^{-1}y = \pi - \cos^{-1}z$

$\Rightarrow \cos^{-1}\left(xy - \sqrt{1-x^2}\ \sqrt{1-y^2}\right) = \pi - \cos^{-1}z$

$\Rightarrow xy - \sqrt{1-x^2}\ \sqrt{1-y^2} = \cos(\pi - \cos^{-1}z)$

$\Rightarrow xy - \sqrt{1-x^2}\ \sqrt{1-y^2} = -\cos(\cos^{-1}z)$

$\Rightarrow xy - \sqrt{1-x^2}\ \sqrt{1-y^2} = -z$

$\Rightarrow (xy+z)^2 = (1-x^2)(1-y^2)$

$\Rightarrow x^2y^2 + z^2 + 2xyz = 1 - x^2 - y^2 + x^2y^2$

$\Rightarrow x^2 + y^2 + z^2 + 2xyz = 1.$

Example-9

Solve $\cot^{-1}x + \sin^{-1}\dfrac{1}{\sqrt{5}} = \dfrac{\pi}{4}$

Solution :

Let $\sin^{-1}\dfrac{1}{\sqrt{5}} = \alpha$ so that $\sin\alpha = \dfrac{1}{\sqrt{5}}$

$\Rightarrow \tan\alpha = \frac{1}{2} \Rightarrow \alpha = \tan^{-1}\frac{1}{2}$

Again $\cot^{-1}x = \tan^{-1}\frac{1}{x}$. Now the equation shall be $\tan^{-1}\frac{1}{x} + \tan^{-1}\frac{1}{2} = \frac{\pi}{4}$

$\Rightarrow \tan^{-1}\dfrac{\frac{1}{x}+\frac{1}{2}}{1-\frac{1}{2x}} = \dfrac{\pi}{4} \Rightarrow \tan^{-1}\dfrac{x+2}{2x-1} = \dfrac{\pi}{4}$

$\Rightarrow \dfrac{x+2}{2x-1} = \tan\frac{\pi}{4} = 1 \quad \Rightarrow x = 3.$

EXERCISES - 2

Note : In the problems involving inverse trigonometric functions it is assumed that the relations exist in suitable intervals.

1. Fill in the blanks choosing correct answer from the brackets :

(i) If $A = \tan^{-1}x$, then the value of $\sin 2A$ = ——. $\left(\dfrac{2x}{1-x^2}, \dfrac{2x}{\sqrt{1-x^2}}, \dfrac{2x}{1+x^2}\right)$

(ii) If the value of $\sin^{-1}x = \dfrac{\pi}{5}$ for some $x \in (-1,1)$ then the value of $\cos^{-1}x$ is ——. $\left(\dfrac{3\pi}{10}, \dfrac{5\pi}{10}, \dfrac{7\pi}{10}\right)$

(iii) The value of $\tan^{-1}(2\cos\dfrac{\pi}{3})$ is ——. $(1, \dfrac{\pi}{4}, \dfrac{\pi}{3})$

(iv) If $x + y = 4$, $xy = 1$, then $\tan^{-1}x + \tan^{-1}y$ = ——. $\left(\dfrac{3\pi}{4}, \dfrac{\pi}{4}, \dfrac{\pi}{2}\right)$

(v) The value of $\cot^{-1}2 + \tan^{-1}\dfrac{1}{3}$ = ——. $\left(\dfrac{\pi}{4}, 1, \dfrac{\pi}{2}\right)$

(vi) The principal value of $\sin^{-1}\left(\sin\frac{2\pi}{3}\right)$ is ——. $\left(\frac{2\pi}{3}, \frac{\pi}{3}, \frac{4\pi}{3}\right)$

(vii) If $\sin^{-1}\frac{x}{5} + \text{cosec}^{-1}\frac{5}{4} = \frac{\pi}{2}$, then the value of $x =$ ——. (2,3, 4)

(viii) The value of $\sin(\tan^{-1}x + \tan^{-1}\frac{1}{x}, x>0 =$ ——. (0, 1, ½)

(ix) $\cot^{-1}\left[\frac{\sqrt{1-\sin x} + \sqrt{1+\sin x}}{\sqrt{1-\sin x} - \sqrt{1+\sin x}}\right] =$ —— $(2\pi - \frac{x}{2}, \frac{x}{2}, \pi - \frac{x}{2})$

(x) $2\sin^{-1}\frac{4}{5} + \sin^{-1}\frac{24}{25} =$ ——. (π, $-\pi$, 0)

(xi) *If* $\theta = \cos^{-1}x + \sin^{-1}x - \tan^{-1}x, x \geq 0$, then the smallest interval in which θ lies is___.

$\left(\left(\frac{\pi}{2}, \frac{3\pi}{4}\right), \left[0, \frac{\pi}{2}\right), \left(0, \frac{\pi}{2}\right]\right)$

(xii) $\sec^2(\tan^{-1}2) + \text{cosec}^2(\cot^{-1}3) =$ ____. (16,14,15)

2. Write whether the following statements are true or false.

(i) $\sin^{-1}\frac{1}{x}\ \text{cosec}^{-1}x = 1$

(ii) $\cos^{-1}\frac{4}{5} + \tan^{-1}\frac{2}{3} = \tan^{-1}\frac{17}{6}$

(iii) $\tan^{-1}\frac{4}{3} + \cot^{-1}\left(\frac{-3}{4}\right) = \pi$

(iv) $\sec^{-1}\frac{1}{2} + \text{cosec}^{-1}\frac{1}{2} = \frac{\pi}{2}$

(v) $\sec^{-1}\left(-\frac{7}{5}\right) = \pi - \cos^{-1}\frac{5}{7}$

(vi) $\tan^{-1}(\tan 3) = 3$

(vii) The principal value of $\tan^{-1}\left(\tan\frac{3\pi}{4}\right)$ is $\frac{3\pi}{4}$

(viii) $\cot^{-1}(-\sqrt{3})$ is in the second quadrant.

(ix) $3\tan^{-1}3 = \tan^{-1}\frac{9}{13}$

(xi) $\tan^{-1}2 + \tan^{-1}3 = -\frac{\pi}{4}$

(x) $2\sin^{-1}\frac{4}{5} = \sin^{-1}\frac{24}{25}$

(xii) The equation $\tan^{-1}(\cot x) = 2x$ has exactly two real solutions.

3. Express the value of the following in simplest form.

(i) $\sin(2\sin^{-1}0.6)$

(ii) $\tan(\frac{\pi}{4} + 2\cot^{-1}3)$

(iii) $\cos(2\sin^{-1}x)$

(iv) $\tan(\cos^{-1}x)$

(v) $\tan^{-1}\left(\frac{x}{y}\right) - \tan^{-1}\frac{x-y}{x+y}$

(vi) $\text{cosec}\left(\cos^{-1}\frac{3}{5} + \cos^{-1}\frac{4}{5}\right)$

(vii) $\sin^{-1}\frac{1}{\sqrt{5}} + \cos^{-1}\frac{3}{\sqrt{10}}$

(viii) $\sin\cos^{-1}\tan\sec^{-1}\sqrt{2}$

(ix) $\sin\left(2\tan^{-1}\sqrt{\frac{1-x}{1+x}}\right)$ (x) $\tan\left\{\frac{1}{2}\sin^{-1}\frac{2x}{1+x^2}+\frac{1}{2}\cos^{-1}\frac{1-y^2}{1+y^2}\right\}$

(xi) $\sin\cot^{-1}\cos\tan^{-1}x.$ (xii) $\tan^{-1}\left(x+\sqrt{1+x^2}\right)$

4. Prove the following statements :

(i) $\sin^{-1}\frac{3}{5}+\sin^{-1}\frac{8}{17}=\cos^{-1}\frac{36}{85}$ (ii) $\sin^{-1}\frac{3}{5}+\cos^{-1}\frac{12}{13}=\cos^{-1}\frac{33}{65}$

(iii) $\tan^{-1}\frac{1}{7}+\tan^{-1}\frac{1}{13}=\tan^{-1}\frac{2}{9}$ (iv) $\tan^{-1}\frac{1}{2}+\tan^{-1}\frac{1}{5}+\tan^{-1}\frac{1}{8}=\frac{\pi}{4}$.

(v) $\tan\left(2\tan^{-1}\frac{1}{5}-\frac{\pi}{4}\right)+\frac{7}{17}=0$

5. Prove the following statements :

(i) $\cot^{-1}9+\operatorname{cosec}^{-1}\frac{\sqrt{41}}{4}=\frac{\pi}{4}$ (ii) $\sin^{-1}\frac{4}{5}+2\tan^{-1}\frac{1}{3}=\frac{\pi}{2}$

(iii) $4\tan^{-1}\frac{1}{5}-\tan^{-1}\frac{1}{70}+\tan^{-1}\frac{1}{99}=\frac{\pi}{4}$ (iv) $2\tan^{-1}\frac{1}{5}-\sec^{-1}\frac{5\sqrt{2}}{7}+2\tan^{-1}\frac{1}{8}=\frac{\pi}{4}$

(v) $\cos^{-1}\frac{12}{13}+2\cos^{-1}\sqrt{\frac{64}{65}}+\cos^{-1}\sqrt{\frac{49}{50}}=\cos^{-1}\frac{1}{\sqrt{2}}$

(vi) $\tan^2\cos^{-1}\frac{1}{\sqrt{3}}+\cot^2\sin^{-1}\frac{1}{\sqrt{5}}=6$ (vii) $\cos\tan^{-1}\cot\sin^{-1}x=x.$

6. Prove the following statements :

(i) $\cot^{-1}(\tan 2x)+\cot^{-1}(-\tan 2x)=\pi$

(ii) $\tan^{-1}x+\cot^{-1}(x+1)=\tan^{-1}(x^2+x+1)$

(iii) $\tan^{-1}\frac{a-b}{1+ab}+\tan^{-1}\frac{b-c}{1+bc}=\tan^{-1}a-\tan^{-1}c$

(iv) $\cot^{-1}\frac{pq+1}{p-q}+\cot^{-1}\frac{qr+1}{q-r}+\cot^{-1}\frac{rp+1}{r-p}=0$

(v) $\tan^{-1}\frac{a-b}{1+ab}+\tan^{-1}\frac{b-c}{1+bc}+\tan^{-1}\frac{c-a}{1+ca}$

$=\tan^{-1}\frac{a^2-b^2}{1+a^2b^2}+\tan^{-1}\frac{b^2-c^2}{1+b^2c^2}+\tan^{-1}\frac{c^2-a^2}{1+c^2a^2}$.

7. Prove the following :

(i) $\tan^{-1}\frac{2a-b}{b\sqrt{3}}+\tan^{-1}\frac{2b-a}{a\sqrt{3}}=\frac{\pi}{3}$

(ii) $\tan^{-1}\frac{1}{x+y}+\tan^{-1}\frac{y}{x^2+xy+1}=\tan^{-1}\frac{1}{x}$

(iii) $\sin^{-1}\sqrt{\frac{x-q}{p-q}} = \cos^{-1}\sqrt{\frac{p-x}{p-q}} = \cot^{-1}\sqrt{\frac{p-x}{x-q}}$

(iv) $\sin^2(\sin^{-1}x + \sin^{-1}y + \sin^{-1}z) = \cos^2(\cos^{-1}x + \cos^{-1}y + \cos^{-1}z)$

(v) $\tan(\tan^{-1}x + \tan^{-1}y + \tan^{-1}z) = \cot(\cot^{-1}x + \cot^{-1}y + \cot^{-1}z)$

8. (i) If $\sin^{-1}x + \sin^{-1}y + \sin^{-1}z = \pi$, show that

$x\sqrt{1-x^2} + y\sqrt{1-y^2} + z\sqrt{1-z^2} = 2xyz$

(ii) If $\tan^{-1}x + \tan^{-1}y + \tan^{-1}z = \pi$ show that $x + y + z = xyz$

(iii) If $\tan^{-1}x + \tan^{-1}y + \tan^{-1}z = \frac{\pi}{2}$, show that $xy + yz + zx = 1$

(iv) If $r^2 = x^2 + y^2 + z^2$, Prove that

$\tan^{-1}\frac{yz}{xr} + \tan^{-1}\frac{zx}{yr} + \tan^{-1}\frac{xy}{zr} = \frac{\pi}{2}$

(v) In a triangle ABC if $m\angle A = 90°$, prove that $\tan^{-1}\frac{b}{a+c} + \tan^{-1}\frac{c}{a+b} = \frac{\pi}{4}$ where a, b, c are sides of the triangle.

9. Solve :

(i) $\cos(2\sin^{-1}x) = \frac{1}{9}$

(ii) $\sin^{-1}x + \sin^{-1}(1-x) = \frac{\pi}{2}$

(iii) $\sin^{-1}(1-x) - 2\sin^{-1}x = \frac{\pi}{2}$

(iv) $\cos^{-1}x + \sin^{-1}\frac{x}{2} = \frac{\pi}{6}$

(v) $\tan^{-1}\frac{x-1}{x-2} + \tan^{-1}\frac{x+1}{x+2} = \frac{\pi}{4}$

(vi) $\tan^{-1}\frac{1}{2x+1} + \tan^{-1}\frac{1}{4x+1} = \tan^{-1}\frac{2}{x^2}$

(vii) $3\sin^{-1}\frac{2x}{1+x^2} - 4\cos^{-1}\frac{1-x^2}{1+x^2} + 2\tan^{-1}\frac{2x}{1-x^2} = \frac{\pi}{3}$

(viii) $\cot^{-1}\frac{1}{x-1} + \cot^{-1}\frac{1}{x} + \cot^{-1}\frac{1}{x+1} = \cot^{-1}\frac{1}{3x}$

(ix) $\cot^{-1}\frac{1-x^2}{2x} = \operatorname{cosec}^{-1}\frac{1+a^2}{2a} - \sec^{-1}\frac{1+b^2}{1-b^2}$

(x) $\sin^{-1}\left(\frac{2a}{1+a^2}\right) + \sin^{-1}\left(\frac{2b}{1+b^2}\right) = 2\tan^{-1}x.$

(xi) $\sin^{-1}x - \cos^{-1}x = \cos^{-1}\frac{\sqrt{3}}{2}$

(xii) $\sin^{-1}2x + \sin^{-1}x = \frac{\pi}{3}.$

10. Rectify the error if any in the following :

$\sin^{-1}\frac{4}{5} + \sin^{-1}\frac{12}{13} + \sin^{-1}\frac{33}{65}$

$$= \sin^{-1}\left[\frac{4}{5}\sqrt{1-\frac{144}{169}} + \frac{12}{13}\sqrt{1-\frac{16}{25}}\right] + \sin^{-1}\frac{33}{65}$$

$$= \sin^{-1}\left(\frac{56}{65}\right) + \cos^{-1}\sqrt{1-\left(\frac{33}{65}\right)^2}$$

$$= \sin^{-1}\left(\frac{56}{65}\right) + \cos^{-1}\left(\frac{56}{65}\right) = \frac{\pi}{2}.$$

11. Prove that :

(i) $\cos^{-1}\left(\dfrac{b + a\cos x}{a + b\cos x}\right) = 2\tan^{-1}\left(\sqrt{\dfrac{a-b}{a+b}}\tan\dfrac{x}{2}\right)$

(ii) $\tan\left(\dfrac{\pi}{4}+\dfrac{1}{2}\cos^{-1}\dfrac{a}{b}\right) + \tan\left(\dfrac{\pi}{4}-\dfrac{1}{2}\cos^{-1}\dfrac{a}{b}\right) = \dfrac{2b}{a}$

(iii) $\tan^{-1}\sqrt{\dfrac{xr}{yz}} + \tan^{-1}\sqrt{\dfrac{yr}{zx}} + \tan^{-1}\sqrt{\dfrac{zr}{xy}} = \pi$

where $r = x + y + z$.

12. (i) If $\cos^{-1}\left(\dfrac{x}{a}\right) + \cos^{-1}\left(\dfrac{y}{b}\right) = \theta$, prove that $\dfrac{x^2}{a^2} - \dfrac{2xy}{ab}\cos\theta + \dfrac{y^2}{b^2} = \sin^2\theta.$

(ii) If $\cos^{-1}\left(\dfrac{x}{2}\right) + \cos^{-1}\left(\dfrac{y}{3}\right) = \theta$, prove that $9x^2 - 12xy\cos\theta + 4y^2 = 36\sin^2\theta.$

(iii) If $\sin^{-1}\left(\dfrac{x}{a}\right) + \sin^{-1}\left(\dfrac{y}{b}\right) = \sin^{-1}\left(\dfrac{c^2}{ab}\right)$, prove that $b^2x^2 + 2xy\sqrt{a^2b^2 - c^4} + a^2y^2 = c^4.$

(iv) If $\sin^{-1}\left(\dfrac{x}{a}\right) + \sin^{-1}\left(\dfrac{y}{b}\right) = \alpha$, prove that $\dfrac{x^2}{a^2} + \dfrac{2xy}{ab}\cos\alpha + \dfrac{y^2}{b^2} = \sin^2\alpha.$

(v) If $\sin^{-1}x + \sin^{-1}y + \sin^{-1}z = \pi$, prove that $x^4 + y^4 + z^4 + 4x^2y^2z^2 = 2(x^2y^2 + y^2z^2 + z^2x^2).$

13. Solve the following equations :

(i) $\tan^{-1}\dfrac{x-1}{x+1} + \tan^{-1}\dfrac{2x-1}{2x+1} = \tan^{-1}\dfrac{23}{36}$

(ii) $\tan^{-1}\dfrac{1}{3} + \tan^{-1}\dfrac{1}{5} + \tan^{-1}\dfrac{1}{7} + \tan^{-1}x = \dfrac{\pi}{4}$

(iii) $\cos^{-1}\left(x+\dfrac{1}{2}\right) + \cos^{-1}x + \cos^{-1}\left(x-\dfrac{1}{2}\right) = \dfrac{3\pi}{2}$

(iv) $3\tan^{-1}\dfrac{1}{2+\sqrt{3}} - \tan^{-1}\dfrac{1}{x} = \tan^{-1}\dfrac{1}{3}.$

●

CHAPTER 3

Linear Programming

It appears to me that if one wants to make progress in mathematics, one should study the masters and not the pupils.

- N.H. Abel

3.1 Introduction

Learning mathematics without knowing its usefulness in solving real life problems is drudgery. We face a variety of problems in the field of manufacturing, farming, commerce, construction work, transportation, assignment of jobs, military operations and so on. In the earlier class you learnt about linear inequalities involving one or two variables and their graphical solutions. In this chapter we shall discuss how to utilise those concepts to solve problems faced in many of the real life situations. For this our first step will be to reduce / formulate the real life problem into the language of mathematics. This requires selecting suitable quantities and designating them as variables, connecting them by means of some equations and inequations depending on various conditions and finally trying to solve them to get the desired values.

Operations research is that branch of mathematics which deals with the study of such processes. Linear programming, in particular, discusses the type of problems in operations research where the equations and inequations involve 'linear functions'. The term 'programming' refers to the plan of action or programme we make for solving the problem.

3.2 General Linear Programming Problem (L.P.P.)

We begin with an example to have an insight into the programming problems.

Example 1

A person having a capital of ₹ 5000 has to buy two types of boxes of toys viz. type I and type II at ₹10 and ₹ 20 per box respectively and sell them at ₹ 14 and ₹ 25 per box respectively. Suppose for reasons of limits on sale he cannot buy more than 150 boxes of type I and 200 boxes of type II but he must buy a minimum of 100 boxes altogether. How many boxes of each type he must buy in order to gain maximum profit assuming that he can sell all the boxes he has bought.

Formulation :

Suppose he buys x boxes of toys of type I and y boxes of type II.

(i) He makes a profit of ₹ 4 per box of type I and ₹ 5 per box of type II. So his total profit would be given by the equation.

$$z = 4x + 5y$$

His objective is to choose non-negative integers x and y such that z is maximum.

(ii) But he cannot make this choice arbitrarily. There are certain constraints to his choice. Because of funds constraint the total cost of buying the toys cannot exceed ₹ 5000. Hence $10x+20y \leq 5000$ or equivalently $x + 2y \leq 500$.

For limits on sale the number of boxes he buyes i.e. x and y must satisfy $x \leq 150$ and $y \leq 200$.

Further and total number of boxes he buys i.e. $x+y$ could be 100 or more.

Thus $x+y \geq 100$.

(iii) Finally the obvious conditions on x and y are given by $x \geq 0$ and $y \geq 0$.

Collecting all the relations which x and y have to satisfy we state below the problem as follows :

Maximize $Z = 4x + 5y$ .... (1)

subject to $\left.\begin{array}{l} x + 2y \leq 500 \\ x \leq 150 \\ y \leq 200 \end{array}\right\}$ .... (2)

$x + y \geq 100$

and $x, y \geq 0.$ .... (3)

From the above example we observe the following :

(i) There is a linear function (i) representing profit which we want to maximize. Sometimes we need to minimise a function, such as cost of productin or time for completion of a job or number of workers needed for certain work and so on.So our objective is to optimise (i.e. maximise or minimise) certain linear function which we shall call the 'objective function'. The objective function may be linear or non linear but we shall consider only linear functions (Hence the term 'linear' in LPP)

(ii) The variables x and y involved in the objective function are known as 'decision variables'.

(iii) There are certain inequations / equations (such as (2)) which need to be satisfied.

We call them 'constraints'.

(iv) Another requirement is that te decision variables x and y must be non-negative $(\text{condition } (3))$

We now define the general LPP.

Definition.

A general linear programming problem LPP is to obtain $x_1, x_2, x_3, ..., x_n$ so as to

Optimize

$$Z = c_1x_1 + c_2x_2 + c_3x_3 + ... + c_nx_n \qquad .... \quad (A)$$

subject to

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + ... \; ... + a_{1n}x_n \leq (\text{or} \geq)\ b_1 \\ a_{21}x_1 + a_{22}x_2 + ... \; ... + a_{2n}x_n \leq (\text{or} \geq)\ b_2 \\ a_{m1}x_1 + a_{m2}x_2 + ... \; ... + a_{mn}x_n \leq (\text{or} \geq)\ b_m \end{array}\right\} \quad .... \quad (B)$$

where $x_1, x_2, ..., x_n \geq ...0 \qquad .... \quad (C)$

and a_{ij}, b_i, c_j with $i = 1, 2, ..., m; j = 1,2, ... n$ are real constants.

In the LPP given above, the function Z in (A) is called the **objective function.** The variables $x_1\ x_3, ... x_n$ are called **decision variables.** The constants $c_1\ c_2, ... c_n$ are called **cost co-efficients.** The inequalities in (B) are called **constraints.** The restrictions in (C) are called **non-negative restrictions.** The solutions which satisfy all the constraints in (B) and the non-negative restrictions in (C) are called **feasible solutions.**

The LPP involves three basic elements :

1. Decision variables whose values we seek to determine,
2. Objective (goal) that we aim to optimize,
3. Constraints and non-negative restrictions that the variables need to satisfy.

3.3 Types of Linear Programming Problems.

As we have already discussed we come across different types of problems which we need solve depending on the objective functions and the constraints. Here we discuss a few important types of LPPs. before learning how to formulate them.

(i) Manufacturing Problem. A manufacturer produces different items so as to maximise his profit. He has to determine the number of units of products he must produce while satisfying a number of constraints because each unit of product requires availability of some amount of raw material, certain manpower, certain machine hours etc.

(ii) Diet Problem. Suppose a person is advised to take vitamins/nutrients of two or more types. The vitamins/nutrients are available in different proportions in different type of foods. If the person has to take a minimum amount of the vitamins/nutrients then the problem is to determine appropriate quantity of food of each type so that cost of food is kept at the minimum.

(iii) Allocation Problem. In this type of problem one has to allocate different resources/tasks to different units / persons depending on the nature of the gain or outcome.

(iv) Transportation Problem. These problems involve transporting materials from sources to destinations for sale or distribution of products or collection of raw materials etc. Here the aim is to use various options for transportation such as distance, time etc so as to keep the cost of transportation to a minimum.

3.4 Formulation of LPP

We have seen in Example-1 how to formulate a simple problem. We take up some more examples to show the method of formulating LPPs.

Example- 2

A firm manufactures two products A and B on which the profits earned per unit are ₹ 13 and ₹ 12 respectively. Each product is processed on three machines M_1, M_2 and M_3. Below is the required processing times in minutes for each machine on each product.

		Product	
	-	A	B
Machine	M_1	4	3
	M_2	2	2
	M_3	3	4

The machines M_1, M_2 and M_3 are available for not more than 9 hours 10 minutes, 10 hours and 8 hours 20 minutes respectively on any working day. Formulate the LPP to find the number of products to be manufactured so as to get maximum profit.

Solution :

Step (i). Let the number of units of products A and B to the produced for optimum profit be x_1 and x_2 respectively.

Step (ii). The objective function is the function that determines the profit (to be maximised) given by

$$Z = 13x_1 + 12x_2$$

Step (iii). The constraints are on the time available for each machine. The machine M_1 takes $4x_1$ minutes to produce x_1 units of A and $3x_2$ minutes to produce x_2 units of B. Hence the restriction of M_1 is given by

$$4x_1 + 3x_2 \le 550$$

Similarly restrictions on machine time of M2 and M3 give.

$$2x_1 + 2x_2 \le 600$$

and $3x_1 + 4x_2 \le 500$

Step (iv). The non-negativity restrictions demand that

$$x_1 \ge 0 \text{ and } x_2 \ge 0.$$

Hence the LPP is formulated as :

Maximise $Z = 13x_1 + 12x_2$

Subject to $4x_1 + 3x_2 \le 550$

$2x_1 + 2x_2 \le 600$

$3x_1 + 4x_2 \le 500$

and $x_1, x_2 \ge 0.$

Example - 3.

An animal feed company must produce 200 kg of a mixture consisting of ingredients A and B. The ingredient A costs ₹ 3 per kg. and B costs ₹ 5 per kg. No more than 80 kg of A can be used and at least 60 kg. of B must be used. Formulate the problem to minimise the cost of mixture.

Solution -

(i) Suppose the company uses x kg. of A and y kg. of B to prepare the mixutre. So the cost or preparation of mixutre is $3x + 5y$ which is to be minimised. Hence the objective function is

$$Z = 3x + 5y$$

(ii) Since quantity of mixture must be 200 kg one condition is

$$x + y \ge 200$$

Again the constraints on the amount of ingredients to be used. requires the following conditions to be satisfied.

$x \le 80$ and $y \ge 60$ .... (3)

(iii) The non negativity conditions are $x \ge 0$ and $y \ge 0$ .... (4)

Hence the LPP is given by

minimize Z= 3x+5y

subject to $x+y \geq 200$

$x \leq 80$

$y \geq 60$

and $x, y \geq 0$.

Exercise - 3 (a)

1. A merchant sells two models X and Y of TV at ₹25000 and ₹50000 per set respectively. He gets a profit of ₹1500 on model X and ₹2000 on model Y. The sales cannot exceed 20 sets in a month. If he cannot invest more than 6 lakh rupees, formulate the problem of determining the number of sets of each type he must keep in stock for maximum profit.

2. A company manufactures and sells two models of lamps L_1 and L_2, the profit being ₹ 15 and ₹10 respectively. The process involves two workers W_1 and W_2 who are available for this kind of work 100 hours and 80 hours per month respectively, W_1 assembles L_1 in 20 and L_2 in 30 minutes. W_2 paints L_1 in 20 and L_2 in 30 minutes. W_2 paints L_1 in 20 and L_2 in 10 minutes. Assuming that all lamps made can be sold, formulate the LPP for determining the productions figures for maximum profit.

3. A factory uses three different resources for the manufacture of two different products, 20 units of the resource A, 12 units of B and 16 units of C being available. One unit of the first product requires 2,2 and 4 units of the resources and one unit of the second product requires 4, 2 and 0 units of the resources taken in order. It is known that the first product gives a profit of ₹20 per unit and the second ₹30 per unit. Formulate the LPP so as to earn maximum profit.

4. A man plans to start a poultry farm by investing at most Rs. 3000. He can buy old hens for ₹80 each and young ones for ₹140 each, but he cannot house more than 30 hens. Old hens lay 4 eggs per week and young ones lay 5 eggs per week, each egg being sold at ₹ 5. It costs ₹5 to feed an old hen and ₹8 to feed a young hen per week. Formulate his problem determining the number of hens of each type he should buy so as to earn a profit of more than ₹300 per week.

5. An agro-based company produces tomato souce and tomato jelly. The quantity of material, machine hour, labor (man-hour) required to produce one unit of each product and the availability of raw material one given in the following table :

	Sauce	Jelly	availability
Man-hour	3	2	10
Machine hour	1	2.5	7.5
Raw material	1	1.2	4.2

A assume that one unit of souce and one unit of jelly each yield a profit of ₹2 and ₹4 respectively. Formulate the LPP so as to yield maximum profit.

6. **(Allocation Problem.)** A farmer has 5 acres of land on which he wishes to grow two crops X and Y. He has to use 4 cart loads and 2 cart loads of manure per acre for crops X and Y respectively. But not more than 18 cart loads of manure is available. Other expenses are ₹200 and ₹500 per acre for the crops X and Y respectively. He estimates profit from crops X and Y at the rates ₹1000 and ₹800 per acre respectively. Formulate the LPP as to how much land he should allocate to each crop for maximum profit.

7. **(Transportation Problem)** A company has two factories at locations X and Y. He has to deliver the products from these factories to depots located at three places A, B and C. The production capacities at X and Y are respectively 12 and 10 units and the requirements at the depots are 8, 8 and 6 units respectively. The cost of transportation from the factories to the depots per unit of the product is given below.

(Cost in Rs.)

	To →	A	B	C
From	X	210	160	250
	Y	170	180	140

The company has to determine how many units of product should be transported from each factory to each depot so that the cost of transportation is minimum. Formulate this LPP.

8. **(Diet Problem)** Two types of food X and Y are mixed to prepare a mixture in such a way that the mixture contains at least 10 units of vitamin A, 12 units of vitamin B and 8 units of vitamin C. These vitamins are available in one kg. of food as per the table given below.

	Vitamins		
food	A	B	C
X	1	2	3
Y	2	2	1

One kg. of food X cost ₹16 and one kg. of food Y costs ₹20. Formulate the LPP so as to determine the least cost of the mixture containing the required amount of vitamins.

9. Special purpose coins each weighing 10gms are to be manufactured using two basic metals M_1 and M_2 and a mix of other metals M_3. M_1, M_2 and M_3 cost ₹500, ₹800 and ₹50 per gram respectively. The strength of a coin demands that not more than 7gm. of M_1 and a minimum of 3 gm of M_2 should be used. The amount of M_3 in each coin is maintained at 25% of that of M_1. Since the demand for the coin is related to its price, formulate the LPP to find the minimum cost of a coin.

10. A company produces three types of cloth A, B and C. Three kinds of wool, say red, green and blue are required for the cloth. One unit length of type A cloth needs 2 meters of red and 3 meters of blue wool; one unit length of type B cloth needs 3 meters of red, 2 meters of green and 2 meters of blue wool and one unit length of type C cloth needs 5 meters of green and 4 meters of blue wool. The firm has a stock of only 80 meters of red, 100 meters of green and 150 meters of blue wool. Assuming that income obtained from one unit length of cloth is ₹30, ₹50 and ₹40 of types A, B and C respectively, formulate the LPP so as to maximize income.

11. A person wants to decide the constituents of a diet which will fulfil his daily requirements of proteins, fats and carbohytrates at minimum cost. The choice is to be made from three different types of food. The yields per unit of these foods are given in the following table.

food	yield / unit			cost/unit
	Protein	Fat	Carbohydrate	
f_1	3	2	6	45
f_2	4	2	3	40
f_3	8	7	7	85
Minimum requirement	1000	200	800	

Formulate the LPP.

3.5 Graphical Solution of LPP.

We now discuss methods of solving an LPP. When the number of decision variables is limited to two, we can use graphical method. Other methods are available for larger number of variables but they are beyond our scope of discussion.

You already know that in two dimensional geometry given an inequation in any one of the forms $ax+by \leq c$, $ax+by \geq c$, $ax + by < c$ or $ax+by > c$, the set of points (x,y) in R^2 satisfying the inequality determines a region which is called the corresponding solution-region. Taking into account all the constraints in an LPP (See inequalities (B) and (C) of art 3.2) we obtain a common region that is the intersection of the solution-regions of each of the inequalities. Every point of this common region satisfies the whole set of constraints including the non-negativity restrictions. This region is called the **feasible region (F.R)** of the L.P.P. Every point of the feasible region is a **feasible solution** of the LPP. All points which are outside the feasible region do not satisfy all the constraints and the nonnegativity restrictions simultaneously. They are called **infeasible solutions** and the region to which they belong is called **infeasible region**. Note that the F.R. is either a region bounded by a convex polygon or anunbounded region (Sec remark (i).

Out of infinitely many points of the feasible region we have to determine those points at which the optimum value of the objective function occurs. This is done as per the following important theorem in linear programming which we state below without proof.

Theorem. If the feasible region for an LPP is bounded by a convex polygon (convex polyhedron, in case of more than two variables) then

(i) at least one of the extreme points (vertices or corner points-where two sides if a polygon meet) gives an optional solution.

and (ii) the objective function has both a maximum and a minimum value each of which occurs at an extreme point of the F.R.

Remark (i) A feasible region is bounded if we can find a constant k such that the distance between any two points of the region cannot exceed k. A region that is not bounded is called unbounded.

(ii) If the F.R. is unbounded then the optimal value of the objective function may or may not exist and if it exists then it must occur at one of the extreme points.

(iii) If an optional value occurs at two different points then it also occurs at each point of the line segment joining these two points.

3.5.1 Working procedure to solue LPP graphically.

Step-1. Taking all inequations of the constaints as equations, draw lines represented by each equation and considering the inequalities of the constraint inequations complete the feasible region.

Step-2. Determine the vertices of the fearible region either by inspection or by & solving the two equations of the intersecting lines.

Step-3. Eraluate the objective function Z= ax+by at each vertex.

case (i) **F.R. is bounded:** The verlex which gives the optional value (maximum or minimum) of Z gives the desired optional solution to the LPP.

case (ii) **F.R. is unbounded:** When M is the maximum value of Z at a vertex Vmax, determine the open half plane corresponding to the inequation $ax + by > M$. If this open half plane has no points in common with the F.R. then M is the mqximum value of Z and the point Vmax gives the desired solution. Otherwise Z has no maximum value.

Similarly consider the open half plane $ax+by < m$ when m is the minimum value of Z at the vertex Vmin. If this half plane has no point common with the F.R. then m is the minimum value of Z and Vmin gives the desired solution. Otherwise Z has no minimum value.

Note. To determine the open half plane in either case (maximum or minimum), draw a dotted line through the corresponding vertex (Vmax or Vmin) parallel to the line $Z=ax+by=0$. Then find the side of the line depending on the inequation $ax+by>M$ or $ax+by<m$.

We now consider some examples to illustrate the above ideas.

Example- 4. Solve the following LPP graphically.

Maximize $Z = 3x+4y$

Subject to

$$4x + 2y \leq 80$$
$$2x + 5y \leq 180$$
$$x, y \geq 0$$

Solution :

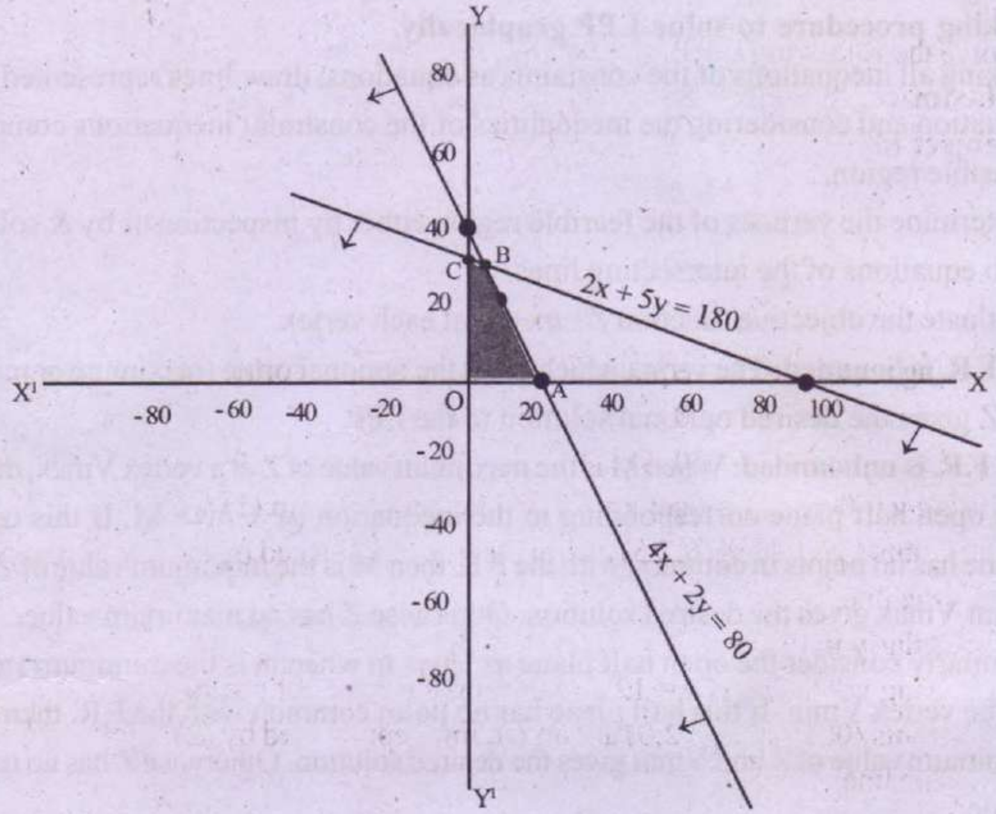
Step-1: Taking inequations of the constraints as equations, draw the straight lines represented by $4x + 2y = 80$ and $2x + 5y = 180$ using Table 1 and Table 2.

Table - 1

x	20	0	10
y	0	40	20

Table - 2

x	90	0	20
y	0	36	28



(Fig. 5)

Step-2 Origin (0,0) satisfies both inequalities

Considering the set of point (x,y) which satisfy the in equations (as shown by the arrows on the lines drawn) along with the condition $x,y \geq 0$ which depicts the first quadrant, we obtain the feasible region.

The the vertices of the feasible region are O(0,0), A (20,0), B (2.5,35) and C (0,36).

Step-3 The values of the objective function at these points are

Z (0) = 0, Z (A) = 60, Z (B) = 147.5 Z (C) = 144.

Thus the maximum value of Z is 147.5 and it occurs at B (2.5, 35), Hence the solution of this LPP is

$$x = 2.5,\ y = 35 \text{ and } Z_{max} = 147.5$$

Note: The coordinates of the vertices O,A,B and C can be obtained by finding the points of inter section of corresponding lines after solving the equations of lines take in pairs. The small arrow marks drawn on a line as in the figure above, show the location of the feasible region on the appropriate side of the line. Refer to the discussion on the Solution Region (SR) in the chapter on linear inequalities in Vol-I to refresh your knowledge in this regard.

Example- 5

Solve the following L.P.P. graphically

Maximze $Z = 40x + 88y$

subject to

$$2x + 8y \leq 60$$
$$5x + 2y \leq 60$$
$$x, y \geq 0$$

Solution :

Step-1 Treating inequations as equations, the constraints become

$$2x + 8y = 60 \text{ ....(1)}$$
$$5x + 2y = 60 \text{ ....(2)}$$

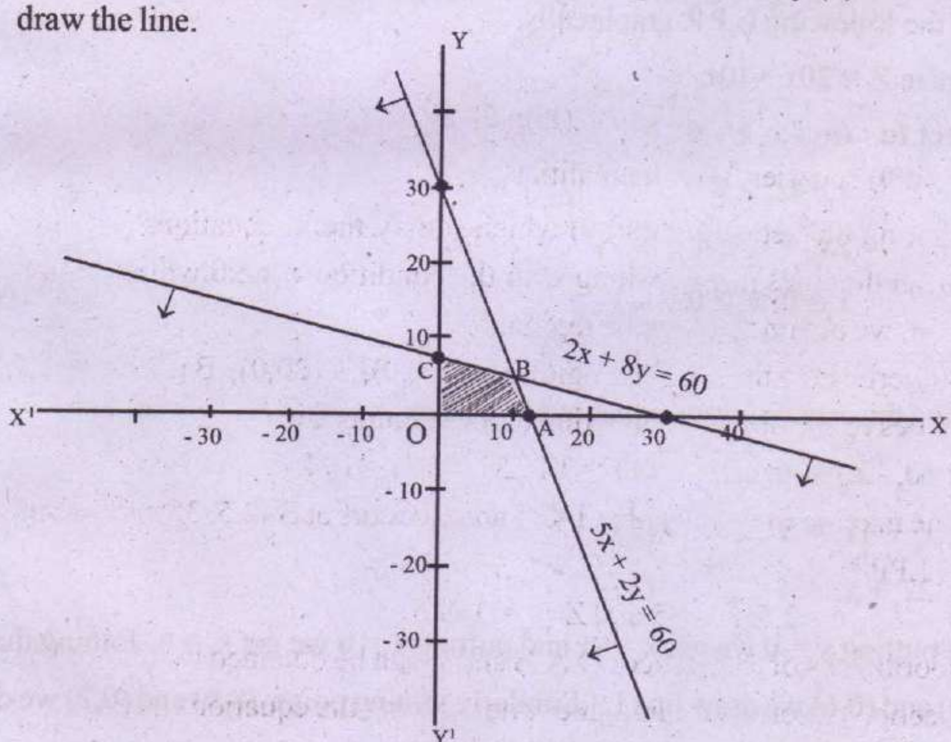
Putting $y = 0$ in (1) we get $x = 30$ and Putting $x = 0$ in (1) we get $y = 7.5$

The points (30, 0) and (0,7.5) are on the line represented by (1). Plot these points and draw the line.

Similarly in (2),

putting $y = 0$ we get $x = 12$ and putting $x = 0$ we get $y = 30$

The points (0, 30) and (12,0) are on the line represented by (2). Plot these points and draw the line.



Step-2 Since origin (0, 0) satisfies the inequation $2x + 8y \leq 60$ and $5x + 2y \leq 60$, the feasible region is below the line $2x + 8y = 60$ and to left of the line $5x + 2y = 60$ Due to the non-negative restrictions $x \geq 0$ and $y \geq 0$, the feasible region is in the first quadrant.

So the shaded region OABC is the feasible region.

In this figure, the point B is the point of intersection of the straight lines

$2x + 8y = 60$ and $5x + 2y = 60$

Solving these equations we get $x = 10, y = 5$ Thus the vertices of the feasible region are O (0, 0), A (12, 0), B (10, 5), C (0, 7.5)

Step-3 The values of the objective function Z at these vertices are

$Z(0) = 40 \times 0 + 88 \times 0 = 0$

$Z(A) = 40 \times 12 + 88 \times 0 = 480$

$Z(B) = 40 \times 10 + 88 \times 5 = 840$

$Z(C) = 40 \times 0 + 88 \times 7.5 = 660$

Thus the maximum value of Z is 840. and it occurs at B (10,5). Hence the solution of this LPP is

$x = 10, y = 5$, and $Z_{max} = 840$

Example- 6

Solve the following L.P.P. graphically.

Optimize $Z = 20x_1 + 40x_2$

Subject to $6x_1 - x_2 \geq -6$

$x_1 + 4x_2 \geq 8$

$2x_1 + x_2 \geq 4$

$x_1 \geq 0; x_2 \geq 0$

Solution

Step-1 Treating the constraints as equations, we have

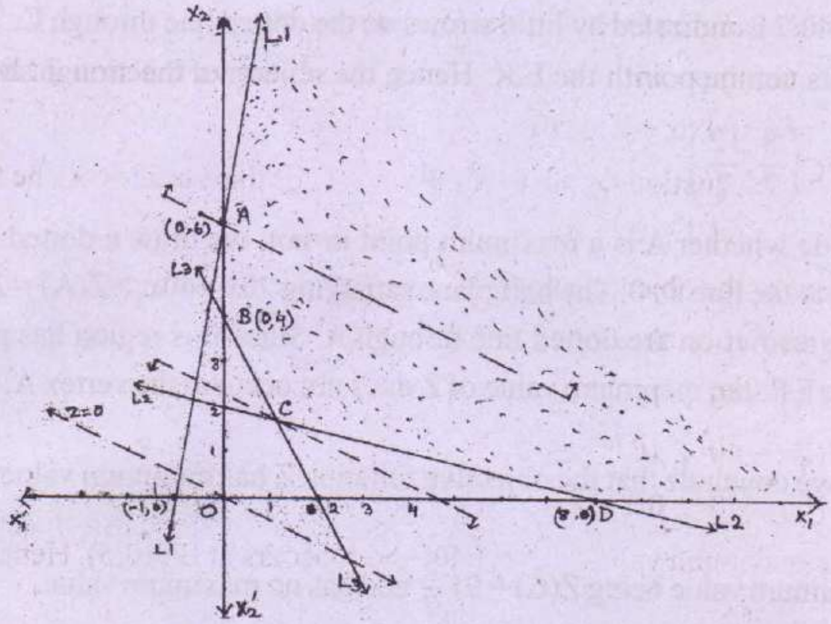
L_1: $6x_1 - x_2 = -6$ .... (1)

L_2: $x_1 + 4x_2 = 8$ .... (2)

L_3: $2x_1 + x_2 = 4$ .... (3)

In (1) putting $x_2 = 0$ we get $x_1 = -1$ and putting $x_1 = 0$ we get $x_2 = 6$. Joining the points (-1, 0) and (0,6) we draw line L_1. Similarly joining points (8,0) and (0,2) we draw line L_2 and by joining points (2,0) and (0,4), line L_3.

The origin (0,0) does not satisfy any of the constraints. Moreover x1 and x2 are non negative. Hence the shaded region ABCD in the first quadrant is the feasible region.



Step-2 The vertices of the fearible region are seen to be the points A, B, C, and D. We have, by inspection, obtained A = (0,6), B=(0,4) and D=(8,0). To obtain C, we solve L_2=0 and L_3= 0. Thus

$$x_1 + 4x_2 - 8 = 0$$

and $$2x_1 + x_2 - 4 = 0$$

$$\Rightarrow \frac{x_1}{-16+8} = \frac{x_2}{-16+} = \frac{1}{1-8} \text{ (by cross Multiplication)}$$

$$\Rightarrow x_1 = \frac{8}{7} \text{ and } x_2 = \frac{12}{7} \Rightarrow C = \left(\frac{8}{7}, \frac{12}{7}\right).$$

Step-3 Now Z= $20x_1 + 40x_2$

$\Rightarrow$ Z(A) = 20×0 + 40×6 = 240 (max.)

Z(B) = 20×0 + 40×4 =160

$Z(C) = 20\times\frac{8}{7} + 40 \times \frac{12}{7} = 640/7 = 91\frac{3}{7}$ (min.)

Z(D) = 20×8 + 40×0 = 160

Here we cannot immediately decide that Z(A) and Z(C) are the maximum and minimum values of Z because the F.R. is **unbounded**. So we draw a dotted line through the origin representing Z=0. Since Z(C) is a possible case of minimum value, we draw a dotted line through C parallel to the line Z=0. The half plane satisfying $20x_1+40x_2 <$

z(c) = 640/7 is indicated by little arrows on the dotted line through C. This region has no points common with the F.R. Hence the objective function Z attains minimum value at $C\left(\frac{8}{7},\frac{12}{7}\right)$.

To decide whether A is a maximum point or not, we draw a dotted line through A parallel to the line Z=0. The half plane satisfying $20x_1+40x_2 > Z(A) = 240$ is indicated by little arrows on the dotted line through A. Since this region has points common with the F.R. the maximum value of Z does not occur at the vertex A.

Hence we conclude that the objective function Z has minimum value of $C=\left(\frac{8}{7},\frac{12}{7}\right)$

the minimum value being $Z(C) = 91\frac{3}{7}$ and has no maximum value.

Example- 7

Solve the following L.P.P. graphically.

Minimize $Z = 6x + 9y$

Subject to $x + 12y \le 65$

$7x - 2y \le 25$

$2x + 3y \ge 10$

$x, y \ge 0$

Solution

Step-1 Treating the given inequations as equations the constraints become

$x + 12y = 65$ .... (1)

$7x - 2y = 25$ .... (2)

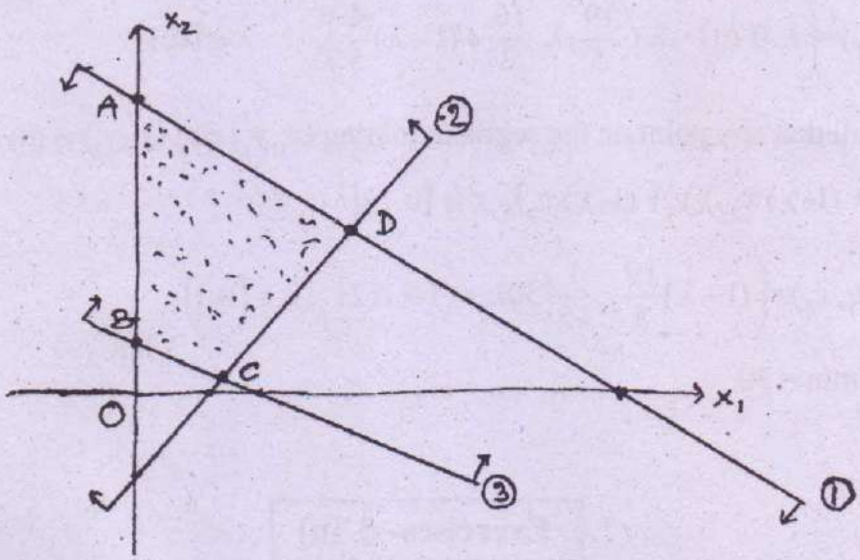
and $2x + 3y = 10$ .... (3)

In equation (1), by putting y = 0 we get $x = 65$ and putting $x = 0$ we get $y = \frac{65}{12}$.

So the points (65,0) and $\left(0, \frac{65}{12}\right)$ are on the line represented by (1). Joining these points we draw the line.

Similarly in equation (2) we get points $\left(\frac{25}{7},0\right)$ and $\left(0,-\frac{25}{2}\right)$ on the line represented by (2). Joining these points we draw the line.

Again we get points (5,0) and $\left(0,\ \frac{10}{3}\right)$ on the line represented by (3) and joining them we draw the line.



Step-2 Since the origin (0,0) does not satisfy the inequation $2x + 3y \geq 10$ and satisfies the other two inequations, the feasible region ABCD is shaded as shown lying in the first quadrant. The point C is obtained by Solving the equations (2) and (3) whence $C=\left(\frac{19}{5},\frac{4}{5}\right)$, The point D is obtained by solving equations (1) and (2) whence D=(5,5). The points A and B are points of intersection of lines (1) and (3) with x=0. Thus the vertices of the feasible region are $A\left(0,\frac{65}{12}\right)$, $B\left(0,\frac{10}{3}\right)$, $C\left(\frac{19}{5},\frac{4}{5}\right)$ and D(5,5).

Step-3 The values of the objective function Z are

$Z(A) = 6\times 0+9\times\frac{65}{12}=\frac{195}{4}$

$Z(B) = 6\times 0+9\times\frac{10}{3}=30$

$Z(C) = 6\times\frac{19}{5}+9\times\frac{4}{5}=30$

$Z(D) = 6\times 5+9\times 5=75$

Here the feasible region is bounded.

Hence Zmin = 30 the minimum value occurs at be and also at C. Hence the minimum value occurs at every print of the segment $\overline{BC}$. (see remark (iii) of section 3.5) The solution of the LPP is given by

$$(x_1, x_2) = \left(\lambda.0 + (1-\lambda)\cdot\frac{19}{5}, \lambda\cdot\frac{10}{3} + (1-\lambda)\frac{4}{5}\right), \qquad \lambda \in [0,1]$$

[Recall that any point on the segment joining (x_1, y_1) and (x_2, y_2) is given by $(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2), \lambda \in [0,1]$]

i.e. $(x_1, x_2) = \left((1-\lambda)\frac{19}{5}, \frac{1}{15}[50\lambda + (1-\lambda)12]\right), \lambda \in [0,1]$

and Zmin = 30.

Exercises- 3 (b)

Solve the following L.P.P.s graphically.

1. Maximize $Z = 5x_1 + 6x_2$
   Subject to $2x_1 + 3x_2 \le 6$
   $x_1, x_2 \ge 0$
2. Minimize $Z = 6x_1 + 7x_2$
   Subject to $x_1 + 2x_2 \ge 4$
   $x_1, x_2 \ge 0$
3. Maximize $Z = 20x_1 + 40x_2$
   Subject to $x_1 + x_2 \le 1$
   $6x_1 + 2x_2 \le 3$
   $x_1, x_2 \ge 0$
4. Minimize $Z = 30x_1 + 45x_2$
   Subject to $2x_1 + 6x_2 \ge 4$
   $5x_1 + 2x_2 \ge 5$
   $x_1, x_2 \ge 0$
5. Maximize $Z = 3x_1 + 2x_2$
   Subject to $-2x_1 + x_2 \le 1$

$x_1 \leq 2$

$x_1 + x_2 \leq 3$

$x_1, x_2 \geq 0$

6. Maximize $Z = 50x_1 + 60x_2$

   Subject to $x_1 + x_2 \leq 5$

   $x_1 + 2x_2 \geq 4$

   $x_1, x_2 \geq 0$

7. Maximize $Z = 5x_1 + 7x_2$

   Subject to $x_1 + x_2 \leq 4$

   $5x_1 + 8x_2 \leq 30$

   $10x_1 + 7x_2 \leq 35$

   $x_1, x_2 \geq 0$

8. Maximize $Z = 14x_1 - 4x_2$

   Subject to $x_1 + 12x_2 \leq 65$

   $7x_1 - 2x_2 \leq 25$

   $2x_1 + 3x_2 \geq 10$

   $x_1, x_2 \geq 0$

   Also find two other points which maximize Z.

9. Maximize $Z = 10x_1 + 12x_2 + 8x_3$

   Subject to $x_1 + 2x_2 \leq 30$

   $5x_1 - 7x_3 \geq 12$

   $x_1 + x_2 + x_3 = 20$

   $x_1, x_2 \geq 0$

   [Hint : Eliminate x_3 from all expressions using the given equation in the set of constraints, so that it becomes an LPP in two variables]

10. Minimize $Z = 20x_1 + 10x_2$

    Subject to $x_1 + 2x_2 \leq 40$

    $3x_1 + x_2 \geq 30$

    $4x_1 + 3x_2 \geq 60$

    $x_1, x_2 \geq 0$

11. Maximize $Z = 4x_1 + 3x_2$

Subject to $x_1 + x_2 \le 50$

$x_1 + 2x_2 \le 80$

$2x_1 + x_2 \ge 20$

$x_1, x_2 \ge 0$

12. Optimize $Z = 5x_1 + 25x_2$

Subject to $-0.5x_1 + x_2 \le 2$

$x_1 + x_2 \ge 2$

$-x_1 + 5x_2 \ge 5$

$x_1, x_2 \ge 0$

13. Optimize $Z = 5x_1 + 2x_2$

Subject to $-0.5x_1 + x_2 \le 2$

$x_1 + x_2 \ge 2$

$-x_1 + 5x_2 \ge 5$

$x_1, x_2 \ge 0$

14. Optimize $Z = -10x_1 + 2x_2$

Subject to $-x_1 + x_2 \ge -1$

$x_1 + x_2 \le 6$

$x_2 \le 5$

$x_1, x_2 \ge 0$

15. Solve the L.P.P.s obtained in Exercise 3(a) Q1 to Q9 by graphical method.

CHAPTER 4

Matrices

It is no paradox to say that in our most theoretical moods we may be nearest to our most practical applications.

- A.N. Whitehead

4.0 Introduction

Brevity is the essence of mathematics. Many a time, in our day to day life, we come across situations which are similar in nature. In stead of giving them individual attention, we would like to deal with them as a collective and systemetic whole. How much interesting it would be if we were able to express a number of equations as a single one and come up ready with their individual solutions in a compact form !

Not only equations, there are several instances from various fields of our activities as diverse as physics, engineering, electronics, finance, business management and over and above all, the software 'matlab' where we need compactification of ideas for a systematic dealing with facts and figures.

4.1 Matrix, what it is :

Suppose there are to 4 unit tests in science, mathematics and literature and your scores in repective subjects and tests are given as under :

Tests/Scores	Science	Mathematics	Literature
I	80	90	76
II	82	89	74
III	88	91	77
IV	89	95	78

From the above we can carve out a rectangular array :

80	90	76
82	89	74
88	91	77
89	95	78

with the understanding that a column gives your scores in a particular subject in the defferent tests whereas a row gives your scores in a particular test in different subjects.

An array as such is an example of a matrix.

Matrix and its order

A matrix is a rectangular array of numbers (any system of numbers), arranged in rows and columns. **If there are *m* rows, and *n* columns in a matrix, it is called an '*m* by *n*' matrix, or a matrix of order $m \times n$.**

Note that m and n in 'm by n' or 'm×n' respectively dinote the number of rows and number of columns in a matrix.

Let us consider a set of numbers, a_{ij} ($i = 1, 2, ...m; j = 1, 2, 3 ....n$). We can arrange them in a rectangular array as follows (denoted by A and described by $[a_{ij}]_{mxn}$). So .

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & a_{m3} & a_{m4} & \cdots & a_{mn} \end{pmatrix}$$

This array is enclosed by parentheses (Alternatively this can also be enclosed by brackets.) $A = [a_{ij}]_{mxn}$ is called a matrix of order mxn.

When $m = n$, we have a square matrix of order nxn (or simply n).

Each a_{ij} is called an element of the matrix and it is the j^{th} element of the i^{th} row.

It can be observed that there are mn elements in a matrix of order $m \times n$.

4.2 Some Definitions:

1. Zero Matrix : If all the elements of a matrix are zero, it is called a **null matrix** or **zero matrix.** The null matrix of order $m \times n$ is denoted by $\mathbf{0}_{mxn}$. If the elements of a matrix are not all zero, it is called a non- zero matrix.

2. Transpose :

The **transpose** of a matrix A is the matrix, obtained from A by changing its rows into columns, and columns into rows. It is denoted by A′ or A^T.

Hence if A is' of order $m \times n$, then A^T or A′ is of order $n \times m$.

The transpose of $\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{bmatrix}$ (2 × 3 matrix) is $\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix}$ (3 × 2 matrix)

Transpose of a transpose :

It is evident that the transpose of the transpose of a matrix is the given matrix itself, that is

$[A^T]^T = A$.

If $A = [a_{ij}]_{m\times n}$, then by definition, $A^T = [a'_{ij}]_{n\times m}$ **where** $a'_{ij} = a_{ji}$.

3. Row Matrix :

A matrix with a single row is called a **row matrix.**

Thus $[a_1 b_1\ c_1]$ is a row matrix. It is a 1×3 matrix.

$[a_1 a_2 ...... a_n]$ is a $1\times n$ (row) matrix.

4. Column Matrix :

A matrix with a single column is called a **column matrix.**

$\begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}$ is a column martix

It is a 3×1 matrix.

In general $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ is an $n\times 1$ column matrix.

N.B. : Transpose of a row matrix is a column matrix and transpose of a column matrix is a row matrix.

5. Diagonal Matrix :

A square martix of which the nondiagonal elements are all zero, is called diagonal matrix.

So $\begin{bmatrix} a_{11} & 0 & 0 & ... & 0 \\ 0 & a_{22} & 0 & ... & 0 \\ 0 & 0 & a_{33}. & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & a_{nn} \end{bmatrix}$ is a diagonal matrix of order *n*.

The elements $a_{11}, a_{22}, a_{33}, ..., a_{nn}$ comprise the diagonal of the square matrix $[a_{ij}]_{n\times n}$.

Hence $[a_{ij}]_{n\times n}$ **is a diagonal matrix if** $a_{ij} = 0$ **for** $i \neq j$**.**

6.Scalar Matrix :

If the **diagonal elements of the diagonal martix are all equal**, it is called a **scalar matrix**. For example, the matrix $\begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix}$ is a scalar matrix of order 3. For any real or complex number α, $[\alpha]$ is also a scalar matrix of order 1.

N.B. : Scalar matrices are necessarity square matrics. **It may be noted that a square matrix of order n actually means a matrix of order n×n as per definition in the beginning.**

7.Unit Matrix :

If the diagonal elements of a diagonal matrix are all unity, it is called **a unit matrix**.

For example $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is a unit matrix of order 3. A unit matrix is also called **Identity matrix.** A unit matrix of order n is denoted by I_n or by I, when the order is understood.

8. Equality of Matrices :

Two matrices A and B are said to be conformable for equality, if they are of same order.

Two matrices of the same order are said to be equal, if and only if the corresponding elements of the two are equal. So the **matrices $[a_{ij}]_{m\times n}$ and $[b_{ij}]_{m\times n}$ will be equal if and only if $a_{ij} = b_{ij}$** for each pair of values of i and j.

Thus the two matrices $\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{bmatrix}$ and $\begin{bmatrix} p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \end{bmatrix}$ are equal if $a_1 = p_1$, $b_1 = q_1$, $c_1 = r_1$, $a_2 = p_2$, $b_2 = q_2$ and $c_2 = r_2$. For example, the matrices $\begin{bmatrix} 1^3 & 2^3 \\ 3^3 & 4^3 \end{bmatrix}$ and $\begin{bmatrix} 1 & 8 \\ 27 & 64 \end{bmatrix}$ are equal. We write $\begin{bmatrix} 1^3 & 2^3 \\ 3^3 & 4^3 \end{bmatrix} = \begin{bmatrix} 1 & 8 \\ 27 & 64 \end{bmatrix}$.

4.3 Algebra of Matrices (Operations on matrices) :

Addition and Subtraction of Matrices :

Two matrices A and B are said to be conformable for addition if they are of the same order. The sum of two matrices of the same order is the matrix of which the elements are the sum of the corresponding elements of the two matrices. Therefore the sum of the two matrices $A = [a_{ij}]_{m\times n}$, and $B = [b_{ij}]_{m\times n}$ is the matrix $C = [c_{ij}]_{m\times n}$ where $c_{ij} = a_{ij} + b_{ij}$ for all values of i and j.

Thus $\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{bmatrix} + \begin{bmatrix} p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \end{bmatrix} = \begin{bmatrix} a_1+p_1 & b_1+q_1 & c_1+r_1 \\ a_2+p_2 & b_2+q_2 & c_2+r_2 \end{bmatrix}$

For example, $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 7 & 8 & 9 \\ 3 & 6 & 9 \end{bmatrix} = \begin{bmatrix} 8 & 10 & 12 \\ 7 & 11 & 15 \end{bmatrix}$

The difference A – B of two matrices A and B of the same order is defined to be the sum A + (– B) of matrices A and – B. Thus $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 9 & 8 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 6 \\ 2 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 2 & -3 \\ -2 & 5 & 3 \end{bmatrix}$

Additive Identity :

It is evident that $[a_{ij}]_{m\times n} + O_{m\times n} = [a_{ij}]_{m\times n}$, and $[a_{ij}]_{m\times n} - O_{m\times n} = [a_{ij}]_{m\times n}$

$\therefore O_{m\times n}$ is the additive identity with respect to addition on m×n matrices.

Additive Inverse :

The additive inverse of a matrix $A = [a_{ij}]_{m\times n}$ is the matrix $[-a_{ij}]_{m\times n}$;

and is denoted by $-A = \begin{bmatrix} -a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & -a_{22} & \cdots & -a_{2n} \\ -a_{31} & -a_{32} & \cdots & -a_{3n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{m1} & -a_{m2} & \cdots & -a_{mn} \end{bmatrix}$

Commutative law holds good for addition of matrices:

Let $A = [a_{ij}]_{m\times n}$, and $B = [b_{ij}]_{m\times n}$ be two matrices, then $A + B = [\alpha_{ij}]_{m\times n}$ where $\alpha_{ij} = a_{ij} + b_{ij}$, and $B + A = [\beta_{ij}]_{m\times n}$, where $\beta_{ij} = b_{ij} + a_{ij}$. But $a_{ij} + b_{ij} = b_{ij} + a_{ij}$. So $\alpha_{ij} = \beta_{ij}$.

Hence A + B = B + A. Thus **matrix addition is commutative.**

Associative law holds good for addition of matrices :

Let $A = [a_{ij}]_{m\times n}$, $B = [b_{ij}]_{m\times n}$, and $C = [c_{ij}]_{m\times n}$ be three matrices, then $(A+B)+C = [\alpha_{ij}]_{m\times n} + [c_{ij}]_{m\times n}$ (where $\alpha_{ij} = a_{ij} + b_{ij}$) $= [p_{ij}]_{m\times n}$

where $p_{ij} = \alpha_{ij} + c_{ij} = a_{ij} + b_{ij} + c_{ij}$ Again $A + (B + C) = [a_{ij}]_{m\times n} + [r_{ij}]_{m\times n}$. (where $r_{ij} = b_{ij} + c_{ij}$) $= [q_{ij}]_{m\times n}$ where $q_{ij} = a_{ij} + r_{ij} = a_{ij} + b_{ij} + c_{ij} = p_{ij}$.

Therefore $[p_{ij}]_{m\times n} = [q_{ij}]_{m\times n}$ Hence (A + B) + C = A + (B + C).

So **matrix addition is associative.**

Transpose of Sum of Matrices :

It can easily be shown that the transpose of the sum of two matrices is equal to the sum of their transposes, provided that the sum is defined. Thus $[A+B]^T = A^T + B^T$

Multiplication of matrices by a scalar

Let A be a matrix $[a_{ij}]_{m\times n}$ and k be a scalar quantity, i.e. a real or comlex number, then the product kA *or* Ak is a matrix $[b_{ij}]_{m\times n}$ where $b_{ij} = k\,a_{ij}$. Hence $k\begin{bmatrix} a & b & c \\ p & q & r \end{bmatrix} = \begin{bmatrix} ka & kb & kc \\ kp & kq & kr \end{bmatrix}$

Thus $2\times\begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 6 & 8 & 10 \end{bmatrix}$.

If A and B are two m × n matrices, and c and d are scalars, then the following results are obvious.

(i) $c(A+B) = cA + cB$ (ii) $(c+d)A = cA + dA$

Corollary :

The transpose of the product kA is the product of k and the transpose of A.

Thus $[kA]^T = k[A]^T$.

Matrix Multiplication :

If the number of columns of a matrix A is equal to the number of rows of another matrix B, then the matrices A and B are said to be conformable for the product AB and the product AB is said to be defined. We denote the product by A. B or AB.

If $A = [a_{ij}]_{m\times n}$ and $B = [b_{jk}]_{n\times p}$, then we have $AB = [c_{ik}]_{m\times p}$ where $c_{ik} = \sum_{j=1}^{n} a_{ij}\cdot b_{jk}$. In other words, if A is a matrix of order $m\times n$ and B is of order $n\times p$, then AB is a matrix of order $m\times p$, and the element of ith row and kth column of AB is the sum of the products of the elements of the ith row of A, and the corresponding elements of the kth column of B.

$$\text{Hence}\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{bmatrix}\times\begin{bmatrix} p_1 & p_2 \\ q_1 & q_2 \\ r_1 & r_2 \end{bmatrix} = \begin{bmatrix} a_1p_1 + b_1q_1 + c_1r_1 & a_1p_2 + b_1q_2 + c_1r_2 \\ a_2p_1 + b_2q_1 + c_2r_1 & a_2p_2 + b_2q_2 + c_2r_2 \end{bmatrix}$$

$$\text{Similarly}\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix}\times\begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1+4+9 & 0+2+6 \\ 0+2+6 & 0+1+4 \end{bmatrix} = \begin{bmatrix} 14 & 8 \\ 8 & 5 \end{bmatrix}$$

In the product AB, A is called the prefactor and B is called the post factor.

Corollary 1 : If A and B are two matrices such that (A + B) and AB are both defined, then A and B are both square matrices of the same order. Since (A + B) is defined, the matrices A and B must be of the same order, say m × n. Again, since AB is defined, the number of columns of A must be equal to the number of rows of B. Hence $m = n$.

Corollary 2 : If A is an $m \times n$ matrix, and if both AB and BA are defined, then B is an $n \times m$ matrix.

Note 1. Non zero matrices may multiply to a zero matrix

If AB = [0] (sometimes the zero matrix [0] is also written as **0**), we cannot say, as in scalar algebra, that either A or B is a zero matrix. For example, if $A = \begin{bmatrix} 1 & 4 & 0 \\ 2 & 5 & 0 \\ 3 & 6 & 0 \end{bmatrix}$, and $B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 7 & 8 & 9 \end{bmatrix}$

we have $AB = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ which shows that **though the product AB is a zero marix, neither A nor B is a zero matrix.**

In the above example both A and B are 3×3 square matrices. However they may be of different orders which are conformable for multiplication. Consider :

(i) $A = \begin{bmatrix} 0 & 3 & 1 \\ 0 & 1 & 4 \end{bmatrix}$ (2×3 matrix), $B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (3×3 matrix)

Here AB = $\mathbf{0}$, a 2×3 zero matrix.

(ii) $A = \begin{bmatrix} -3 & 2 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 0 \\ 3 & 0 \end{bmatrix}$ (Square matrices of order 2)

Here AB= 0, a 2×2 zero matrix.

Observe that neither A, nor B is a zero matrix.

Note 2. If **AB = AC, we cannot say, as in scalar algebra, that B = C**, even if $A \neq 0$. For example, if $A = \begin{bmatrix} 1 & 4 & 0 \\ 2 & 5 & 0 \\ 3 & 6 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, and $C = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{bmatrix}$ we have $AB = \begin{bmatrix} 7 & 10 & 13 \\ 11 & 14 & 17 \\ 15 & 18 & 21 \end{bmatrix} = AC$

whre as $B \neq C$.

Properties of matrix multiplication

(a) Matrix Multiplication is not commutative in general

In order that the matrices A and B be conformable for the products AB and BA, we take

$A = [a_{ij}]_{m \times n}$, and $B = [b_{ij}]_{n \times m}$. Then the product AB is a matrix of order $m \times m$ while the product BA is a matrix of order $n \times n$. Hence the question of equality of matrices AB and BA does not arise unless $m = n$.

If $m = n$, we have $A = [a_{ij}]_{n \times n}$, and $B = [b_{ij}]_{n \times n}$. Then $AB = [c_{ij}]_{n \times n}$,

where $[c_{ij}] = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}$ and $BA = [d_{ij}]_{n \times n}$, where $d_{ij} = \sum_{k=1}^{n} b_{ik}\, a_{kj}$. Since $c_{ij} \neq d_{ij}$ in general, $AB \neq BA$ in general.

You can also numerically verify the assertion by taking

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 3 \\ 4 & -2 \end{bmatrix}$$

that $AB \neq BA$.

Thus matrix multiplication is noncommutative in general.

Note : The $n \times n$ unit matrix I_n behaves as the multiplicative identity during multiplication with $n \times n$ square matrices, i.e. $A I_n = A = I_n A$ where A is any square matrix of order n.

You can verify this by taking specific examples.

We shall make use of this fact in our subsequent developments.

(b) Matrix multiplication is associative

Let $A = [a_{ij}]_{m \times n}$, $B = [b_{jk}]_{n \times p}$ and $C = [c_{kl}]_{p \times q}$ be three matrices, so that they are conformable for the products AB and BC. Now $AB = [\alpha_{ik}]_{m \times p}$, wehre $\alpha_{ik} = \sum_{j=1}^{n} a_{ij}\ b_{jk}$, and $C = [c_{kl}]_{p \times q}$.

So the marices AB and C are conformable for the product (AB). C; hence $(AB)\,C = [\beta_{il}]_{m \times q}$

where $\beta_{il} = \sum_{k=1}^{p} \alpha_{ik}\, c_{kl}$

$$= \sum_{k=1}^{p} \left[\left(\sum_{j=1}^{n} a_{ij}\, b_{jk} \right) \right] c_{kl} = \sum_{k=1}^{p} \sum_{j=1}^{n} a_{ij}\, b_{jk}\, c_{kl} \quad \text{.................(1)}$$

Similarly, $BC = [\vartheta_{jl}]_{n \times q}$, where $\vartheta_{jl} = \sum_{k=1}^{p} b_{jk}\, c_{kl}$.

So the matrices A and BC are conformable for the product A (BC), and hence A (BC) =

$[\delta_{il}]_{m \times q}$ where $\delta_{il} = \sum_{j=1}^{n} a_{ij}\, \vartheta_{jl} = \sum_{j=1}^{n} a_{ij} \left(\sum_{k=1}^{p} b_{jk} c_{kl} \right)$

$$= \sum_{j=1}^{n} \left(\sum_{k=1}^{p} a_{ij}\, b_{jk}\, c_{kl} \right) = \sum_{k=1}^{p} \sum_{j=1}^{n} a_{ij}\, b_{jk}\, c_{kl} \quad (\because \text{ the number of terms is finite})$$

$= \beta_{il}$ (from (1))

Since $\beta_{il} = \delta_{il}$ for all pairs of i and l ($i = 1, 2, 3, \ldots\ldots m, l = 1, 2, 3, \ldots\ldots q$)

we have $[\beta_{il}]_{m \times q} = [\vartheta_{il}]_{m \times q}$. So (AB) C = A (BC)

Thus matrix multiplication is associative.

You can also numerically verify the result taking suitable examples maintaining conformability for multiplication.

(c) Matrix multiplication is distributive with respect to addition :

Let $A = [a_{ij}]_{m \times n}$, $B = [b_{jk}]_{n \times p}$, and $C = [c_{jk}]_{n \times p}$, be three matrices, so that they are conformable for the sum (B+C), and for the products AB and AC. We have $B + C = [x_{jk}]_{n \times p}$ where $x_{jk} = b_{jk} + c_{jk}$. So the matrices A and B + C are conformable for the product A(B+C).

Again A (B+C) = $[y_{ik}]_{m \times p}$, (where $y_{ik} = \sum_{j=1}^{n} a_{ij} x_{jk}$)

$$= \sum_{j=1}^{n} a_{ij}\,(b_{jk} + c_{jk}) = \sum_{j=1}^{n} a_{ij}\, b_{jk} + \sum_{j=1}^{n} a_{ij}\, c_{jk} \qquad (1)$$

Again $AB = [u_{ik}]_{m \times p}$ (where $u_{ik} = \sum_{j=1}^{n} a_{ij}\, b_{jk}$)

and $AC = [\upsilon_{ik}]_{m \times p}$ (where $\upsilon_{ik} = \sum_{j=1}^{n} a_{ij}\, c_{jk}$)

So the matrices AB and AC are conformable for the addition AB + AC; hence

$AB + AC = [\omega_{ik}]_{m \times p}$, where $\omega_{ik} = u_{ik} + \upsilon_{ik} = \sum_{j=1}^{n} a_{ij}\, b_{jk} + \sum_{j=1}^{n} a_{ij}\, c_{jk} = y_{ik}$ [from (1)]

Since $y_{ik} = \omega_{ik}$ for all pairs of i and k ($i = 1, 2, 3, \ldots\ldots m; k = 1, 2, 3, \ldots\ldots p$),

we have $[y_{ik}]_{m \times p} = [w_{ik}]_{m \times p}$,

So **A (B + C) = AB + AC**

Thus **matrix multiplication is distributive with respect to addition.**

Corollary Assuming the conformability of the operations,

We have **A (B – C) = AB – AC,**

(A + B) (C +D) = A (C + D) + B (C +D) = AC + AD + BC + BD

(d) In matrix multiplication, a scalar matrix behaves like a scalar multiplier.

Let $K = \begin{bmatrix} k & 0 & \dots & 0 \\ 0 & k & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & k \end{bmatrix}_{n \times n}$ be a scalar matrix.

and $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$ be a matrix of order n.

Then $KA = \begin{bmatrix} k & 0 & \dots & 0 \\ 0 & k & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & k \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$

$$= \begin{bmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \dots & \dots & \dots & \dots \\ ka_{n1} & ka_{n2} & \dots & ka_{nn} \end{bmatrix} = k \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

It can easily be seen that KA = AK. In other words, multiplication of a matrix with a scalar matrix is commutative as in case of multiplication by a scalar.

(e) The transpose of the product of two matrices is the product of their transposes, taken in the reverse order.

Let $A = [a_{ij}]_{m \times n}$ and $B = [b_{jk}]_{n \times p}$ be two matrices so that they are conformable of the product AB.

Then $AB = [c_{ik}]_{m \times p}$ where $c_{ik} = \sum_{j=1}^{n} a_{ij}\, b_{jk}$

$\therefore$ Transpose of AB is $[AB]^T = = [c'_{ki}]_{p \times m}$

Again $B^T = [b'_{kj}]_{p \times n}$ where $b'_{kj} = b_{jk}$, and $A^T = [a'_{ji}]_{n \times m}$ where $a'_{ji} = a_{ij}$, so that they are conformable for the product $B^T A^T$.

Then $B^T A^T = [d'_{ki}]_{p \times m}$ where $d'_{ki} = \sum_{j=1}^{n} b'_{kj} a'_{ji} = \sum_{j=1}^{n} b_{jk} a_{ij} = \sum_{j=1}^{n} a_{ij} b_{jk} = c_{ik} = c'_{ki}$

Since $c'_{ki} = d'_{ki}$ for all pairs of k and i ($k = 1, 2, \dots\dots p;\ i = 1, 2, \dots m$)

we have $[c'_{ki}]_{p\times m} = [d'_{ki}]_{p\times m}$, Hence $[AB]^T = B^TA^T$

It is known as the law of reversal for a transpose.

Corollary If A, B, C... K are square matrices of same order, then $[ABC]^T = C^T[AB]^T = C^TB^TA^T$, In general $[ABC\ldots\ldots K]^T = K^T \ldots\ldots C^TB^TA^T$.

Some further definitions based on matrix-multiplication :

i. Orthogonal Matrix :

A square matrix A of order n is said to be orthogonal if $AA' = A'A = I_n$

ii. Idempotent Matrix : A square matrix A is called idempotent matrix if A^2 i.e. $A\times A = A$. For example $A = \begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}$ is idempotent.

iii. Involuntary Matrix : A matrix such that $A^2 = I$ (unit matrix the same order as A) is called an involuntary matrix.

iv. Nilpotent matrix : A square matrix A is called a nilpotent matrix if there exists a positive integer m such that $A^m = A\times A\times\ldots\times A$ (m times) $= \mathbf{0}_{m\times m}$

Illustrative Examples

Example 1 :

Given $A = \begin{bmatrix} 4 & 2 & -1 \\ 3 & -7 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} & 2 & 3 \\ - & 3 & 0 \\ - & 1 & 5 \end{bmatrix}$, Find A + B, A – B, AB and BA wherever it is possible. State reasons for the operations which are not possible.

Here A is a matrix of order 2×3, and B is a matrix of order 3×2. Since the matrices are not of the same order, the operations (A + B) and (A – B) are not possible.

Since the number of columns of A is equal to the number of rows of B, AB is defined; and

$$AB = \begin{bmatrix} 4 & 2 & -1 \\ 3 & -7 & 1 \end{bmatrix} \times \begin{bmatrix} 2 & 3 \\ -3 & 0 \\ -1 & 5 \end{bmatrix} = \begin{bmatrix} 8-6+1 & 12+0-5 \\ 6+21-1 & 9+0+5 \end{bmatrix} = \begin{bmatrix} 3 & 7 \\ 26 & 14 \end{bmatrix}$$

Similarly, BA is defined and $BA = \begin{bmatrix} 2 & 3 \\ -3 & 0 \\ -1 & 5 \end{bmatrix} \times \begin{bmatrix} 4 & 2 & -1 \\ 3 & -7 & 1 \end{bmatrix}$

$$= \begin{bmatrix} 8+9 & 4-21 & -2+3 \\ -12+0 & -6+0 & 3+0 \\ -4+15 & -2-35 & 1+5 \end{bmatrix} = \begin{bmatrix} 17 & -17 & 1 \\ -12 & -6 & 3 \\ 11 & -37 & 6 \end{bmatrix}$$

Example 2 :

Find the values of x,y, z, t for which the matrices $\begin{bmatrix} x+y & y-z \\ 5-t & 7+x \end{bmatrix}$ and $\begin{bmatrix} t-x & z-t \\ z-y & x+z+t \end{bmatrix}$ may be equal.

If $\begin{bmatrix} x+y & y-z \\ 5-t & 7+x \end{bmatrix} = \begin{bmatrix} t-x & z-t \\ z-y & x+z+t \end{bmatrix}$, then $x+y=t-x$, $y-z=z-t$, $5-t=z-y$ and $7+x=x+z+t$. Solving these four equations we get $x=1$, $y=2$, $z=3$ and $t=4$.

Example 3 :

Determine the matrices A and B where $A+2B = \begin{bmatrix} 1 & 2 & 0 \\ 6 & -3 & 3 \\ -5 & 3 & 1 \end{bmatrix}$ and $2A-B = \begin{bmatrix} 2 & -1 & 5 \\ 2 & -1 & 6 \\ 0 & 1 & 2 \end{bmatrix}$

$$2(2A-B)+A+2B = \begin{bmatrix} 4 & -2 & 10 \\ 4 & -2 & 12 \\ 0 & 2 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 0 \\ 6 & -3 & 3 \\ -5 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 10 \\ 10 & -5 & 15 \\ -5 & 5 & 5 \end{bmatrix} = 5\begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ -1 & 1 & 1 \end{bmatrix}$$

$$\therefore 5A = 5\begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ -1 & 1 & 1 \end{bmatrix} \Rightarrow A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ -1 & 1 & 1 \end{bmatrix}$$

From the second equation, we have

$$B = 2A - \begin{bmatrix} 2 & -1 & 5 \\ 2 & -1 & 6 \\ 0 & 1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & 4 \\ 4 & -2 & 6 \\ -2 & 2 & 2 \end{bmatrix} - \begin{bmatrix} 2 & -1 & 5 \\ 2 & -1 & 6 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -1 \\ 2 & -1 & 0 \\ -2 & 1 & 0 \end{bmatrix}$$

Example 4 :

If $A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix}$, find the value of $A^3 - A^2 + I_3$.

$$A^2 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$A^3 = A^2A = \begin{bmatrix} -1 & 0 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{Therefore } A^3 - A^2 + I = \begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} -1 & 0 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

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

Example 5 :

Verify that $[AB]^T = B^T A^T$

$$\text{Where } A = \begin{bmatrix} 1 & 2 & 3 \\ 6 & 7 & 8 \\ 6 & -3 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 2 \\ 5 & 6 & 1 \end{bmatrix}$$

$$\text{We have } AB = \begin{bmatrix} 1 & 2 & 3 \\ 6 & 7 & 8 \\ 6 & -3 & 4 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 2 \\ 5 & 6 & 1 \end{bmatrix} = \begin{bmatrix} 22 & 28 & 10 \\ 67 & 88 & 40 \\ 17 & 24 & 16 \end{bmatrix}$$

$$\therefore [AB]^T = \begin{bmatrix} 22 & 67 & 17 \\ 28 & 88 & 24 \\ 10 & 40 & 16 \end{bmatrix}$$

$$\text{Again } A^T = \begin{bmatrix} 1 & 6 & 6 \\ 2 & 7 & -3 \\ 3 & 8 & 4 \end{bmatrix} \text{ and } B^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 2 & 1 \end{bmatrix}$$

$$\therefore B^TA^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 2 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 6 & 6 \\ 2 & 7 & -3 \\ 3 & 8 & 4 \end{bmatrix} = \begin{bmatrix} 22 & 67 & 17 \\ 28 & 88 & 24 \\ 10 & 40 & 16 \end{bmatrix}$$

Hence $[AB]^T = B^T A^T$

Applications

Example 6 :

Transform the following system of equations to a single equation, using matrices :

(i) $2x+3y-5 = 0$

$3x+5y-7 = 0$

(ii) $3x-y+z = 6$

$x+2y+z = 1$

$2x+7y+5z = 5$

Solution : Equation (i) can be written as

$2x+3y = 5$

$3x+5y = 7$

Take the matrix of coefficients as

$A = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$, the matrix of variables as $X=\begin{bmatrix} x \\ y \end{bmatrix}$ and the constants on RHS as $B = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$

(We take the matrices in the above particular forms in order to get conformability for matrix multiplication)

Using matrix multiplication, we get $\begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} 5 \\ 7 \end{bmatrix}$

or $AX = B$, which represents the system of equations as a single matrix equation.

(ii) As before, the matrix equation in this case is

$AX = B$, where

$$A=\begin{bmatrix} 3 & -1 & 1 \\ 1 & 2 & 1 \\ 2 & 7 & 5 \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ and } B = \begin{bmatrix} 6 \\ 1 \\ 5 \end{bmatrix}$$

N.B. The solution of AX=B shall be dealt with in the next chapter.

Example 7 :

The total investment in farming and business by two persons P_1 and P_2 is ₹10 lakhs and ₹15 lakhs respectively. If these two sectors yield profit annually at the rate of 20% and 25% respectively, determine how can they divide their investments in the respective sectors so as to get total annual profit of ₹2.3 lakhs and ₹3.5 lakhs respectively.

Solution :

Taking investment of P_1 and P_2 in farming as x and y respectively, the 'investment matrix', in lakhs is given by

$$A = \begin{array}{c} \\ P_1 \\ P_2 \end{array}\begin{array}{c} \text{Farming} \quad\quad \text{Business} \\ \begin{bmatrix} x & 10-x \\ y & 15-y \end{bmatrix} \end{array}$$

and the 'rate of profit matrix' is given by

$$B = \begin{bmatrix} 20\% \\ 25\% \end{bmatrix}$$

$\therefore$ The 'total profit matrix', in lakhs, is given by

$$AB = \begin{bmatrix} \dfrac{20x}{100} + \dfrac{25(10\text{-}x)}{100} \\ \dfrac{20y}{100} + \dfrac{25(15-y)}{100} \end{bmatrix} = \begin{bmatrix} 2.3 \\ \\ 3.5 \end{bmatrix}$$

(You can choose the matrices in some other way and, depending upon their conformability for multiplication, find out the profit matrix)

Solving for x and y applying definition of equality of matrices

$x = 4$ lakhs, $y = 5$ lakhs.

$\Rightarrow P_1$ should invest ₹4 lakhs in farming and ₹6 lakhs in business and P_2 should invest ₹5 lakhs in farming and ₹10 lakhs in business in order to get the total desired profit.

EXERCISES 4 (a)

1. State the order of the following matrices.

   (i) $[a\ b\ c]$

   (ii) $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

   (iii) $\begin{bmatrix} x & y \\ y & z \\ z & x \end{bmatrix}$

   (iv) $\begin{vmatrix} 1 & 0 & 1 & 4 \\ 2 & 1 & 3 & 0 \\ -3 & 2 & 1 & 3 \end{vmatrix}$

2. How many entries are there in a

   (i) 3×3 matrix

   (ii) 3×4 matrix

   (iii) $p \times q$ matrix

   (iv) a square matrix of order p ?

3. Give an example of

   (i) 3×1 matrix

   (ii) 2×2 matrix

   (iii) 4×2 matrix

   (iv) 1×3 matrix

4. Let $A = \begin{bmatrix} 1 & 2 & 3 & 4 & 1 \\ 4 & 5 & 6 & 1 & 2 \\ 3 & 9 & 1 & 1 & 6 \end{bmatrix}$

(i) what is the order of A ?

(ii) Write down the entries a_{31}, a_{25}, a_{23}.

(iii) Write down A^T.

(iv) What is the order of A^T ?

5. Matrices A and B are given below. Find A + B, B + A, A – B and B – A. Verify that A + B = B + A and B – A = – (A – B).

(i) $A = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$ $\qquad B = \begin{bmatrix} -6 \\ 9 \end{bmatrix}$

(ii) $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$ $\qquad B = \begin{bmatrix} 4 & 1 \\ -3 & 2 \end{bmatrix}$

(iii) $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{5} \end{bmatrix}$ $\qquad B = \begin{bmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & \frac{4}{5} \end{bmatrix}$

(iv) $A = \begin{bmatrix} 1 & a-b \\ a+b & -3 \end{bmatrix}$ $\qquad B = \begin{bmatrix} 1 & b \\ -a & 5 \end{bmatrix}$

(v) $A = \begin{bmatrix} 1 & -2 & 5 \\ -1 & 4 & 3 \\ 1 & 2 & -3 \end{bmatrix}$ $\qquad B = \begin{bmatrix} -1 & 2 & -5 \\ 1 & -3 & -3 \\ 1 & 2 & 4 \end{bmatrix}$

6. (i) Find the 2 × 2 matrix X if $X + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$.

(ii) Given $[x\ y\ z] - [-4\ 3\ 1] = [-5\ 1\ 0]$, determine x, y, z.

(iii) If $\begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} - \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$, determine x_1, x_2, y_1, y_2.

(iv) Find the matrix which when added to $\begin{bmatrix} 2 & -3 \\ -4 & 7 \end{bmatrix}$ gives $\begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix}$

7. Calculate whenever possible, the following products.

(i) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ $\qquad$ (ii) $\begin{bmatrix} 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$

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

8. If $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$, $C = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$

Calculate (i) AB (ii) BA (iii) BC, (iv) CB, (v) AC, (vi) CA

9. Find the following products.

(i) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix}$ (iv) $\begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

(v) $\begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}^2$ where $i = \sqrt{-1}$ (vi) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

(vii) $\begin{bmatrix} 0 & k \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ (viii) $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

(ix) $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ (x) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

10. Write true or false in the following cases :

(i) The sum of a 3×4 matrix with a 3×4 matrix is a 3×3 matrix

(ii) $k\,[0] = 0, k \in R$

(iii) A – B = B – A if one of A and B is zero and A and B are of the same order.

(iv) A + B = B + A, if A and B are matrices of the same order

(v) $\begin{bmatrix} 1 & 0 \\ -2 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 2 & 0 \end{bmatrix} = 0$ (vi) $\begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix} = 3 \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$

(vii) With five elements a matrix cannot be constructed.

(viii) The unit matrix is its own transpose.

11. If $A = \begin{bmatrix} 2 & 4 \\ 3 & 13 \end{bmatrix}$ and $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ find $A - \alpha I, \alpha \in R$.

12. Find x and y in the following

(i) $\begin{bmatrix} x & -2y \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & -8 \\ 0 & -2 \end{bmatrix}$ (ii) $\begin{bmatrix} x+3 \\ 2-y \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$

(iii) $\begin{bmatrix} 2x-y \\ x+y \end{bmatrix} = \begin{bmatrix} 3 \\ -9 \end{bmatrix}$ (iv) $\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

(v) $[2x \quad -y] + [y \quad 3x] = 5\,[1 \quad 0]$

13. The element of *i*th row and *j*th column of the following matrix is $i + j$. Complete the matrix.

$$\begin{bmatrix} 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & - \\ 4 & 5 & - & - \\ 5 & - & - & - \\ - & - & - & - \end{bmatrix}$$

14. Write down the matrix $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$ if $a_{ij} = 2i + 3j$.

15. Construct a 2×3 matrix having elements given by

(i) $a_{ij} = i + j$ (ii) $a_{ij} = i - j$

(iii) $a_{ij} = i \times j$ (iv) $a_{ij} = \frac{i}{j}$

16. If $\begin{bmatrix} 2x & y \\ 1 & 3 \end{bmatrix} + \begin{bmatrix} 4 & 2 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 8 & 3 \\ 1 & 2 \end{bmatrix}$; find x and y.

17. Find A such that $\begin{bmatrix} 2 & 3 & 4 \\ 1 & 0 & -2 \\ 3 & 1 & -1 \end{bmatrix} + A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & -1 & 0 \\ 1 & 3 & 2 \end{bmatrix}$

18. If $\begin{bmatrix} x+y & x-z \\ 2x-y & 0 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}$; find the values of x, y, z.

19. What is the order of the matrix B if $[\,3\ 4\ 2\,]\,B = [\,2\ \ 1\ \ 0\ \ 3\ \ 6\,]$

20. Find A if, $\begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} A = \begin{bmatrix} -4 & 8 & 4 \\ -1 & 2 & 1 \\ -3 & 6 & 3 \end{bmatrix}$

21. Find B if $B^2 = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$

22. Find x and y when $\begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$

23. Find AB and BA given that

(i) $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$ (ii) $A = \begin{bmatrix} 7 & 5 \\ 6 & 4 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(iii) $A = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}, B = \begin{bmatrix} 4 & 2 \\ -1 & -2 \end{bmatrix}$ (iv) $A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, B = \begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix}$

24. Evaluate

(i) $[[2, 1] + 2\,[0 - 2]] \begin{bmatrix} 2 & 1 & 3 \\ 1 & -2 & 0 \end{bmatrix}$ (ii) $\left[\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix} + \begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix}\right] \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

25. If $A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ -1 & 4 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \\ 2 & 2 & 2 \end{bmatrix}$ show that AB = AC, though B ≠ C.

Verify that (i) A + (B + C) = (A + B) + C,
(ii) A (B + C) = AB + AC (iii) A (BC) = (AB) C

26. Find A and B where

$2A + B = \begin{bmatrix} 2 & 2 & 5 \\ 5 & 4 & 3 \\ 1 & 1 & 4 \end{bmatrix}$ and $A - 2B = \begin{bmatrix} 1 & 6 & 5 \\ 5 & 2 & -1 \\ -2 & -2 & 2 \end{bmatrix}$

27. If $A = \begin{bmatrix} 4 & 2 \\ -1 & 1 \end{bmatrix}$ and I is the 2 × 2 unit matrix, find (A – 2I) (A – 3I).

28. Verify that $[AB]^T = B^T A^T$ where (i) $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & -2 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ -1 & 1 \end{bmatrix}$ and

(ii) $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}, B = \begin{bmatrix} 2 & 3 & 0 \\ 1 & 2 & 3 \end{bmatrix}$

29. Verify that $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ satisfies the equation $x^2 - (a + d)x + (ad - bc)\,I = 0$ where I is the 2 × 2 unit matrix.

30. If $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & -2 & 1 \\ 4 & 2 & 1 \end{bmatrix}$, show that $A^3 - 23A - 40\,I = 0$

31. Simplify : $[x\,y\,z] \times \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

32. If A and B are matrices of the same order and AB = BA, then prove that
(i) $A^2 - B^2 = (A - B)(A + B)$
(ii) $A^2 + 2AB + B^2 = (A + B)^2$
(iii) $A^2 - 2AB + B^2 = (A - B)^2$

33. If α and β are scalars and A is a square matrix then prove that
$(A - \alpha I).(A - \beta I) = A^2 - (\alpha + \beta)A + \alpha\beta I$ where I is a unit matrix of same order as A.

34. If α and β are scalars such that $A = \alpha B + \beta I$ where A, B and the unit matrix I are of the same order, then prove that AB = BA.

35. If $A = \begin{bmatrix} -1 & 3 & 5 \\ 1 & -3 & -5 \\ -1 & 3 & 5 \end{bmatrix}$, show that $A^2 = A$.

36. If $A = \begin{bmatrix} 0 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 7 & 6 & 3 \\ 1 & 4 & 5 \end{bmatrix}$, find 2A + 3B, 2A – 3B.

37. If $A = \begin{bmatrix} 1 & 0 & -2 \\ 2 & 3 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 4 & -1 & 3 \\ 0 & 2 & 1 \end{bmatrix}$ and $C = \begin{bmatrix} 2 & -3 & 0 \\ 1 & 4 & 5 \end{bmatrix}$, find
(i) $A - 3B + 2C$ (ii) $(A + B - C)^T$ (iii) $B^T - C^T$.

38. If $A = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 3 & 0 \end{bmatrix}$, verify
(i) (A + B) C = AC + BC (ii) (AB) C = A (BC)

39. $\begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x & 2 \\ 1 & y \end{bmatrix} = \begin{bmatrix} -3 & 4 \\ -1 & 4 \end{bmatrix}$, find x and y.

40. If $A = \begin{bmatrix} 1 & -2 & 2 \\ 3 & 1 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 4 \\ 1 & 2 \\ 3 & -1 \end{bmatrix}$, verify that $(AB)^T = B^T A^T$

41. If A, B, C are matrices of order 2 × 2 each and $2A + B + C = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$, $A + B + C = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$,
$A + B - C = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$, find A, B and C.

42. If $\begin{bmatrix} x & y \\ x & \frac{x}{2} + t \end{bmatrix} + \begin{bmatrix} y & x + t \\ x + 2 & \frac{x}{2} \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$, find x, y, z and t.

43. There are two families A and B. In family A there are 4 men, 6 women and 2 children and in family B there are 2 men, 2 women and 4 children. The recommended daily amount of calories is 2400 for men, 1900 for women, 1800 for children and 45 gram of protein for men, 55 grams for women and 33 grams for children. Represent the above information using matrices. By matrix multiplication, calculate the total requirement of calories and proteins for each of the two families.

44. A trust fund has ₹50,000 that is to be invested in two types of bonds. The first and second bonds respectively pay annual interest at the rate of 5% and 6% respectively. Using matrix multiplication, determine how to invest the money in these bonds so as to get a total annual interest of ₹2780.

4.4 Symmetric and Skew Symmetric Matrix :

Definition (Symmetric matrix) : A matrix which is equal to its transpose, is a symmetric matrix.

Obviously, **for a symmatric matrix $A = [a_{ij}]$, $a_{ij}=a_{ji}$ for all i and j.**

Corollary :

A symmetric matrix is necessarily a square matrix.

For if $A = A'$ and A is of order m x n then A' is of order n x m and

$A = A' \Rightarrow m = n$ so that that A and A' are both square matrices.

So the above definition is in the best possible form.

Example :

$\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ are symmetric matrices.

Definition (Skew Symmetric matrix)

A matrix A such that $A = -A'$ is called skew symmetric.

Obviously, for a **skew symmetric matrix $A = [a_{ij}]$, $a_{ij} = -a_{ji}$ for all i and j.**

Also, a skew symmetric matrix is necessarily a square matrix.

A skew symmetric matrix enjoys an additional property :

The diagonal elements of a skew symmetric matrix are zero.

Proof : If $A=[a_{ij}]$ is skew symmetric, then $a_{ij} = -a_{ji}$ for all i and j, so for diagonal elements we have $a_{ii} = -a_{ii}$ for all i,

which implies $a_{ii} = 0$ for all i.

Hence the result.

Note: (1) It follows from definition that if A is a symmetric/skew symmetric matrix then α A is also symmetric/ skew symmetrc for any scalar α.

(2) The zero matrix is both symmetric and skew symmetric. The converse is also true:

A matrix which is both symmetric and skew symmetric, must be the zero matrix.

Let A=$[a_{ij}]$ be both symmetric and skew symmetric. Then we have

$a_{ij} = a_{ji}$ (by symmetricity)

$= - a_{ij}$ (by skew symmetricity)

$\Rightarrow a_{ij} = 0$ for all i & j.

$\Rightarrow$ A is a zero matrix.

Now let us see what see what happens to A+A′ and A-A′ !

A-A′

Take an example :

$$A = \begin{bmatrix} 1 & -2 & 0 \\ -3 & 2 & 1 \\ 3 & 5 & 3 \end{bmatrix}$$

$$\text{Then } A' = \begin{bmatrix} 0 & 1 & -4 \\ -1 & 0 & -4 \\ 4 & 4 & 0 \end{bmatrix}$$

$$\text{and } -A' = \begin{bmatrix} -1 & 3 & -4 \\ 2 & -2 & -5 \\ 0 & -1 & -3 \end{bmatrix}$$

$$\text{Thus } A+A' = \begin{bmatrix} 2 & -5 & 4 \\ -5 & 4 & 6 \\ 4 & 6 & 6 \end{bmatrix}$$

$$\text{and } A-A' = \begin{bmatrix} 0 & 1 & -4 \\ -1 & 0 & -4 \\ 4 & 4 & 0 \end{bmatrix}$$

Observe that A+A′ is symmetric, whereas A-A′ is skew-symmetric.

However, it is not an isolated phenomenon with particular matrices.

Our next two theorems establish some general results along this direction.

Theorem - 1

For a square matrix A

(i) $A+A'$ is a symmetric matrix and

(ii) $A-A'$ is a skew symmetric matrix.

Proof : (i) $(A+A')' = A' + (A')'$ ($\because$ Transpose of sum = sum of transposes)

$= A' + A$ ($\because$ Transpose of transpose is the original matrix)

$= A + A'$ ($\because$ matrix addition is commutative)

Since $(A+A')' = A+A'$, it follows that $A+A'$ is a symmetric matrix.

(ii) $(A-A')' = A' + (-A')' = A' - (A')'$ (by prperty of transpose of scalar multiple, taking scalar as -1)

$= A' - A = -(A-A')$ (by property of multiplication by a scalar)

Since $(A-A')' = -(A-A')$, it follows that $A-A'$ is skew symmetric matrix.

Theorem-2

A square matrix can be uniquely expressed as a sum of a symmetric and a skew symmetric matrix.

Proof:

Let A be a square matrix.

By the property of matrix-addition and multiplication by a scalar,

$$A = \frac{1}{2}(A + A') + \frac{1}{2}(A - A')$$

$A+A'$ is symmetric and $A-A'$ is skew symmetric. (By Theorem-1)

$\Rightarrow \frac{1}{2}(A + A')$ and $\frac{1}{2}(A - A')$ are respectively symmetric and skew symmetric.

($\because$ If A is symmetric/skew symmetric, then α A is symmetric/skew symmetric for any scalar α)

To prove uniqueness, suppose $A = R + S$ where R is symmetric and S is skew symmetric.

$\therefore R = R'$ and $S = -S'$

So $A = R + S = R' - S'$

$\Rightarrow A' = (R' - S')' = (R')' - (S')' = R - S.$

Thus

$R + S = A$ and $R - S = A'$

$\Rightarrow R = \frac{1}{2}(A + A')$ and $S = \frac{1}{2}(A - A')$

So $A = R + S = \frac{1}{2}(A + A') + \frac{1}{2}(A - A')$

and the expression is unique.

Illustrative Examples

Example - 7

If A and B are both symmetric matrices of the same order, then AB is symmetric if and only if AB=BA.

Proof:

Suppose $AB = BA$

$(AB)' = B'A'$ (transpose of product)

$= BA$ ($\because$ A and B are symmetric)

$= AB$ (by hypothesis)

$\Rightarrow$ AB is symmetric.

Conversely suppose that AB is symmetric.

We shall show that AB=BA.

$AB = (AB)'$ ($\because$ AB is symmetric)

$= B'A'$ (transpose of product)

$= BA$ ($\because$ A and B are shymmetric)

It follows that $AB = BA$.

Example-8

Express $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \\ -1 & 5 & -2 \end{bmatrix}$ as a sum of a symmetric and skew symmetric matrices.

Solution :

$$\text{Taking } A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \\ -1 & 5 & -2 \end{bmatrix}$$

$$A + A' = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \\ -1 & 5 & -2 \end{bmatrix} + \begin{bmatrix} 1 & 4 & -1 \\ 2 & 0 & 5 \\ 3 & 1 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 6 & 2 \\ 6 & 0 & 6 \\ 2 & 6 & -4 \end{bmatrix}$$

$$A\text{-}A' = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \\ -1 & 5 & -2 \end{bmatrix} - \begin{bmatrix} 1 & 4 & -1 \\ 2 & 0 & 5 \\ 3 & 1 & -2 \end{bmatrix} = \begin{bmatrix} 0 & -2 & 4 \\ 2 & 0 & -4 \\ -4 & 4 & 0 \end{bmatrix}$$

$$\therefore\ A = \frac{1}{2}(A+A') + \frac{1}{2}(A\text{-}A') = \begin{bmatrix} 1 & 3 & 1 \\ 3 & 0 & 3 \\ 1 & 3 & -2 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & -2 \\ -2 & 2 & 0 \end{bmatrix}$$

$\begin{bmatrix} 1 & 3 & 1 \\ 3 & 0 & 3 \\ 1 & 3 & -2 \end{bmatrix}$ is symmetric and $\begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & -2 \\ -2 & 2 & 0 \end{bmatrix}$ is skew symmetric.

4.5 Transformation of Matrices

(Elementary Row and Column Operations)

A matrix can be transformed into another by charging its rows or columns through some elementary operations as illustrated below :

1. **Interchange of any two rows/columns, symbolically represented as $R_i \leftrightarrow R_j$ or $C_i \leftrightarrow C_j$. R_i, R_j stand for i^{th} and j^{th} rows and C_i, C_j stand for i^{th} and j^{th} columns.**

Example-9

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \xrightarrow{C_1 \leftrightarrow C_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \xrightarrow{C_2 \leftrightarrow C_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(a) Applying $C_1 \leftrightarrow C_3$ to $\begin{bmatrix} \sqrt{2} & 2 & 0 \\ 3 & -1 & \sqrt{5} \\ 1 & 3 & 5 \end{bmatrix}$

we get $\begin{bmatrix} 0 & 2 & \sqrt{2} \\ \sqrt{5} & -1 & 3 \\ 5 & 3 & 1 \end{bmatrix}$.

This process of transformation of the matrices is described as

$$\begin{bmatrix} \sqrt{2} & 2 & 0 \\ 3 & -1 & \sqrt{5} \\ 1 & 3 & 5 \end{bmatrix} \xrightarrow{C_1 \leftrightarrow C_3} \begin{bmatrix} 0 & 2 & \sqrt{2} \\ \sqrt{5} & -1 & 3 \\ 5 & 3 & 1 \end{bmatrix}$$

(b) $$\begin{bmatrix} \sqrt{2} & 2 & 0 \\ 3 & -1 & \sqrt{5} \\ 1 & 3 & 5 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 3 & -1 & \sqrt{5} \\ \sqrt{2} & 2 & 0 \\ 1 & 3 & 5 \end{bmatrix}$$

Example-10

Transform the following into unit matrices :

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Solution:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \xrightarrow{C_1 \leftrightarrow C_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(we can also apply $R_1 \leftrightarrow R_2$)

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

We can also apply column operations :

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \xrightarrow{C_1 \leftrightarrow C_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \xrightarrow{C_2 \leftrightarrow C_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2. **Multiplying the elements of a column/row by a ronzero scalar, symbolically represented as $R_i \to kR_i$ or $C_j \to kC_j$; $k \neq 0$.**

3. **Replacing a row/column by adding to it a nonzero scalar multiplis of elements of another row/column, symbolically :**

$R_i \to R_i + kR_j$ or $C_i \to C_i\ kC_j$.

Example-11

Transform into unit matrix :

(i) $\begin{bmatrix} 6 & 3 \\ 2 & 2 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

Solution :

(i) $$\begin{bmatrix} 6 & 3 \\ 2 & 2 \end{bmatrix} \xrightarrow[R_2 \to \frac{1}{2}R_2]{R_1 \to \frac{1}{3}R_1} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \xrightarrow{R_1 \to R_1 - R_2} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \xrightarrow{R_2 \to R_2 - R_1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(ii) $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \xrightarrow[R_3 \to \frac{1}{3}R_3]{R_1 \to R_1 - 2R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

4.6 Inverse of a Matrix

Definition (Inverse of a matrix) :

If A and B are two square matrices of the same order such that AB=BA=I, the unit matrix of the same order as A or B, then B is called the multiplicative inverse of A or simply inverse of A, written as A^{-1}. A is also called inverse of B, written as B^{-1}.

Note:

1. We do not define inverse of a matrix which is not square because of the requirement AB=BA in the definition of inverse, which requires that both must be square matrices.
2. A matrix whose inverse exists, is called an **invertible matrix**.

Illustrative Emaples :

Example-12

$$\begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}\begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

So both the matrices on the LHS are invertible and are inverses of each other.

Example-13

$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ is not invertible.

You can see this from the fact that

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} x & y \\ z & t \end{bmatrix} = \begin{bmatrix} x+2z & y+2t \\ 2x+4z & 2y+4t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

gives rise to inconsistent equations

x+2y=1, 2x+4z=0 and y+2t=0, 2y+4t=1 in terms of the unknowns x,y,z,t.

Just what types of matrices are invertible and what are not, shall be one of the objects of our study in the next chapter.

However, it may be remarked that **if, in the course of applying elementary operations to a matrix we get all the zeros in a row/column, then its inverse does not exist.**

Further discussion regarding this shall be done in the next chapter.

Uniqueness of Inverse

Theorem-3

Inverse of a matrix, if it exists, is unique.

Proof :

Let a square matrix A have inverses B and C. Then, by definition,

$AB = BA = I$

and $AC = CA = I$

Now, $C = CI = C(AB) = (CA)B = IB = B.$

$\Rightarrow C = B$ and there is only one inverse of A, in other words, inverse of a matrix is unique.

Inverse of Product :

Theorem - 4

If A and B are invertible matrices of the same order, then

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

Proof :

$$(AB)(B^{-1}A^{-1}) = A\left(B(B^{-1}A^{-1})\right) = A\left((BB^{-1})A^{-1}\right) = A(IA^{-1}) = AA^{-1} = I$$

(by associative property of matrix-multiplication and property of inverse of a matrix)

Since $(AB)(B^{-1}A^{-1}) = I$, it follows that

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

Inversion of Matrices :

Given that a matrix is invertible, we now discuss methods of finding its inverse.

Let A be a square matrix of order n and I, a unit matrix of the same order. Supposing that A is invertible, the procedure for finding A^{-1} is detailed below :

1. **Elementary row operations:**

We use the equation

$A = IA$

and transform the matrices involved in such a manner that

(i) The post-factor A in RHS is left unchanged.

(ii) A in LHS gets transformed into the unit matrix I through row operations which are to be reflected onto the pre-factor I in RHS, ultimately giveing us

$I = BA$.

Since we have supposed that A is invertible A^{-1} uniquely exists such that $AA^{-1}=A^{-1}A=I$.

So post-multiplying A^{-1} to

$I = BA$

$IA^{-1} = (BA)\,A^{-1} = B\,(AA^{-1}) = BI$

$\Rightarrow A^{-1} = B.$

N.B.: In the format A = IA we can apply only row operations on A and consequently on I because it is the rows of I in RHS which get involved in the multiplication.

2. **Elementary column operations:**

We take the equation

A = AI, leave A on RHS unchanged, carry out elementary column operations on A until we convert it to the unit matrix, simultaneously applying the same column operations on I on RHS; thus getting

$I = AB$, which gives

$B = A^{-1}$.

N.B.: We take the equation A=AI, because it is only the columns of I which get involved in the product AI.

Illustrative Examples :

Example - 14

Find inverse of $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$ applying elementary operations.

Solution :

To apply column operations we write :

$$A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} = A \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

or $$\begin{bmatrix} 1 & 5 \\ 0 & 3 \end{bmatrix} = A \begin{bmatrix} 3 & 0 \\ -1 & 1 \end{bmatrix}$$

$(C_1 \rightarrow 3C_1 - C_2)$

or $$\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} = A \begin{bmatrix} 3 & -15 \\ -1 & 6 \end{bmatrix}$$

$(C_2 \rightarrow C_2 - 5C_1)$

or $$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = A \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$$

$(C_2 \rightarrow \frac{1}{3} C_2)$

$$\therefore A^{-1} = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}.$$

(You can also apply elementary row operation in the order $R_1 \to R_1 - R_2$, $R_2 \to R_2 - R_1$ and $R_1 \to R_1 - 2R_2$ in the equation A= IA and arrive at the same result)

Example - 15

Applying elementary operations, find the inverse of

$$\text{(i)} \begin{bmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix}, \quad \text{(ii)} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 1 & 0 & 2 \end{bmatrix}$$

Solution :

(i) In order to apply column operations we write

$$A = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix} = A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{or} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = A \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \ (C_1 \leftrightarrow C_3)$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = A \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix}.$$

$$\left(C_1 \to \frac{1}{2}C_1,\ C_2 \to \frac{1}{2}C_2,\ C_3 \to \frac{1}{2}C_3\right)$$

$$\therefore A^{-1} = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix}$$

$$\text{(ii)} \quad A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A$$

$$\text{or} \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 4 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} A$$

$(R_1 \leftrightarrow R_3)$

$$\text{or}\begin{bmatrix}1&0&2\\0&1&0\\1&2&3\end{bmatrix}=\begin{bmatrix}0&0&1\\0&1&-2\\1&0&0\end{bmatrix}A$$

$(R_2 \to R_2 - 2R_1)$

$$\text{or}\begin{bmatrix}1&0&2\\0&1&0\\1&2&1\end{bmatrix}=\begin{bmatrix}0&0&1\\0&1&-2\\1&0&-1\end{bmatrix}A$$

$(R_3 \to R_3 - R_1)$

$$\text{or}\begin{bmatrix}1&0&2\\0&1&0\\0&0&1\end{bmatrix}=\begin{bmatrix}0&0&1\\0&1&-2\\1&-2&3\end{bmatrix}A$$

$(R_3 \to R_3 - 2R_2)$

$$\text{or}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}=\begin{bmatrix}-2&4&-5\\0&1&-2\\1&-2&3\end{bmatrix}A$$

$(R_1 \to R_1 - 2R_3)$

$$\therefore A^{-1}=\begin{bmatrix}1&2&3\\2&1&4\\1&0&2\end{bmatrix}^{-1}=\begin{bmatrix}-2&4&-5\\0&1&-2\\1&-2&3\end{bmatrix}$$

Exercises - 4 (b)

1. State which of the following matrices are symmetric, skew symmetric, both or not either :

(i) $\begin{bmatrix}0&1\\1&0\end{bmatrix}$ (ii) $\begin{bmatrix}x&1&2\\-1&y&3\\-2&-3&z\end{bmatrix}$, $(x,y,z) \neq (0,0,0)$

(iii) $\begin{bmatrix}1&0&3\\0&-1&2\\3&2&1\end{bmatrix}$ (iv) $\begin{bmatrix}0&1&2\\-1&0&3\\-2&-3&0\end{bmatrix}$ (v) $\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}$

(vi) $\begin{bmatrix}1&-1&2\\1&1&-3\\-2&3&1\end{bmatrix}$ (vii) $\begin{bmatrix}0&0&1\\0&0&-2\\-1&2&0\end{bmatrix}$

2. State 'True' or 'False':

(i) If A and B are symmetric matrices of the same order and $AB - BA \neq 0$, then AB is not symmetric.

(ii) For any square matrix A, AA' is symmetric.

(iii) If A is any skew symmetric matrix, then A^2 is also skew symmetric.

(iv) If A is symmetric, then $A^2, A^3, ..., A^n$ are all symmetric.

(v) If A is symmetric then $A - A^1$ is both symmetric and skew symmetric.

(vi) For any square matrix $(A - A^1)^2$ is skew symmetric.

(vii) A matrix which is not symmetric is skew symmetric.

3. (i) If A and B are symmetric matrices of the same order with $AB \neq BA$, final whether $AB - BA$ is symmetric or skew symmetric.

(ii) If a symmetric/skew symmetric matrix is expressed as a sum of a symmetric and a skew symmetric matrix then prove that one of the matrices in the sum must be zero matrix.

4. A and B are square matrices of the same order, prove that

(i) If A, B and AB are all symmetric, then $AB - BA = 0$

(ii) If A, B and AB are all skew symmetric then $AB + BA = 0$

5. If $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ -2 & 5 & 3 \end{bmatrix}$, then verify that

(i) A+A' is symmetric

(ii) A-A' is skew symmetric

6. Prove that a unit matrix is its own inverse. Is the converse true ?

If $A = \begin{bmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{bmatrix}$ show that $A^2 = I$ and hence $A = A^{-1}$.

7. (Here A is an involuntary matrix, recall the definition given earlier)

8. Show that $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is its own inverse.

9. Express as a sum of a symmetric and a skew symmetric matrix.

(i) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \\ -1 & 5 & -2 \end{bmatrix}$ (ii) $\begin{bmatrix} 2 & -1 & 3 \\ 5 & 7 & -2 \\ 1 & 4 & 6 \end{bmatrix}$ (iii) $\begin{bmatrix} x & a & b \\ a & y & c \\ b & c & z \end{bmatrix}$

(iv) $\begin{bmatrix} 0 & x \\ -x & 0 \end{bmatrix}$ (v) $\begin{bmatrix} 1 & 5 \\ 7 & -3 \end{bmatrix}$ (vi) $\begin{bmatrix} 4 & -3 \\ 1 & 2 \end{bmatrix}$

(vii) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

10. What is the inverse of

(i) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

11. Find inverse of the following matrices by elementary row/column operation (transformations):

(i) $\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$ (ii) $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$ (iii) $\begin{bmatrix} 4 & -2 \\ 3 & 1 \end{bmatrix}$

(iv) $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$ (v) $\begin{bmatrix} 1 & 0 \\ 2 & -3 \end{bmatrix}$ (vi) $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

12. Find the inverse of the following matrices using elementary transformation :

(i) $\begin{bmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix}$ (ii) $\begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}$

(iii) $\begin{bmatrix} 3 & -2 & 3 \\ 2 & 1 & -1 \\ 4 & -3 & 2 \end{bmatrix}$ (iv) $\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$

(v) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 1 & 0 & 2 \end{bmatrix}$

CHAPTER 5

Determinants

In the sky, there is no distinction of east and west; people create distinctions out of their own minds and then believe them to be true.

- Buddha

5.0 Introduction

Around 1100 B.C., the Chinese had first used the concept of determinants in solving linear equations. After a long gap of time certain rules of determinants were given by Leibnitz (1646 – 1716). Later other mathematicians, notably, G. Cramer (1704–1752), A.T. Lagrange (1738 – 1813), C.G.J.Jacobi (1804–1851), J.J.Sylvester (1814–1890), and Cayley (1821 – 1895) made significant contributions to the theory of determinants. Sylvester was the first to use the word matrix and Cayley made extensive study of the theory of martices. Besides its use in the solution of linear equation, martices are now being used as a tool in various disciplines.

We have already discussed matrices and in this chapter, associated with determinants, we shall study their use in the solution of systems of linear equations.

5.1 Determinant of a square matrix

Let us solve the two linear equations

$a_1x + b_1y + c_1 = 0$ .... (i)

and $a_2x + b_2y + c_2 = 0$ .... (ii)

To eliminate y, we multiply (i) by b_2 and (ii) by b_1 and then by subtraction we get

$$x = \frac{b_1c_2 - b_2c_1}{a_1b_2 - a_2b_1}$$

Similarly eliminating x from equations (i) and (ii) we get

$$y = \frac{c_1a_2 - c_2a_1}{a_1b_2 - a_2b_1}$$

These solutions exist provided $a_1b_2 - a_2b_1 \neq 0$

The quantity $a_1b_2 - a_2b_1$ determines whether a solution of the linear equations (i) and (ii) exists or not. It is especially denoted by the symbol $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$ which is called a determinant (of order two). Thus

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1 \text{ .... (1)}$$

In fact, it is called the **determinant of the square matrix** $\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$.

Similarly we can show by the method of elimination that the set of equations

$$a_1x + b_1y + c_1z + d_1 = 0$$
$$a_2x + b_2y + c_2z + d_2 = 0$$
$$a_3x + b_3y + c_3z + d_3 = 0$$

produces a solution if

$$a_1b_2c_3 - a_1b_3c_2 + a_3b_1c_2 - a_2b_1c_3 + a_2b_3c_1 - a_3b_2c_1 \neq 0$$

The above quantity on the left is again denoted

by $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

which is called a determinant (of order three)

Thus

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1b_2c_3 - a_1b_3c_2 + a_3b_1c_2 - a_2b_1c_3 + a_2b_3c_1 - a_3b_2c_1 \text{ .... (2)}$$

$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$ **is the determinant of the square matrix** $\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$

Given a determinant as on the L.H.S of (2) how do we write the value as on the R.H.S ? We shall study below the rules of 'expansion' of a determinant. But before we proceed let us get ourselves acquainted with some new terminology.

5.2 Minors, Cofactors and Expansion of a determinant.

Let $\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

be a determinant. The entries a_1, b_1, c_1, a_2, b_2 etc. are real or complex numbers and are called the **elements** of the determinant. The elements which are placed horizontally form a **row** and those placed vertically form a **column. The number of rows (which is same as the number of columns) is called the order of the determinant**. Here Δ is a determinant of order three (we also call Δ a third order determinant.) The elements along the 'diagonal' of the determinant viz. a_1, b_2, c_3 in Δ, are called **diagonal elements.**

A determinant can be expressed as a compact symbol $|a_{ij}|$ where a_{ij} stands for the element in the *ith* row and *jth* column. The suffixes *i* and *j* which represent row-number and column-number respectively vary from 1 to *n*, where *n* is the order of the determinant.

Thus in Δ above $a_{11}=a_1$, $a_{12}=b_1$, $a_{13}=c_1$, $a_{21}=a_2$, $a_{22}=b_2$, $a_{23}=c_2$, $a_{31}=a_3$, $a_{32}=b_3$, $a_{33}=c_3$.

Now Δ can be expressed as the sum of products of its elements as on the *r.h.s.* of (2). This expression on the *r.h.s.* of (2) is called the **expansion** of the determinant Δ. These terms in the expansion, of which three are preceeded by positive sign and three are preceeded by negative sign, are obtained by fixing *a, b, c* in natural order and arranging the suffixes 1, 2, 3 in all possible orders (Hence $\lfloor 3 = 6$ terms). The term $a_1b_2c_3$ which is called the **leading term** is always preceeded by positive sign.

However the sign preceeding a term is determined as per the following principle :

A pair of positive integers (p,q) is called an **inversion** if $p > q$. A term is preceeded by positive sign or negative sign according as the number of inversions in the order of the suffixes is even or odd. For example, the ordered arrangements (2, 1,3) and (3, 2, 1) have one inversion viz. (2, 1) and three inversions viz. (3, 2), (3, 1) and (2, 1) respectively. Hence the terms $a_2\,b_1\,c_3$ and $a_3\,b_2\,c_1$ are preceeded by negative sign. We shall learn an easier way of expanding a determinant using the concept of minor and cofactor.

It must be borne in mind that a term being preceeded by a positive or negative sign does not mean that the term is positive or negative after numerical evaluation. For example : $a_2b_1c_3$ is preceeded by negative sign, where as $-a_2b_1c_3$ becomes -(-1).23=6, which is positive for $a_2=-1$, $b_1=2$, $c_3=3$.

Minor : The minor of any element of a determinant Δ is the determinat obtained from Δ by deleting the row and column in which the given element occurs. For example in Δ,

minor of $a_1 = \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}$, minor of $c_2 = \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}$ and minor of $b_3 = \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$;

we denote the minor of an element a_{ij} **of Δ by $\mathbf{M}_{ij}$**.

The minor of any element in a third order determinant is a second order determinant.

Cofactor : The cofactor of an element is equal to its minor with appropriate sign. The sign is positive if the row-number and column- number of the element add up to an even integer, otherwise it is negative. Usually the cofactor of an element is denoted by the corresponding capital letter. For example, in Δ above,

$$\text{Cofactor of } a_1 = A_1 = + \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}$$

$$\text{Cofactor of } c_2 = C_2 = - \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}$$

and Cofactor of $a_3 = A_3 = + \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}$

Sometimes we adopt the notation :

Cofactor of $a_{ij} = C_{ij} = (-1)^{i+j} M_{ij}$ ..... (3)

Expansion of Determinant

A determinant of order two is evaluated by (1)

A determinant of higher order is evaluated by 'expanding' the determinant by the elements of any row (or any column) as the sum of products of the elements of the row (column) with the cofactors of the respective elements of the same row (column).

$$\text{Thus } \Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

$$= a_1 (b_2c_3 - b_3c_2) - b_1(a_2c_3 - a_3c_2) + c_1(a_2b_3 - a_3b_2)$$

which is same as r.h.s of (2). Hence we have

$$\Delta = a_1A_1 + b_1B_1 + c_1C_1 \qquad ......(4)$$

The above expansion has been done using the elements of the first row. We can similarly verify that

$$\Delta = a_2A_2 + b_2B_2 + c_2C_2$$
$$= a_3A_3 + b_3B_3 + c_3C_3$$

and also $\Delta = a_1A_1 + a_2A_2 + a_3A_3 = b_1B_1 + b_2B_2 + b_3B_3$ etc.

N.B. If the elements along any row or column are all zero, then the value of the determinant is zero.

5.3 Application of determinants in finding the area of a triangle

In chapter-11, Vol-I, the formula for the area of a triangle with vertices $A(x_1,y_1)$, $B(x_2,y_2)$, and $C(x_3,y_3)$ was proved to be $|\Delta|$, where

$$\Delta = \frac{1}{2}\{x_1(y_2-y_3)+x_2(y_3-y_1)+x_3(y_1-y_2)\}$$

which is the value of the determinant $\frac{1}{2}\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{vmatrix}$.

Since area of a triangle is a positive real number, it is written as $|\Delta|$.

Condition for Collinearity

$$\Delta = 0 \text{ or } \frac{1}{2}\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{vmatrix} = 0 \text{ or } \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{vmatrix} = 0$$ also turns out to be the condition of collinearity

of the points $A(x_1,y_1)$, $B(x_2,y_2)$, and $C(x_3,y_3)$.

This may be seen from the fact that, when A, B and C are collinear, slope of $\overline{AB}$ = slope of $\overleftrightarrow{BC}$ (or slope of $\overleftrightarrow{AC}$)

By equating any two of the above slopes we can get the condition for collinearity.

N.B. When three points form a triangle, they are necessarily noncollinear, so taking area of the triangle being zero as a condition for collinearity, makes no sense as area is always a positive real number.

Example 1

Applying determinants prove that the points A(2,3), B(5,4), C(7,6) form a traingle. Find the area of ΔABC.

Solution

We have $\begin{vmatrix} 2 & 5 & 7 \\ 3 & 4 & 6 \\ 1 & 1 & 1 \end{vmatrix} = 4$

Since the determinant is non zero, the points are not collinear, hence form a triangle.

$$\Delta ABC = \frac{1}{2}\begin{vmatrix} 2 & 5 & 7 \\ 3 & 4 & 6 \\ 1 & 1 & 1 \end{vmatrix} = 2 \text{ sq. units.}$$

5.4 Some properties of Determinants

(a) A determinant remains unaltered by changing rows into columns, and columns into rows

(The determinant obtained after such change is called the transpose of the original determinant)

Let $\Delta_1 = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$ and $\Delta_2 = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$ = transpose of Δ_1, denoted as Δ_1'

Δ_2 is obtained from Δ_1 by the interchange of rows and columns.

From definition, $\Delta_1 = a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$

$= a_1(b_2c_3 - b_3c_2) - b_1(c_3a_2 - c_2a_3) + c_1(a_2b_3 - a_3b_2)$

and $\Delta_2 = a_1\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2\begin{vmatrix} b_1 & b_3 \\ a_1 & c_3 \end{vmatrix} + a_3\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$

$= a_1(b_2c_3 - b_3c_2) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1)$

$= a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1$

$= a_1(b_2c_3 - b_3c_2) - b_1(c_3a_2 - a_3c_2) + c_1(a_2b_3 - a_3b_2)$

Therefore $\Delta_1 = \Delta_2 = \Delta_1'$

Alternative Proof :

As the leading term of each of Δ_1 and Δ_2 is $a_1b_2c_3$, and the remaining 5 terms of each of Δ_1 and Δ_2 are derived from $a_1b_2c_3$ by writing *a, b, c* in the natural order, and arranging the subscripts 1, 2, 3, in all possible orders with a sign determined by the same rule of signs, the two determinants Δ_1 and Δ_2 have the same terms. Hence $\Delta_1 = \Delta_2$

(b) The interchange of two adjacent rows or columns of a determinant changes the sign of the determinant without changing its absolute value.

Let $\Delta_1 = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$ and $\Delta_2 = \begin{vmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix}$

Δ_2 is obtained from Δ_1 by interchanging the first and the second rows.

From definition, $\Delta_1 = a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$

$= a_1(b_2c_3 - b_3c_2) + b_1(c_2a_3 - a_2c_3) + c_1(a_2b_3 - a_3b_2)$

and $\Delta_2 = a_2\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} - b_2\begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} + c_2\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}$

$= a_2(b_1c_3 - b_3c_1) + b_2(c_1a_3 - c_3a_1) + c_2(a_1b_3 - a_3b_1)$

$= a_2b_1c_3 - a_2b_3c_1 + b_2c_1a_3 - b_2c_3a_1 + c_2a_1b_3 - c_2a_3b_1$

$= -a_1(b_2c_3 - b_3c_2) - b_1(c_2a_3 - c_3a_2) - c_1(a_2b_3 - a_3b_2)$

$\therefore\ \Delta_2 = -\Delta_1$

Thus $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = -\begin{vmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \\ a_1 & b_1 & c_1 \end{vmatrix}$

Similarly, $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = -\begin{vmatrix} b_1 & a_1 & c_1 \\ b_2 & a_2 & c_2 \\ b_3 & a_3 & c_3 \end{vmatrix} = \begin{vmatrix} b_1 & c_1 & a_1 \\ b_2 & c_2 & a_2 \\ b_3 & c_3 & a_3 \end{vmatrix}$

Alternative proof :

Since the interchange of two adjacent rows is the interchange of two subscripts and the interchange of two adjacent columns is the interchange of two letters, the sign of each term of the determinant is changed in either case. Hence $\Delta_1 = -\Delta_2$ and so on.

It can be shown that interchange of any two rows (or columns) changes the sign of the determinant without changing its absolute value.

While interchanging the p[th] row and q[th] row, observe that if p[th] row needs to move k steps down (each such move in nothing but one interchange of two adjacent rows), q[th] row having moved one step up in the last interchange of p[th] row, needs to move k – 1 steps up so that we need k + (k – 1) interchanges, which is odd. Other rows between p[th] row and q[th] row move one up and once down, thus retaining their position. Thus an odd number of interchanges yields a change of sign at the end.

(c) **If two rows, or two columns of a determinant are identical, then the value of the determinant is zero.**

The interchange of two identical rows or columns does not change the numerical value of the determinants. But by the property (*b*), the sign of determinant is changed by such interchange.

Hence $\Delta = -\Delta$, or $2\Delta = 0 \therefore \Delta = 0$.

A few more properties of determinants

(d) If every element of any row (or column) of a determinant is multiplied by a factor, the determinant is multiplied by the same factor.

Since we have $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1A_1 + b_1B_1 + c_1C_1$, and A_1, B_1, C_1

are respectively indedependent of a_1, b_1, c_1, we get

$$\begin{vmatrix} ka_1 & kb_1 & kc_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = ka_1A_1 + kb_1B_1 + kc_1C_1$$

$$= k(a_1A_1 + b_1B_1 + c_1C_1)$$

$$= k\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

$$= k\,(a_1A_1 + a_2A_2 + a_3A_3) = \begin{vmatrix} ka_1 & b_1 & c_1 \\ ka_2 & b_2 & c_2 \\ ka_3 & b_3 & c_3 \end{vmatrix}$$

Therefore **if every element of any row (or column) of a determinant has a common factor k, then k is a factor of the determinant and can be taken outside.**

(e) If every element of a row (or a column) of a determinant can be expressed as the sum of two numbers, then the determinant can be expressed as the sum of two determinants.

As we have $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1A_1 + b_1B_1 + c_1C_1$, and A_1, B_1, C_1 are respectively

independent of a_1, b_1, c_1 we get $\begin{vmatrix} a_1+\alpha_1 & b_1+\beta_1 & c_1+\gamma_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

$$= (a_1+\alpha_1)\,A_1 + (b_1+\beta_1)\,B_1 + (c_1+\gamma_1)\,C_1$$
$$= (a_1A_1 + b_1B_1 + c_1C_1) + (\alpha_1A_1 + \beta_1B_1 + \gamma_1C_1)$$

$$= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Similarly we can write $\begin{vmatrix} a_1-\alpha_1 & b_1 & c_1 \\ a_2-\alpha_2 & b_2 & c_2 \\ a_3-\alpha_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} - \begin{vmatrix} \alpha_1 & b_1 & c_1 \\ \alpha_2 & b_2 & c_2 \\ \alpha_3 & b_3 & c_3 \end{vmatrix}$

(f) A determinant remains unchanged by adding k times the elements of any row (or column) to the corresponding elements of any other row (or column), where k is any given number.

By the preceding properties of the determinant, we have

$$\begin{vmatrix} a_1+kb_1 & b_1 & c_1 \\ a_2+kb_2 & b_2 & c_2 \\ a_3+kb_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} kb_1 & b_1 & c_1 \\ kb_2 & b_2 & c_2 \\ kb_3 & b_3 & c_3 \end{vmatrix},$$

$$= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + k\begin{vmatrix} b_1 & b_1 & c_1 \\ b_2 & b_2 & c_2 \\ b_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

as the value of the second determinant is zero, since the first and the second columns are identical.

(g) The sum of the products of the elements of any row (or column) of a determinant and the cofactors of the corresponding elements of any other row (or column) of the determinant is zero.

It can be verified that $a_1A_1 + b_1B_2 + c_1C_2 = 0 = a_1B_1 + a_2B_2 + a_3B_3$ etc.

(h) If all the elements below the leading diagonal or above leading diagonal or except leading elements are zero, then the value of the determinant becomes equal to product of all the diagonal elements.

$$\text{e.g. : } \begin{vmatrix} a_1 & b_1 & c_1 \\ 0 & b_2 & c_2 \\ 0 & 0 & c_3 \end{vmatrix} = a_1b_2c_3 = \begin{vmatrix} a_1 & 0 & 0 \\ a_2 & b_2 & 0 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

(i) If the elements of a determinant D that involves x **are polynomials in x and if D vanishes for $x = a$, then $(x-a)$ is a factor of D. In other words, if two rows (or two columns) become identical for $x=a$ then $(x-a)$ is a factor of D.**

If r rows (r>1) become identical on substitution of a for x, then $(x-a)^{r-1}$ is a factor of D. (This can be proved by expansion of determinant and application of remainder theorem)

5.5 Some special types of Determinants

1. Symmetric determinant :

It is a determinant with $a_{ij} = a_{ji}$, i.e. the elements symmetrically situated about the main diagonal are equal.

Example

$$\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix} = abc + 2fgh - af^2 - bg^2 - ch^2$$

2. Skew symmetric determinant

It is a determinant with $a_{ij} = -a_{ji}$. This obviously renders all diagonal elements equal to zero (since $a_{ii} = -a_{ii} \Rightarrow a_{ii} = 0$) and all other elements symmetrically situated about the main diagonal, additive inverses of one another.

Note that the value of a **skew symmetric determinant of odd order is zero,**

$$\text{e.g. } \begin{vmatrix} 0 & b & -c \\ -b & 0 & a \\ c & -a & 0 \end{vmatrix} = 0$$

3. Circulant determinant

It is a determinant in which the elements of the rows (or columns) are in cyclic arrangement,

e.g. $\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} = -(a^3+b^3+c^3-3abc)$

The corresponding matrix $\begin{bmatrix} a & b & c \\ b & c & a \\ c & a & b \end{bmatrix}$ is also known as a circulant matrix.

*5.6 Product of Determinants

(Additional topic for interested students; not for examination)

Let |A| and |B| be two determinants of same order.

Let |C| = |A||B|. The product determinant |C| can be obtained by applying any one of the following methods :

(i) Row by Column product :

The (i-j)th element of |C| is obtained by the sum of the products of corresponding elements (inner product) of ith row of |A| and jth column of |B|.

(ii) Row by Row Product :

The (i-j)th element of |C| is obtained by the 'inner product' of ith row of |A| and jth row of |B|.

Similarly, (iii) **Column by Row** and (iv) **Column by Column** products are defined.

For example :

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \begin{vmatrix} -1 & 2 \\ -4 & 3 \end{vmatrix} = \begin{vmatrix} -1-8 & 2+6 \\ -3-16 & 6+12 \end{vmatrix} = \begin{vmatrix} -9 & 8 \\ -19 & 18 \end{vmatrix} = -10, \quad \text{by method (i)}$$

$$= \begin{vmatrix} -1+4 & -4+6 \\ -3+8 & -12+12 \end{vmatrix} = \begin{vmatrix} 3 & 2 \\ 5 & 0 \end{vmatrix} = -10, \quad \text{by method (ii)}$$

$$= \begin{vmatrix} -1+6 & -4+9 \\ -2+8 & -8+12 \end{vmatrix} = \begin{vmatrix} 5 & 5 \\ 6 & 4 \end{vmatrix} = -10, \quad \text{by method (iii)}$$

$$= \begin{vmatrix} -1-12 & 2+9 \\ -2-16 & 4+12 \end{vmatrix} = \begin{vmatrix} -13 & 11 \\ -18 & 16 \end{vmatrix} = -10, \quad \text{by method (iv)}$$

Let $|A^T|$ denote the determinant of A^T, the transpose of A, obtained by interchanging rows to columns and vice-versa. Then by the property of determinants $|A^T| = |A|$. The fact that $|A|\,|B| = |A^T|\,|B| = |A|\,|B^T| = |A^T||B^T|$ gives us the justification of the above methods of multiplication.

Warning :

We shall define product of two matrices A and B in a particular way by taking ith row of A and jth column of B. It should be clear that no product such as row by row, column by row or column by column exists in the c4ontext of matrices.

In case of determinants, other methods of multiplication besides the row by column method are some times helpful in obtaining factorisation of a given determinant.

5.7 Illustrative Examples :

Example 2

Find the minors and cofactors of the elements of the determinant $\begin{vmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \\ 1 & 4 & 2 \end{vmatrix}$

The determinant is of the form $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$

So $M_{11} = \begin{vmatrix} 1 & 3 \\ 4 & 2 \end{vmatrix} = 2 \times 1 - 3 \times 4 = -10$

$M_{12} = \begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} = 2 \times 2 - 3 \times 1 = 1$

$M_{13} = \begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = 2 \times 4 - 1 \times 1 = 7$

$M_{21} = \begin{vmatrix} 2 & 1 \\ 4 & 2 \end{vmatrix} = 2 \times 2 - 1 \times 4 = 0$

$M_{22} = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 \times 2 - 1 \times 1 = 1$

$M_{23} = \begin{vmatrix} 1 & 2 \\ 1 & 4 \end{vmatrix} = 1 \times 4 - 2 \times 1 = 2$

$M_{31} = \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = 2 \times 3 - 1 \times 1 = 5$

$M_{32} = \begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} = 1 \times 3 - 1 \times 2 = 1$

$M_{33} = \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = 1 \times 1 - 2 \times 2 = -3$

We know that the cofactor $C_{ij} = (-1)^{i+j} M_{ij}$

So, $C_{11} = (-1)^{1+1} M_{11} = -10$

$C_{12} = (-1)^{1+2} M_{12} = -1$

$C_{13} = (-1)^{1+3} M_{13} = 7$

$C_{21} = (-1)^{2+1} M_{21} = 0$

$C_{22} = (-1)^{2+2} M_{22} = 1$

$C_{23} = (-1)^{2+3} M_{23} = -2$

$C_{31} = (-1)^{3+1} M_{31} = 5$

$C_{32} = (-1)^{3+2} M_{32} = -1$

$C_{33} = (-1)^{3+3} M_{33} = -3$

Example 3

If $x + y + z = 0$, show that $\begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^3 & y^3 & z^3 \end{vmatrix} = 0$

$\begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^3 & y^3 & z^3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ x & y-x & z-x \\ x^3 & y^3-x^3 & z^3-x^3 \end{vmatrix}$ (Subtracting 1st column from the 2nd and 3rd column)

$= \begin{vmatrix} y-x & z-x \\ y^3-x^3 & z^3-x^3 \end{vmatrix} = (y-x)(z-x) \begin{vmatrix} 1 & 1 \\ x^2+xy+y^2 & x^2+zx+z^2 \end{vmatrix}$

$= (y-x)(z-x)(x^2 + zx + z^2 - x^2 - xy - y^2)$

$= (y-x)(z-x)\{x(z-y) + (z^2 - y^2)\}$

$= (y-x)(z-x)(z-y)(x + y + z) = 0 \ (\because x + y + z = 0)$

Example 4

Prove that $(a-1)$ is a factor of the determinant $\begin{vmatrix} a+1 & 2 & 3 \\ 3 & a+2 & 4 \\ 4 & 4 & a+3 \end{vmatrix}$.

If we put $a = 1$ in the given determinant, it becomes $\begin{vmatrix} 2 & 2 & 3 \\ 3 & 3 & 4 \\ 4 & 4 & 4 \end{vmatrix} = 0 \ (\because C_1 = C_2)$

On the other hand after expansion the given determinant is a cubic expression which reduces to zero if $a = 1$. Therefore, $a = 1$ is a root of the resulting cubic equation in a. So $(a-1)$ is a factor of the determinant.

Example 5

Prove without expanding that

(i) $$\begin{vmatrix} bc & a & a^2 \\ ca & b & b^2 \\ ab & c & c^2 \end{vmatrix} = \begin{vmatrix} 1 & a^2 & a^3 \\ 1 & b^2 & b^3 \\ 1 & c^2 & c^3 \end{vmatrix},$$

(ii) $$\begin{vmatrix} 1 & a & a^2-bc \\ 1 & b & b^2-ca \\ 1 & c & c^2-ba \end{vmatrix} = 0$$

(i) $$\begin{vmatrix} bc & a & a^2 \\ ca & b & b^2 \\ ab & c & c^2 \end{vmatrix} = \begin{vmatrix} \frac{1}{a}(abc) & \frac{a^2}{a} & \frac{a^3}{a} \\ \frac{1}{b}(abc) & \frac{b^2}{b} & \frac{b^3}{b} \\ \frac{1}{c}(abc) & \frac{c^2}{c} & \frac{c^3}{c} \end{vmatrix}$$

$$= abc \begin{vmatrix} \frac{1}{a} & \frac{a^2}{a} & \frac{a^3}{a} \\ \frac{1}{b} & \frac{b^2}{b} & \frac{b^3}{b} \\ \frac{1}{c} & \frac{c^2}{c} & \frac{c^3}{c} \end{vmatrix} = \frac{1}{abc}\, abc \begin{vmatrix} 1 & a^2 & a^3 \\ 1 & b^2 & b^3 \\ 1 & c^2 & c^3 \end{vmatrix} = \begin{vmatrix} 1 & a^2 & a^3 \\ 1 & b^2 & b^3 \\ 1 & c^2 & c^3 \end{vmatrix}$$

(ii) $$\begin{vmatrix} 1 & a & a^2-bc \\ 1 & b & b^2-ca \\ 1 & c & c^2-ba \end{vmatrix} = \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} - \begin{vmatrix} 1 & a & bc \\ 1 & b & ca \\ 1 & c & ba \end{vmatrix} \qquad \text{.............(1)}$$

of these two determinants, the second one is $\begin{vmatrix} 1 & a & bc \\ 1 & b & ca \\ 1 & c & ba \end{vmatrix} = \frac{1}{abc} \begin{vmatrix} a & a^2 & abc \\ b & b^2 & abc \\ c & c^2 & abc \end{vmatrix}$

$$= \frac{1}{abc}.abc \begin{vmatrix} a & a^2 & 1 \\ b & b^2 & 1 \\ c & c^2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix}$$

Substituting the value of second determinant on the right hand side of (1), we have

$$\begin{vmatrix} 1 & a & a^2-bc \\ 1 & b & b^2-ca \\ 1 & c & c^2-ba \end{vmatrix} = 0$$

Example 6

Without expanding find the value of the determinant $\begin{vmatrix} 17 & 58 & 97 \\ 19 & 60 & 99 \\ 18 & 59 & 98 \end{vmatrix}$

Subtracting the elements of the third row from the corresponding elements of the second row, and subtracting the element of the first row from the corresponding elements of the third row, we

have $\begin{vmatrix} 17 & 58 & 97 \\ 19 & 60 & 99 \\ 18 & 59 & 98 \end{vmatrix} = \begin{vmatrix} 17 & 58 & 97 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 0$ ($\because$ second and third rows are identical)

Example 7

Factorize the determinant

$\begin{vmatrix} x^3-a^3 & x^2 & x \\ b^3-a^3 & b^2 & b \\ c^3-a^3 & c^2 & c \end{vmatrix}$ without expanding.

we have $\begin{vmatrix} x^3-a^3 & x^2 & x \\ b^3-a^3 & b^2 & b \\ c^3-a^3 & c^2 & c \end{vmatrix} = \begin{vmatrix} x^3 & x^2 & x \\ b^3 & b^2 & b \\ c^3 & c^2 & c \end{vmatrix} - \begin{vmatrix} a^3 & x^2 & x \\ a^3 & b^2 & b \\ a^3 & c^2 & c \end{vmatrix}$

$$= xbc \begin{vmatrix} x^2 & x & 1 \\ b^2 & b & 1 \\ c^2 & c & 1 \end{vmatrix} - a^3 \begin{vmatrix} 1 & x^2 & x \\ 1 & b^2 & b \\ 1 & c^2 & c \end{vmatrix} = (xbc - a^3) \begin{vmatrix} x^2 & x & 1 \\ b^2 & b & 1 \\ c^2 & c & 1 \end{vmatrix}$$

(Interchanging the first column and the second column, then the second column and the third column in the second determinant).

$$= (xbc - a^3) \begin{vmatrix} x^2-b^2 & x-b & 0 \\ b^2-c^2 & b-c & 0 \\ c^2 & c & 1 \end{vmatrix}$$

(Subtracting the second row from the first, and the third row from the second)

$$= (xbc - a^3)(x - b)(b - c)\begin{vmatrix} x+b & 1 & 0 \\ b+c & 1 & 0 \\ c^2 & c & 1 \end{vmatrix}$$

$$= (xbc - a^3)(x - b)(b - c)\begin{vmatrix} x+b & 1 \\ b+c & 1 \end{vmatrix}$$

$$= (xbc - a^3)(x - b)(b - c)(x + b - b - c)$$

$$= (x - b)(x - c)(xbc - a^3)(b - c)$$

Example 8

Express $\begin{vmatrix} a^2 & 2ab & b^2 \\ b^2 & a^2 & 2ab \\ 2ab & b^2 & a^2 \end{vmatrix}$ in the form of a perfect square.

we have $\begin{vmatrix} a^2 & 2ab & b^2 \\ b^2 & a^2 & 2ab \\ 2ab & b^2 & a^2 \end{vmatrix} = \begin{vmatrix} a^2+2ab+b^2 & 2ab+b^2+a^2 & b^2+a^2+2ab \\ b^2 & a^2 & 2ab \\ 2ab & b^2 & a^2 \end{vmatrix}$

(Adding third and second rows to the first row)

$$= (a^2 + 2ab + b^2)\begin{vmatrix} 1 & 1 & 1 \\ b^2 & a^2 & 2ab \\ 2ab & b^2 & a^2 \end{vmatrix}$$

$$= (a + b)^2\begin{vmatrix} 1 & 0 & 0 \\ b^2 & a^2-b^2 & 2ab-b^2 \\ 2ab & b^2-2ab & a^2-2ab \end{vmatrix}$$

(Subtracting first column from the second, and first column from the third)

$$= (a + b)^2\begin{vmatrix} a^2-b^2 & 2ab-b^2 \\ b^2-2ab & a^2-2ab \end{vmatrix}$$

$$= (a + b)^2 \{(a^2 - b^2)(a^2 - 2ab) + (b^2 - 2ab)^2\}$$

$$= (a + b)^2 \{a^4 - 2a^3b - a^2b^2 + 2ab^3 + b^4 - 4ab^3 + 4a^2b^2\}$$

$= (a+b)^2\{a^4 - 2a^3b + 3a^2b^2 - 2ab^3 + b^4\}$

$= (a+b)^2\{a^4 + a^2b^2 + b^4 - 2a^3b - 2ab^3 + 2a^2b^2\}$

$= (a+b)^2(a^2 - ab + b^2)^2 = \{(a+b)(a^2 - ab + b^2)\}^2$

$= (a^3 + b^3)^2$

Alternative method using product of determinants :

$$\begin{vmatrix} a^2 & 2ab & b^2 \\ b^2 & a^2 & 2ab \\ 2ab & b^2 & a^2 \end{vmatrix} = \begin{vmatrix} a & b & 0 \\ 0 & a & b \\ b & 0 & a \end{vmatrix}\begin{vmatrix} a & 0 & b \\ b & a & 0 \\ 0 & b & a \end{vmatrix} \text{ (Row by Row product)}$$

$$= \begin{vmatrix} a & b & 0 \\ 0 & a & b \\ b & 0 & a \end{vmatrix}\begin{vmatrix} a & b & 0 \\ 0 & a & b \\ b & 0 & a \end{vmatrix}$$ (Interchanging row and column in the second determinant)

$$= \begin{vmatrix} a & b & 0 \\ 0 & a & b \\ b & 0 & a \end{vmatrix}^2 = [a.a^2 - b(-b^2)]^2$$

$$= (a^3 + b^3)^2$$

Example 9

If $A + B + C = \pi$, show that $\begin{vmatrix} -1 & \cos C & \cos B \\ \cos C & -1 & \cos A \\ \cos B & \cos A & -1 \end{vmatrix} = 0$

We have $\begin{vmatrix} -1 & \cos C & \cos B \\ \cos C & -1 & \cos A \\ \cos B & \cos A & -1 \end{vmatrix} = \begin{vmatrix} -1 & \cos C & \cos B \\ 0 & \cos^2 C - 1 & \cos C.\cos B + \cos A \\ 0 & \cos B.\cos C + \cos A & \cos^2 B - 1 \end{vmatrix}$

(Adding cos C times the first row to the second row, and cos B times the first row to the third row).

$$= -\begin{vmatrix} -\sin^2 C & \cos C.\cos B + \cos A \\ \cos B.\cos C + \cos A & -\sin^2 B \end{vmatrix}$$

$= -\sin^2 B \sin^2 C + (\cos B.\cos C + \cos A)^2$

$= -\sin^2 B.\sin^2 C + \{\cos(\pi - \overline{B+C}) + \cos B.\cos C\}^2 \ (\because A + B + C = \pi)$

$= -\sin^2 B.\sin^2 C + \{-\cos(B+C) + \cos B.\cos C\}^2$

$= -\sin^2 B.\sin^2 C + \sin^2 B.\sin^2 C = 0$ (As $\cos(B+C) = \cos B.\cos C - \sin B.\sin C$)

EXERCISES 5 (a)

1. Evaluate the following determinants.

(i) $\begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix}$ (ii) $\begin{vmatrix} 2 & -3 \\ 1 & -4 \end{vmatrix}$ (iii) $\begin{vmatrix} \sec\theta & \tan\theta \\ \tan\theta & \sec\theta \end{vmatrix}$

(iv) $\begin{vmatrix} 0 & x \\ 2 & 0 \end{vmatrix}$ (v) $\begin{vmatrix} 1 & \omega \\ -\omega & \omega \end{vmatrix}$ (vi) $\begin{vmatrix} 4 & -1 \\ 3 & 2 \end{vmatrix}$

(vii) $\begin{vmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{vmatrix}$ (viii) $\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}$ (ix) $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$

(x) $\begin{vmatrix} 2 & 3 & 1 \\ 0 & 0 & 0 \\ -1 & 2 & 0 \end{vmatrix}$ (xi) $\begin{vmatrix} 1 & x & y \\ 0 & \sin x & \sin y \\ 0 & \cos x & \cos y \end{vmatrix}$ (xii) $\begin{vmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 3 & 4 & 5 \end{vmatrix}$

(xiii) $\begin{vmatrix} 0.2 & 0.1 & 3 \\ 0.4 & 0.2 & 7 \\ 0.6 & 0.3 & 2 \end{vmatrix}$ (xiv) $\begin{vmatrix} 1 & \omega & \omega^2 \\ \omega & \omega^2 & 1 \\ \omega^2 & 1 & \omega \end{vmatrix}$ (xv) $\begin{vmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{vmatrix}$

(xvi) $\begin{vmatrix} -6 & 0 & 0 \\ 3 & -5 & 7 \\ 2 & 8 & 11 \end{vmatrix}$ (xvii) $\begin{vmatrix} 1 & 0 & 0 \\ 2 & 3 & 5 \\ 4 & 1 & 3 \end{vmatrix}$ (xviii) $\begin{vmatrix} -18 & 17 & 19 \\ 3 & 0 & 0 \\ -14 & 5 & 2 \end{vmatrix}$

2. State true or false.

(i) If the first and second rows of a determinant be interchanged then the sign of the determinant is changed.

(ii) If first and third rows of a determinant be interchanged then the sign of the determinant does not change.

(iii) If in a third order determinant first row be changed to second column, second row to first column and third row to third column, then the value of the determinant does not change.

(iv) A row and a column of a determinant can have two or more common elements.

(v) The minor and the cofactor of the element a_{32} of a determinant of third order are equal.

(vi) $\begin{vmatrix} 3 & 1 & 3 \\ 0 & 4 & 0 \\ 1 & 3 & 1 \end{vmatrix} = 0$ (vii) $\begin{vmatrix} 6 & 4 & 2 \\ 4 & 0 & 7 \\ 5 & 3 & 4 \end{vmatrix} = \begin{vmatrix} 6 & 4 & 5 \\ 4 & 0 & 3 \\ 2 & 7 & 4 \end{vmatrix}$ (viii) $\begin{vmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 4 & 2 & 3 \\ 7 & 5 & 6 \\ 3 & 1 & 2 \end{vmatrix}$

3. Fill in the blanks with appropriate answer from the brackets.

(i) The value of $\begin{vmatrix} 0 & 8 & 0 \\ 25 & 520 & 25 \\ 1 & 410 & 0 \end{vmatrix} =$ ______ $(0, 25, 200, -250)$

(ii) If ω is the cube root of unity, then $\begin{vmatrix} 1 & \omega & \omega^2 \\ \omega & \omega^2 & 1 \\ \omega^2 & 1 & \omega \end{vmatrix} =$ ______ $(1, 0, \omega, \omega^2)$

(iii) The value of the determinant $\begin{vmatrix} 1 & a & b+c \\ 1 & b & c+a \\ 1 & c & a+b \end{vmatrix} =$ ______ $(a+b=c, (a+b+c)^2, 0, 1+a+b+c)$

(iv) If $\begin{vmatrix} a & b & c \\ b & a & b \\ x & b & c \end{vmatrix} = 0$, then $x =$ ______ $(a, b, c, a+b+c)$

(v) $\begin{vmatrix} a_1+a_2 & a_3+a_4 & a_5 \\ b_1+b_2 & b_3+b_4 & b_5 \\ c_1+c_2 & c_3+c_4 & c_5 \end{vmatrix}$ can be expressed at the most as ______ different third order determinants. $(1, 2, 3, 4)$

(vi) The minimum value of $\begin{vmatrix} \sin x & \cos x \\ -\cos x & 1+\sin x \end{vmatrix}$ is ______. $(-1, 0, 1, 2)$

(vii) The determinant $\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{vmatrix}$ is not equal to ______.

$$\left(\begin{vmatrix} 2 & 1 & 1 \\ 2 & 2 & 3 \\ 2 & 3 & 6 \end{vmatrix}, \begin{vmatrix} 2 & 1 & 1 \\ 3 & 2 & 3 \\ 4 & 3 & 6 \end{vmatrix}, \begin{vmatrix} 1 & 2 & 1 \\ 1 & 5 & 3 \\ 1 & 9 & 6 \end{vmatrix}, \begin{vmatrix} 3 & 1 & 1 \\ 6 & 2 & 3 \\ 10 & 3 & 6 \end{vmatrix} \right)$$

(viii) With 4 different elements we can construct —— number of different determinants of order 2. (1, 6, 8, 24)

4. Solve the following :

(i) $\begin{vmatrix} 4 & x+1 \\ 3 & x \end{vmatrix} = 5$

(ii) $\begin{vmatrix} x & a & a \\ m & m & m \\ b & x & b \end{vmatrix} = 0$

(iii) $\begin{vmatrix} 7 & 6 & x \\ 2 & x & 2 \\ x & 3 & 7 \end{vmatrix} = 0$

(iv) $\begin{vmatrix} 0 & x-a & x-b \\ x+a & 0 & x-c \\ x+b & x+c & 0 \end{vmatrix} = 0$

(v) $\begin{vmatrix} x+a & b & c \\ b & x+c & a \\ c & a & x+b \end{vmatrix} = 0$

(vi) $\begin{vmatrix} 1+x & 1 & 1 \\ 1 & 1+x & 1 \\ 1 & 1 & 1+x \end{vmatrix} = 0$

(vii) $\begin{vmatrix} 1 & 4 & 20 \\ 1 & -2 & 5 \\ 1 & 2x & 5x^2 \end{vmatrix} = 0$

(viii) $\begin{vmatrix} x+1 & \omega & \omega^2 \\ \omega & x+\omega^2 & 1 \\ \omega^2 & 1 & x+\omega \end{vmatrix} = 0$

(ix) $\begin{vmatrix} 2 & 2 & x \\ -1 & x & 4 \\ 1 & 1 & 1 \end{vmatrix} = 0$

(x) $\begin{vmatrix} x & 1 & 3 \\ 1 & x & 1 \\ 3 & 6 & 3 \end{vmatrix} = 0$

5. **Evaluate the following .**

(i) $\begin{vmatrix} 2 & 3 & 4 \\ 1 & -1 & 3 \\ 4 & 1 & 10 \end{vmatrix}$

(ii) $\begin{vmatrix} x & 1 & 2 \\ y & 3 & 1 \\ z & 2 & 2 \end{vmatrix}$

(iii) $\begin{vmatrix} x & 1 & -1 \\ 2 & y & 1 \\ 3 & -1 & z \end{vmatrix}$

(iv) $\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}$

(v) $\begin{vmatrix} 8 & -1 & -8 \\ -2 & -2 & -2 \\ 3 & -5 & -3 \end{vmatrix}$

(vi) $\begin{vmatrix} \sin^2\theta & \cos^2\theta & 1 \\ \cos^2\theta & \sin^2\theta & 1 \\ -10 & 12 & 2 \end{vmatrix}$

(vii) $\begin{vmatrix} -1 & 3 & 2 \\ 1 & 3 & 2 \\ 1 & -3 & -1 \end{vmatrix}$

(viii) $\begin{vmatrix} 11 & 23 & 31 \\ 12 & 19 & 14 \\ 6 & 9 & 7 \end{vmatrix}$

(ix) $\begin{vmatrix} 37 & -3 & 11 \\ 16 & 2 & 3 \\ 5 & 3 & -2 \end{vmatrix}$ (x) $\begin{vmatrix} 2 & -3 & 4 \\ -4 & 2 & -3 \\ 11 & -15 & 20 \end{vmatrix}$

6. Show that $x = 1$ is a solution of $\begin{vmatrix} x+1 & 3 & 5 \\ 2 & x+2 & 5 \\ 2 & 3 & x+4 \end{vmatrix} = 0$

7. Show that $(a+1)$ is a factor of $\begin{vmatrix} a+1 & 2 & 3 \\ 1 & a+1 & 3 \\ 3 & -6 & a+1 \end{vmatrix}$

8. Show that $\begin{vmatrix} a_1 & b_1 & -c_1 \\ -a_2 & -b_2 & c_2 \\ a_3 & b_3 & -c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

9. Prove the following .

(i) $\begin{vmatrix} a & b & c \\ x & y & z \\ p & q & r \end{vmatrix} = \begin{vmatrix} y & b & q \\ x & a & p \\ z & c & r \end{vmatrix} = \begin{vmatrix} x & y & z \\ p & q & r \\ a & b & c \end{vmatrix}$

(ii) $\begin{vmatrix} 1+a & 1 & 1 \\ 1 & 1+b & 1 \\ 1 & 1 & 1+c \end{vmatrix} = abc\left(1+\frac{1}{a} + \frac{1}{b} + \frac{1}{c}\right)$

(iii) $\begin{vmatrix} b+c & c+a & a+b \\ q+r & r+p & p+q \\ y+z & z+x & x+y \end{vmatrix} = 2\begin{vmatrix} a & b & c \\ p & q & r \\ x & y & z \end{vmatrix}$

(iv) $\begin{vmatrix} (a+1)(a+2) & a+2 & 1 \\ (a+2)(a+3) & a+3 & 1 \\ (a+3)(a+4) & a+4 & 1 \end{vmatrix} = -2$

(v) $\begin{vmatrix} a+d & a+d+k & a+d+c \\ c & c+b & c \\ d & d+k & d+c \end{vmatrix} = abc$

(vi) $\begin{vmatrix} 1 & 1 & 1 \\ b+c & c+a & a+b \\ b^2+c^2 & c^2+a^2 & a^2+b^2 \end{vmatrix} = (b-c)(c-a)(a-b)$

(vii) $\begin{vmatrix} a & a^2 & a^3 \\ b & b^2 & b^3 \\ c & c^2 & c^3 \end{vmatrix} = abc\,(a-b)\,(b-c)\,(c-a)$

(viii) $\begin{vmatrix} b+c & a & c \\ b & c+a & b \\ c & c & a+b \end{vmatrix} = 4abc$

(ix) $\begin{vmatrix} b^2+c^2 & ab & ac \\ ab & c^2+a^2 & bc \\ ca & cb & a^2+b^2 \end{vmatrix} = 4a^2b^2c^2$

(x) $\begin{vmatrix} \text{a} & b & c \\ a^2 & b^2 & c^2 \\ bc & ca & ab \end{vmatrix} = (b-c)\,(c-a)\,(a-b)\,(bc+ca+ab)$

(xi) $\begin{vmatrix} a-b-\text{c} & 2a & 2a \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix} = (a+b+c)^3$

(xii) $\begin{vmatrix} (v+w)^2 & u^2 & u^2 \\ v^2 & (w+u)^2 & v^2 \\ w^2 & w^2 & (u+v)^2 \end{vmatrix} = 2\,uvw\,(u+v+w)^3$

10. Factorize the following :

(i) $\begin{vmatrix} x+a & b & c \\ b & x+c & a \\ c & a & x+b \end{vmatrix}$

(ii) $\begin{vmatrix} a & b & c \\ b+c & c+a & a+b \\ a^2 & b^2 & c^2 \end{vmatrix}$

(iii) $\begin{vmatrix} x & 2 & 3 \\ 1 & x+1 & 3 \\ 1 & 4 & x \end{vmatrix}$

11. Show that by eliminating α and β from the equations $a_i\alpha + b_i\beta + c_i = 0$, i = 1, 2, 3 we get

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0$$

12. Prove the following :

(i) $\begin{vmatrix} 1 & bc & a(b+c) \\ 1 & ca & b(c+a) \\ 1 & ab & c(a+b) \end{vmatrix} = 0$

(ii) $\begin{vmatrix} x+4 & 2x & 2x \\ 2x & x+4 & 2x \\ 2x & 2x & x+4 \end{vmatrix} = (5x+4)(4-x)^2$

(iii) $\begin{vmatrix} \sin\alpha & \cos\alpha & \cos(\alpha+\delta) \\ \sin\beta & \cos\beta & \cos(\beta+\delta) \\ \sin\gamma & \cos\gamma & \cos(\gamma+\delta) \end{vmatrix} = 0$

(iv) $\begin{vmatrix} 1 & x & x^2 \\ x^2 & 1 & x \\ x & x^2 & 1 \end{vmatrix} = (1-x^3)^2$

13. Prove that the points (x_1, y_1), (x_2, y_2), (x_3, y_3) are collinear if $\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0$

14. If $A + B + C = \pi$, prove that $\begin{vmatrix} \sin^2 A & \cot A & 1 \\ \sin^2 B & \cot B & 1 \\ \sin^2 C & \cot C & 1 \end{vmatrix} = 0$

15. Eliminate x, y, z from $a = \frac{x}{y-z}$, $b = \frac{y}{z-x}$, $c = \frac{z}{x-y}$

16. Given the equations $x = cy + bz$, $y = az + cx$ and $z = bx + ay$ whrer x, y and z are not all zero, prove that $a^2 + b^2 + c^2 + 2abc = 1$ by determinant method.

17. If $ax + hy + g = 0$, $hx + by + f = 0$ and $gx + fy + c = \lambda$ find the value of λ in the form of a determinant.

5.8 Consistency, Inconsistency and number of solutions of a system of linear equations.

Let us consider the three linear simultaneous equations.

$$a_1x + b_1y + c_1z = k_1$$
$$a_2x + b_2y + c_2z = k_2$$
$$a_3x + b_3y + c_3z = k_3.$$

The determinant of the matrix of coefficients is given by

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \neq 0.$$

and let A_1, A_2, A_3 ... denote the respective cofactors of a_1, a_2, a_3, ...in Δ. Multiplying the equations by A_1, A_2, A_3 respectively and adding, we have $(a_1A_1 + a_2A_2 + a_3A_3)x + (b_1A_1 + b_2A_2 + b_3A_3)y + (c_1A_1 + c_2A_2 + c_3A_3)z = k_1A_1 + k_2A_2 + k_3A_3$

Since $b_1A_1 + b_2A_2 + b_3A_3 = 0$ and $c_1A_1 + c_2A_2 + c_3A_3 = 0$, we have

$(a_1A_1 + a_2A_2 + a_3A_3)x = k_1A_1 + k_2A_2 + k_3A_3$

$\Rightarrow \Delta x = \Delta_x$

where $\Delta_x = \begin{vmatrix} k_1 & b_1 & c_1 \\ k_2 & b_2 & c_2 \\ k_3 & b_3 & c_3 \end{vmatrix}$

(refer expansion of a determinant using cofactors and property (g) of a diterminant)

Similarly we can obtain

$\Delta y = \Delta_y$ and $\Delta z = \Delta_z$;

where $\Delta_y = \begin{vmatrix} a_1 & k_1 & c_1 \\ a_2 & k_2 & c_2 \\ a_3 & k_3 & c_3 \end{vmatrix}$, $\Delta_z = \begin{vmatrix} a_1 & b_1 & k_1 \\ a_2 & b_2 & k_2 \\ a_3 & b_3 & k_3 \end{vmatrix}$.

Now three exclusive and exhaustive cases can arise

(i) $\Delta \neq 0$

(ii) $\Delta = 0$ and $\Delta_x, \Delta_y, \Delta_z$ are not all zero

(iii) $\Delta, \Delta_x, \Delta_y, \Delta_z$ are all zero.

Under case (i)

$x = \frac{\Delta_x}{\Delta}, y = \frac{\Delta_y}{\Delta}$ and $z = \frac{\Delta_z}{\Delta}$

which is a unique solution such a system.

Such a system is called a **consistent system**

Example -10

Test whether the following system has a unique solution

$2x + 3y - z = 9$

$x + 2y + z = 4$

$5x - y + 2z = 1$

Solution

$$\text{Here } \Delta = \begin{vmatrix} 2 & 3 & -1 \\ 1 & 2 & 1 \\ 5 & -1 & 2 \end{vmatrix} = 30$$

Since $\Delta \neq 0$, The system is consistent and has a unique solution.

(In fact, the solution is $x = 1$, $y = 2$, $z = -1$)

Under case (ii)

Division by zero being meaningless, we cannot get values of $\frac{\Delta_x}{\Delta}, \frac{\Delta_y}{\Delta}, \frac{\Delta_z}{\Delta}$ and hence the system lacks a solution and is called **inconsistent**.

Example - 11

Examine solvability of the system :

$x + y - z = 1$
$2x + y + z = 7$
$3x + 3y - 3z = 2$

Solution

$$\Delta = \begin{vmatrix} 1 & 1 & -1 \\ 2 & 1 & 1 \\ 3 & 3 & -3 \end{vmatrix} = 0,\ \Delta_x = \begin{vmatrix} 1 & 1 & -1 \\ 7 & 1 & 1 \\ 2 & 3 & -3 \end{vmatrix} = -38 \neq 0$$

$\Rightarrow$ System is **inconsistent** and hence not solvable.

(The inconsistency of the system can be immediately realised from $x + y - z = 1$ and $3x + 3y - 3z = 2$

Under case (iii)

$\Delta x = \Delta_x, \Delta y = \Delta_y, \Delta z = \Delta_z$

and in this case the system of equations is not inconsistent as both the sides in each of the above involve zeros. However, we do not get any definite solution since determinant of *x, y, z* involves division by Δ which is zero.

Example - 12

$2x - y = 5$
$6x - 3y = 15$

Solution

$$\Delta = \begin{vmatrix} 2 & -1 \\ 6 & -3 \end{vmatrix} = 0,\ \Delta_x = \begin{vmatrix} 5 & -1 \\ 15 & -3 \end{vmatrix} = 0$$

$$\Delta_y = \begin{vmatrix} 2 & 5 \\ 6 & 15 \end{vmatrix} = 0$$

$\therefore$ The system has infinite number of solutions.

Infact, by giving any particular value of x either equation, we can get a corresponding value of y and in this way, obtain infinite number of solutions.

Note : An interesting case occurs when we come across systems like

$$a_1x+b_1y+c_1z=0$$
$$a_2x+b_2y+c_2z=0$$
$$a_3x+b_3y+c_3z=0$$

$$\text{with } \Delta=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}=0$$

Such a system as above, with all zeros in RHS is known as a **homogeneous linear system** and $x = y = z = 0$ is an obvious solution.

But since $\Delta = 0$ and obviously $\Delta_x = \Delta_y = \Delta_z = 0$, the system comes under case (iii), thus admitting indefinite number of solutions.

Out of these solutions, the obvious solution $x = y = z = 0$ is called a **trivial solution**. But a solution in which not all of x, y, z are zero; is called a **nontrivial solution**.

Note that **a nontrivial solution of a homogeneous system occurs only when the determinant of coefficients $\Delta = 0$** ,

otherwise (if $\Delta \neq 0$) it comes under case (i) having $x = y = z = 0$ as the unique solution.

Take an example :

Example - 13

Examine whether the following system has any nontrivial solution. If so, find one.

$$x + 2y + z = 0$$
$$3x + 5y + 2z = 0$$
$$4x + 3y - z = 0$$

Solution

In this case it is easy to check that

$$\Delta_x = \Delta_y = \Delta_z = 0$$

$$\text{Also } \Delta=\begin{vmatrix} 1 & 2 & 1 \\ 3 & 5 & 2 \\ 4 & 3 & -1 \end{vmatrix}=0$$

So the system has nontrivial solution apart from the trivial one, $x = y = z = 0$.

Adding 1st and 3rd equation, we get

$5x + 5y = 0$, which gives

$y = -x$

Putting $y = -x$, the system reduces to

$-x + z = 0$

$-2x + 2z = 0$

$x - z = 0$, which are one and the same, i.e. $z = x$.

Every non zero value of x gives rise to a nontrivial solution, e.g.

$(x, y, z) = (1, -1, 1)$ and in general $(k, -k, k)$, $k \in R$ constitute the infinite set of nutrivial solutions.

N.B.(i) Though we have discussed 2 or 3 equations in 2 or 3 unknowns, the discussions apply to n number of equations in n unknowns for $n > 2$.

(ii) Solving linear equations by the above method is known as **Cramer's Rule.**

5.9 Inverse of a square matrix

We have already discussed inverse of a square matrix in the previous chapter. Now we shall introduce the concept of adjoint of a square matrix and discuss another method of finding the inverse. We shall also discuss precisely under what condition inverse of square matrix exists we begin with the definition

Singular matrix

A square matrix A is called **singular if det A = 0**. **If det A ≠ 0** the matrix A is called a **non-singular matrix** or a **regular matrix**.

Adjoint of a matrix :

If $A = [a_{ij}]_{n \times n}$ is a square matrix, then the transpose of the matrix $[A_{ij}]_{n \times n}$ of which the elements are cofactors of the corresponding element in $|A|$, is called the **adjoint or adjugate** of A. It is denoted by ***adj* A**.

Thus $adj\, A = [A_{ji}]_{n \times n}$

If $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, then $adj\, A = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix}$ where

A_{ij} is the cofactor of a_{ij} in $|A|$.

Theorem 1 : If A is a square matrix, then $A.(adj A) = |A|\, I = (adj A)\, A$

Proof : Let $A = [a_{ij}]_{n \times n}$. The element on the ith row and j th column of $A\,(adj\, A)$ is equal to $a_{i1} A_{j1} + a_{i2} A_{j2} + \ldots\ldots + a_{in} A_{jn} = |A|$ or zero according as $i = j$ or $i \neq j$. (ref. property of determinant)

$$\therefore A\ (adj A) = \begin{bmatrix} |A| & 0 & 0 & .... & 0 \\ 0 & |A| & 0 & .... & 0 \\ ... & ... & ... & .... & ... \\ 0 & 0 & 0 & .... & |A| \end{bmatrix} = |A|\ I$$

Similarly $(adj A).A = |A|\ I$

Hence the theorem follows.

Note : The **product of a singular matrix and its adjoint is zero matrix, since in this case $|A| = 0$**

Theorem 2 :

Let A be a square matrix. Then A^{-1} **exists if A is nonsingular and the inverse is given by** $A^{-1} = \frac{1}{|A|}\ adj\ A$

Proof : We know that

$A.\ (adj A) = (adj\ A)\ A = |A|.\ I$ since $|A| \neq 0$,

$$\Rightarrow A.\left(\frac{1}{|A|} adj A\right) = \left(\frac{1}{|A|} adj A\right).\ A = I$$

$$\Rightarrow A^{-1} = \frac{1}{|A|}\ adj\ A.$$

5.10 System of linear equations and solution (Matrix method)

Suppose we have the following system of equations :

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned} \qquad (1)$$

We write $A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$, $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $B = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$ (2)

The equation (1) in the matrix form becomes $AX = B$.

Theorem 3 :

The system of equations (1) has a unique solution if and only if the matrix A is non singular; and in this case the matrix solution of (1) is given by $X = \frac{1}{|A|}\ (adj\ A).\ B$.

Proof. The system given by (1) is $AX = B$

Since A is nonsingular, by theorem 2, A^{-1} exists. Hence $A^{-1}(AX) = A^{-1}B$

$$\Rightarrow (A^{-1}A)X = IX = X = A^{-1}B \Rightarrow X = \frac{AdjA}{|A|} . B.$$

5.11 Illustrative Examples

Example 14

Find the adjoint of the matrix $\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & 3 & 1 \end{bmatrix}$

The given matrix is $A = (a_{ij})_{3\times 3}$
A_{ij} and M_{ij} respectively are cofactor and minor of a_{ij}.

Thus $A_{11} = (-1)^{1+1} M_{11} = \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} = -5$

$$A_{12} = (-1)^{1+2} M_{12} = -\begin{vmatrix} 2 & 2 \\ 1 & 1 \end{vmatrix} = 0$$

$$A_{13} = (-1)^{1+3} M_{13} = \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = 5$$

$$A_{21} = (-1)^{2+1} M_{21} = -\begin{vmatrix} 1 & 1 \\ 3 & 1 \end{vmatrix} = 2$$

$$A_{22} = (-1)^{2+2} M_{22} = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0$$

$$A_{23} = (-1)^{2+3} M_{23} = -\begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} = -2$$

$$A_{31} = (-1)^{3+1} M_{31} = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1$$

$$A_{32} = (-1)^{3+2} M_{32} = -\begin{vmatrix} 1 & 1 \\ 2 & 2 \end{vmatrix} = 0$$

$$A_{33} = (-1)^{3+3} M_{33} = \begin{vmatrix} 1 & 1 \\ 2 & 1 \end{vmatrix} = -1$$

Adjoint of A i.e. Adj A = The transpose of $[A_{ij}]_{3\times 3}$

$$= \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix} = \begin{bmatrix} -5 & 2 & 1 \\ 0 & 0 & 0 \\ 5 & -2 & -1 \end{bmatrix}$$

Example 15

Find the inverse of the matrix $\begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$.

The given matrix is $A = [a_{ij}]_{2\times 2}$

$$|A| = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} = -5 \neq 0$$

So the matrix is invertible.

Now $A_{11} = (-1)^{1+1}(1) = 1$ $\quad A_{21} = (-1)^{2+1}(2) = -2$

$A_{12} = (-1)^{1+2}\ (3) = -3$. $\quad A_{22} = (-1)^{2+2}(1) = 1$

$$\text{adj } A = \begin{bmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ -3 & 1 \end{bmatrix}.$$

$$\text{So } A^{-1} = \frac{adj A}{|A|} = -\frac{1}{5}\begin{bmatrix} 1 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{5} & \frac{2}{5} \\ \frac{3}{5} & -\frac{1}{5} \end{bmatrix}.$$

Example 16

Solve the equation $x + 2y = 3$, $3x + y = 4$ by matrix method.

The given system of equations is of the form AX = B.

$$\text{Where } A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}, B = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, X = \begin{bmatrix} x \\ y \end{bmatrix}$$

$|A| = -5 \neq 0$. So A^{-1} exists.

$$\text{Now } AX = B \Rightarrow X = A^{-1}B = \begin{bmatrix} -\frac{1}{5} & \frac{2}{5} \\ \frac{3}{5} & -\frac{1}{5} \end{bmatrix}\begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$ [A^{-1} is taken from Example 17]

$$= \begin{bmatrix} -\frac{3}{5} + \frac{8}{5} \\ \frac{9}{5} - \frac{4}{5} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$\Rightarrow x = y = 1$.

Example 17

Solve the system $x + 2y + 3z = 8$, $2x + y + z = 8$, $x + y + 2z = 6$.

The given system is of the form AX = B.

where $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 8 \\ 8 \\ 6 \end{bmatrix}$, $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

Now $|A| = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{vmatrix} = -2 \neq 0.$

Here $A_{11} = (-1)^{1+1} \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1, A_{12} = (-1)^{1+2} \begin{vmatrix} -2 & 1 \\ 1 & 2 \end{vmatrix} = -3.$

$$A_{13} = (-1)^{1+3} \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1, A_{21} = (-1)^{2+1} \begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} = -1.$$

$$A_{22} = (-1)^{2+2} \begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix} = -1, A_{23} = (-1)^{2+3} \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = 1$$

$$A_{31} = (-1)^{3+1} \begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = -1, A_{32} = -(-1)^{3+2} = \begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} = 5$$

$$A_{33} = (-1)^{3+3} \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = -3$$

$$adj\ A = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ -3 & -1 & 5 \\ 1 & 1 & -3 \end{bmatrix}$$

The solution of $AX = B$ is $X = A^{-1} B$

$$= -\frac{1}{2} \begin{bmatrix} 1 & -1 & -1 \\ -3 & -1 & 5 \\ 1 & 1 & -3 \end{bmatrix} \begin{bmatrix} 8 \\ 8 \\ 6 \end{bmatrix}$$

$$= -\frac{1}{2} \begin{bmatrix} 8 & -8 & -6 \\ -24 & -8 & +30 \\ 8 & +8 & -18 \end{bmatrix}$$

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

$$\therefore \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

$\Rightarrow x = 3, y = 1, z = 1$

EXERCISES 5 (b)

1. Write the number of solutions of the following system of equations.

(i) $x - 2y = 0$

(ii) $x - y = 0$ and $2x - 2y = 1$

(iii) $2x + y = 2$ and $-x - \frac{1}{2}y = 3$

(iv) $3x + 2y = 1$ and $x + 5y = 6$

(v) $2x + 3y + 1 = 0$ and $x - 3y - 4 = 0$

(vi) $x + y + z = 1$
$x + y + z = 2$
$2x + 3y + z = 0$

(vii) $x + 4y - z = 0$
$3x - 4y - z = 0$
$x - 3y + z = 0$

(viii) $x + y - z = 0$
$3x - y - z = 0$
$x - 3y + z = 0$

(ix) $a_1x + b_1y + c_1z = 0$
$a_2x + b_2y + c_2z = 0$
$a_3x + b_3y + c_3z = 0$
and $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0$

2. Show that the following system is inconsistent.

$(a - b)x + (b - c)y + (c - a)z = 0$

$(b - c)x + (c - a)y + (a - b)z = 0$

$(c - a)x + (a - b)y + (b - c)z = 1$

3.(i) The system of equations

$x+2y+3z = 4$

$2x + 3y + 4z = 5$

$3x + 4y + 5z = 6$ has

(a) infinitely many solutions (b) no solution

(c) a unique solution (d) none of the three

(ii) If the system of equations

$2x + 5y + 8z = 0$

$x + 4y + 7z = 0$

$6x + 9y - \lambda z = 0$

has a nontrivial solution, then λ is equal to

(a) 12 (b) -12

(c) 0 (d) none of the three

(iii) The system of linear equations

$x+y+z = 2$

$2x + y - z = 3$

$3x + 2y + kz = 4$

has a unique solution if.

(a) $k \neq 0$ (b) $-1 < k < 1$

(c) $-2 < k < 2$ (d) $k = 0$

(iv) The equations

$x+y+z=6$

$x+2y+3z = 10$

$x+2y+mz = n$

give infinite number of values of the triplet (x, y, z) if

(a) $m = 3, n \in R$ (b) $m = 3, n \neq 10$

(c) $m = 3, n = 10$ (d) none of the three

(v) The system of equations

$2x - y + z = 0$

$x - 2y + z = 0$

$\lambda x - y + 2z = 0$

has infinite number of nontrivial solutions for

(a) $\lambda = 1$ (b) $\lambda = 5$

(c) $\lambda = -5$ (d) no real value of λ

*(vi) The system of equations

$a_1x + b_1y + c_1z = 0$

$a_2x + b_2y + c_2z = 0$

$a_3x + b_3y + c_3z = 0$

with $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0$ has

(a) more than two solutions (b) one trivial and one nontrivial solutions

(c) No solution (d) only trivial solutions

4. Can the inverses of the following matrices be found ?

(i) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ (iv) $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

(v) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

5. Find the inverse of the following :

(i) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (ii) $\begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}$

(iii) $\begin{bmatrix} 4 & -2 \\ 3 & 1 \end{bmatrix}$ (iv) $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$

(v) $\begin{bmatrix} 1 & 0 \\ 2 & -3 \end{bmatrix}$ (vi) $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

(vii) $\begin{bmatrix} i & -i \\ i & i \end{bmatrix}$ (viii) $\begin{bmatrix} x & -x \\ x & x^2 \end{bmatrix}$, where $x \neq 0, x \neq -1$

6. Find the adjoint of the following matrics.

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

(iii) $\begin{bmatrix} 2 & 1 & 2 \\ 2 & 2 & 1 \\ 1 & 2 & 2 \end{bmatrix}$ (iv) $\begin{bmatrix} 1 & 3 & 0 \\ 2 & -1 & 6 \\ 5 & 2 & 1 \end{bmatrix}$

7. Which of the following martices are invertible ?

(i) $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & -1 & 1 \end{bmatrix}$ (ii) $\begin{bmatrix} 2 & 1 & -2 \\ 1 & 2 & 1 \\ 3 & 6 & 4 \end{bmatrix}$

(iii) $\begin{bmatrix} -1 & -2 & 3 \\ 2 & 1 & -4 \\ -1 & 0 & 2 \end{bmatrix}$ (iv) $\begin{bmatrix} 1 & 0 & 1 \\ 2 & -2 & 1 \\ 3 & 2 & 4 \end{bmatrix}$

8. Examining consistency and solvability, solve the following equations by matrix method.

(i) $x - y + z = 4$
$2x + y - 3z = 0$
$x + y + z = 2$

(ii) $x + 2y - 3z = 4$
$2x + 4y - 5z = 12$
$3x - y + z = 3$

(iii) $2x - y + z = 4$
$x + 3y + 2z = 12$
$3x + 2y + 3z = 16$

(iv) $x + y + z = 4$
$2x + 5y - 2z = 3$
$x + 7y - 7z = 5$

(v) $x + y + z = 4$
$2x - y + 3z = 1$
$3x + 2y - z = 1$

(vi) $x + y - z = 6$
$2x - 3y + z = 1$
$2x - 4y + 2z = 1$

(vii) $x - 2y = 3$
$3x + 4y - z = -2$
$5x - 3z = -1$

(viii) $x + 2y + 3z = 14$
$2x - y + 5z = 15$
$2y + 4z - 3x = 13$

(ix) $2x + 3y + z = 11$
$x + y + z = 6$
$5x - y + 10z = 34$

9. Given the matrices

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & -2 & 1 \\ 4 & 2 & 1 \end{bmatrix}, X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ and } C = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

write down the linear equations given by AX = C and solve it for *x, y, z* by matrix method.

10. Find X if $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 2 & 1 & -1 \end{bmatrix} X = \begin{bmatrix} 6 \\ 0 \\ 1 \end{bmatrix}$, where $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

11. Answer the following :

(i) If every element of a third order matrix is multiplied by 5, then how many times its determinant value becomes ?

(ii) What is the value of x if $\begin{vmatrix} 4 & 1 \\ 2 & 1 \end{vmatrix}^2 = \begin{vmatrix} 3 & 2 \\ 1 & x \end{vmatrix} - \begin{vmatrix} x & 3 \\ -2 & 1 \end{vmatrix}$?

(iii) What are the values of x and y if $\begin{vmatrix} x & y \\ 1 & 1 \end{vmatrix} = 2, \begin{vmatrix} x & 3 \\ y & 2 \end{vmatrix} = 1$?

(iv) What is the value of x if $\begin{vmatrix} x+1 & 1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{vmatrix} = 4$?

(v) What is the value of $\begin{vmatrix} 0 & -h & -g \\ h & 0 & -f \\ g & f & 0 \end{vmatrix}$?

(vi) What is the value of $\begin{vmatrix} \frac{1}{a} & 1 & bc \\ \frac{1}{b} & 1 & ca \\ \frac{1}{c} & 1 & ab \end{vmatrix}$?

(vii) What is the cofactor of 4 in the determinant $\begin{vmatrix} 1 & 2 & -3 \\ 4 & 5 & 0 \\ 2 & 0 & 1 \end{vmatrix}$?

(viii) In which interval does the determinant $A = \begin{vmatrix} 1 & \sin\theta & 1 \\ -\sin\theta & 1 & \sin\theta \\ -1 & -\sin\theta & 1 \end{vmatrix}$ lie ?

(ix) If $x + y + z = \pi$, what is the value of $\Delta = \begin{vmatrix} \sin(x+y+z) & \sin B & \cos C \\ -\sin B & 0 & \tan A \\ \cos(A+B) & -\tan A & 0 \end{vmatrix}$

where A, B, C are the angles of a triangle.

12. Evaluate the following determinants ;

(i) $\begin{vmatrix} 14 & 3 & 28 \\ 17 & 9 & 34 \\ 25 & 9 & 50 \end{vmatrix}$

(ii) $\begin{vmatrix} 16 & 19 & 13 \\ 15 & 18 & 12 \\ 14 & 17 & 11 \end{vmatrix}$

(iii) $\begin{vmatrix} 224 & 777 & 32 \\ 735 & 888 & 105 \\ 812 & 999 & 116 \end{vmatrix}$

(iv) $\begin{vmatrix} 1 & 1 & 1 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{vmatrix}$

(v) $\begin{vmatrix} 1 & 2 & 3 \\ 3 & 5 & 7 \\ 8 & 14 & 20 \end{vmatrix}$

(vi) $\begin{vmatrix} 1^2 & 2^2 & 3^2 \\ 2^2 & 3^2 & 4^2 \\ 3^2 & 4^2 & 5^2 \end{vmatrix}$

(vii) $\begin{vmatrix} 1 & 0 & -5863 \\ -7361 & 2 & 7361 \\ 1 & 0 & 4137 \end{vmatrix}$

(viii) $\begin{vmatrix} 265 & 240 & 219 \\ 240 & 225 & 198 \\ 219 & 198 & 181 \end{vmatrix}$

(ix) $\begin{vmatrix} 0 & a^2 & b \\ b^2 & 0 & a^2 \\ a & b^2 & 0 \end{vmatrix}$

(x) $\begin{vmatrix} a-b & b-c & c-a \\ x-y & y-z & z-x \\ p-q & q-r & r-p \end{vmatrix}$

(xi) $\begin{vmatrix} a-b & b-c & c-a \\ b-c & c-a & a-b \\ c-a & a-b & b-c \end{vmatrix}$

(xii) $\begin{vmatrix} -\cos^2\theta & \sec^2\theta & -0.2 \\ \cot^2\theta & -\tan^2\theta & 1.2 \\ -1 & 1 & 1 \end{vmatrix}$

13. If $\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1+x & 1 \\ 1 & 1 & 1+y \end{vmatrix} = 0$, what are x and y ?

14. For what value of x $\begin{vmatrix} 2x & 0 & 0 \\ 0 & 1 & 2 \\ -1 & 2 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 2 & 3 & 4 \\ 0 & 3 & 5 \end{vmatrix}$?

15. Solve $\begin{vmatrix} x+a & 0 & 0 \\ a & x+b & 0 \\ a & 0 & x+c \end{vmatrix} = 0$

16. Solve $\begin{vmatrix} a+x & a-x & a-x \\ a-x & a+x & a-x \\ a-x & a-x & a+x \end{vmatrix} = 0$

17. Solve $\begin{vmatrix} x+a & b & c \\ a & x+b & c \\ a & b & x+c \end{vmatrix} = 0$

18. Show that $x = 2$ is a root of $\begin{vmatrix} x & -6 & -1 \\ 2 & -3x & x-3 \\ -3 & 2x & x+2 \end{vmatrix} = 0$. Solve this completely.

19. Evaluate $\begin{vmatrix} 1 & a & bc \\ 1 & b & ca \\ 1 & c & ab \end{vmatrix} - \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix}$

20. Evaluate $\begin{vmatrix} a & a^2-bc & 1 \\ b & b^2-ac & 1 \\ c & c^2-ab & 1 \end{vmatrix}$

21. For what value of λ the system of equations $x + y + z = 6$, $4x + \lambda y - \lambda z = 0$, $3x + 2y - 4z = -5$ does not possess a solution ?

22. If A is a 3 × 3 matrix and $|A| = 2$, then which matrix is represented by A × *adj* A ?

23. If $A = \begin{bmatrix} 0 & -\tan\frac{\alpha}{2} \\ \tan\frac{\alpha}{2} & 0 \end{bmatrix}$, show that $(I + A)(I - A)^{-1} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$ where $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

24. Prove the following :

(i) $\begin{vmatrix} a^2+1 & ab & ac \\ ab & b^2+1 & bc \\ ac & bc & c^2+1 \end{vmatrix} = 1 + a^2 + b^2 + c^2$

(ii) $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^3 & b^3 & c^3 \end{vmatrix} = (b-c)(c-a)(a-b)(a+b+c)$

(iii) $\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} = 3abc - a^3 - b^3 - c^3$ (iv) $\begin{vmatrix} b^2-ab & b-c & bc-ac \\ ab-a^2 & a-b & b^2-ab \\ bc-ac & c-a & ab-a^2 \end{vmatrix} = 0$

(v) $\begin{vmatrix} -a^2 & ab & ac \\ ab & -b^2 & bc \\ ac & bc & -c^2 \end{vmatrix} = 4a^2b^2c^2$ (vi) $\begin{vmatrix} (b+c)^2 & a^2 & bc \\ (c+a)^2 & b^2 & ca \\ (a+b)^2 & c^2 & ab \end{vmatrix} = (a^2+b^2+c^2)(a+b+c)(b-c)(c-a)(a-b)$

(vii) $\begin{vmatrix} b+c & a+b & a \\ c+a & b+c & b \\ a+b & c+a & c \end{vmatrix} = a^3 + b^3 + c^3 - 3abc$

(viii) $\begin{vmatrix} a+b+c & -c & -b \\ -c & a+b+c & -a \\ -b & -a & a+b+c \end{vmatrix} = 2(b+c)(c+a)(a+b)$

(ix) $\begin{vmatrix} ax-by+cz & ay+bx & az+cx \\ bx+ay & by-cz-ax & bz+cy \\ cx+az & ay+bz & cz-ax-by \end{vmatrix} = (a^2+b^2+c^2)(ax+by+cz)(x^2+y^2+z^2)$

25. If $2s = a+b+c$, show that $\begin{vmatrix} a^2 & (s-a)^2 & (s-a)^2 \\ (s-b)^2 & b^2 & (s-b)^2 \\ (s-c)^2 & (s-c)^2 & c^2 \end{vmatrix} = 2s^3(s-a)(s-b)(s-c)$

26. If $\begin{vmatrix} x & x^2 & x^3-1 \\ y & y^2 & y^3-1 \\ z & z^2 & z^3-1 \end{vmatrix} = 0$, then prove that $xyz = 1$ when x, y, z are non-zero and unequal.

27. Without expanding show that the following determinant is equal to $Ax + B$ where A and B are determinants of order 3 not involving x.

$$\begin{vmatrix} x^2+x & x+1 & x-2 \\ 2x^2+3x-1 & 3x & 3x-3 \\ x^2+2x+3 & 2x-1 & 2x-1 \end{vmatrix}$$

28. If x, y, z are positive and are the pth, qth and rth terms of a G.P.,

then prove that $\begin{vmatrix} \log x & p & 1 \\ \log y & q & 1 \\ \log z & r & 1 \end{vmatrix} = 0$

29. If $D_j = \begin{vmatrix} j & a & n(n+1)/2 \\ j^2 & b & n(n+1)(2n+1)/6 \\ j^3 & c & n^2(n+1)^2/4 \end{vmatrix}$, then prove that $\sum_{j=1}^{n} D_j = 0$

30. If $a_1, a_2, \ldots\ldots a_n$ are in G.P, and $a_i > 0$ for every i, then find the value of

$$\begin{vmatrix} \log a_n & \log a_{n+1} & \log a_{n+2} \\ \log a_{n+1} & \log a_{n+2} & \log a_{n+3} \\ \log a_{n+2} & \log a_{n+3} & \log a_{n+4} \end{vmatrix}$$

31. If $f(x) = \begin{vmatrix} 1+\sin^2 x & \cos^2 x & 4\sin^2 x \\ \sin^2 x & 1+\cos^2 x & 4\sin 2x \\ \sin^2 x & \cos^2 x & 1+4\sin^2 x \end{vmatrix}$, what is the least value of $f(x)$?

32. If $f_r(x)$, $g_r(x)$, $h_r(x)$, $r = 1, 2, 3$ are polynomials in x such that $f_r(a) = g_r(a) = h_r(a)$ and

$F(x) = \begin{vmatrix} f_1(x) & f_2(x) & f_3(x) \\ g_1(x) & g_2(x) & g_3(x) \\ h_1(x) & h_2(x) & h_3(x) \end{vmatrix}$, find $F'(x)$ at $x = a$.

33. If $f(x) = \begin{vmatrix} \cos x & \sin x & \cos x \\ \cos 2x & \sin 2x & 2\cos 2x \\ \cos 3x & \sin 3x & 3\cos 3x \end{vmatrix}$, find $f^1\left(\frac{\pi}{2}\right)$.

[Hint. In Q 32, $F^1(x) = \begin{vmatrix} f_1'(x) & f_2'(x) & f_3'(x) \\ g_1(x) & g_2(x) & g_3(x) \\ h_1(x) & h_2(x) & h_3(x) \end{vmatrix} + \begin{vmatrix} f_1(x) & f_2(x) & f_3(x) \\ g_1'(x) & g_2'(x) & g_3'(x) \\ h_1(x) & h_2(x) & h_3(x) \end{vmatrix} +$

$$\begin{vmatrix} f_1(x) & f_2(x) & f_3(x) \\ g_1(x) & g_2(x) & g_3(x) \\ h_1'(x) & h_2'(x) & h_3'(x) \end{vmatrix}.$$

CHAPTER 6

Probability

If friendship is your weakest point, then you are the strongest person in the world.

- Abraham Lincoln

6.0 Introduction

In the first volume we have dealt with probability of random or nondeterministic events.

We have introduced some fndamental concepts like statistical experiment, sample space of an experiment, elementary events, event as a subset of a sample space, mutually exclusive events, equiprobable or uniform sample spaces and above all, the definition of probability in a uniform sample space and have proved some propositions related to the properties of probability.

A recapitulation of the above discussions and the techniques involved will be helpful in following the contents that come in the sequel.

6.1 Conditional Probability

Consider the experiment of choosing a two digit number from N. The sample space of the outcomes of this experiment is given as

$S = \{10,11,...,19;20,21,...,29;...;90,91,...,99\}$

Clearly $|S| = 90$

Under this experiment let events E_2 and E_3 be as follows :

E_2 = event of choosing an even number

$= \{10,12,..., 18;20,22,...,28;30,32,...,38;...;90,92,...98\}$

E_3 = event of choosing a number divisible by 3

$= \{12,15,18, 21,24,27, 30,33,36,39,...90,93,96,99\}$

Clearly $|E_2| = 45$ and $|E_3| = 30$.

Out of E_2 and E_3 we now contemplate two events $E_2|E_3$ = the event that the choice is an even number, given that a number divisible by 3 has already been choosen

$= \{12,18,24,30,36,42,48,54,60,66,72,78,84,90,96\}$

Clearly $|E_2 | E_3| = 15$

In writing $E_2|E_3$ we precisely choose those numbers of E_3 which are members of E_2, which is nothing but $E_2 \cap E_3$. Thus the choice of $E_2|E_3$ is actually 15 $(=|E_2 | E_3|)$ out of 30 $(=|E_3|)$.

Such an event as $E_2|E_3$ is called a **conditional event** and its probability is obviously 15 out of 30, i.e. $\frac{15}{30}=\frac{|E_2 | E_3|}{|E_3|}=\frac{1}{2}$ or, for that matter, $P(E_2|E_3)=\frac{|E_2 \cap E_3|}{|E_3|}$, which is known as a conditional probability - the probabilityof E_2 subject to E_3, i.e. the probability of the choice being an even number subject to the condition that a number divisible by 3 has already been chosen.

In the same vein we can now consider $E_3|E_2$, the event of the choice being a number divisible by 3, given that an even number has already been chosen.

As before it can very well be seen that

$$P(E_3|E_2) = \frac{|E_3 | E_2|}{|E_2|} = \frac{|E_3 \cap E_2|}{|E_2|} = \frac{15}{45} = \frac{1}{3}$$

In the context of the sample space S we make the modification

$$P(E_2|E_3) = \frac{|E_2 | E_3|}{|E_3|} = \frac{|E_3 \cap E_3|}{|E_3|} = \frac{\frac{|E_2 \cap E_3|}{|S|}}{\frac{|E_3|}{|S|}} = \frac{P(E_2 \cap E_3)}{P(E_3)}$$

Similarly we can get

$$P(E_3|E_2) = \frac{P(E_3 \cap E_2)}{P(E_2)}$$

In the above example it is easy to observe that

$P(E_2|E_3) = \frac{15}{30} = \frac{1}{2}$ and

$P(E_3|E_2) = \frac{15}{45} = \frac{1}{3}$.

The above discussions prompt the following definition.

Definition (Conditional Probability)

When A and B are events in a sample space the conditional probability of B subject to A is defined as

$$\mathbf{P(B|A) = \frac{P(B \cap A)}{P(A)}}$$

Which is the probability of B subject to he condition that A has already occurred.

Similarly

$$\mathbf{P(A|B) = \frac{P(A \cap B)}{P(B)}}.$$

Note : The definition P(B|A) and P(A|B) respectively presuppose that $P(A) \neq 0$ and $P(B) \neq 0$; P(B|A) and P(A|B) are meaningful only when $A \neq \phi$ and $B \neq \phi$ respectively.

Corollary :

The following properties follow from the definition of conditional probability :

(i) **For any $E \subset S$ (sample space), with $F \neq \phi$**

$P(S|F) = 1$ and also $P(F|F) = 1$.

Proof :

$$P(S|F) = \frac{|S \cap F|}{|F|} = \frac{|F|}{|F|} = 1$$

Proof of $P(F|F) = 1$ is left to the reader.

(ii) **For any three events E, F and G in sample space S with $G \neq \phi$,**

$$((E \cup F)|G) = P(E|G) + P(F|G) - P((E \cap F)|G)$$

Proof follows easily from $|E \cup F| = |E| + |F| - |E \cap F|$.

(iii) **For events E and F with $F \neq \phi$**

$$P(E^c|F) = 1 - P(E|F)$$

Proof : From Property (i)

$P(S|F) = 1$

$$\Rightarrow P((E \cup E^c)|F) = P(E|F) + P(E^c|F) = 1$$

($\because$ E and E^c are disjoint and $S = E \cup E^c$)

$$\Rightarrow P(E^c|F) = 1 - P(E|F).$$

(iv) **For events E, F and G in sample space S, with $G \neq \phi$,**

$$E \subseteq F \Rightarrow P(E|G) \leq P(F|G)$$

Proof:

$$E \subseteq F \Rightarrow F = E \cup (F - E)$$

$$\Rightarrow |F| = |E| + |F - E|$$

($\because$ E and F-E are disjoint sets)

Now $F \cap G = ((E \cup (F - E)) \cap G = (E \cap G) \cup ((F - E) \cap G)$

$$\Rightarrow |F \cap G| = |E \cap G| + |(F - E) \cap G|$$

($\because E \cap (F - E) = \phi \Rightarrow (E \cap G) \cap ((F - E) \cap G) = \phi$)

$$\Rightarrow |F \cap G| = |E \cap G| + |(F - E) \cap G|$$

$$\geq |E \cap G|$$

($\because |(F - E) \cap G| \geq 0$)

$$\because \frac{|F \cap G|}{|G|} \geq \frac{|E \cap G|}{|G|}$$

So $P(E|G) \leq P(F|G)$

Multiplication theorem for conditional probability

(Multiplication Rule)

For events A and B with $A \neq \phi$,

$$P(B \cap A) = P(A).P(B \mid A)$$

Proof:

We have, by definition of conditional probability

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)}$$

Multiplying P(A) on both sides

$$P(B \cap A) = P(A).P(B \mid A) \qquad \textbf{(A)}$$

Interpreted verbally, formula (A) says that probability of joint occurrence of A and B (or B and A) (Some times also written as P(BA) is given by probability of A multiplied by probability of B subject to the occurrence of A.

Formula (A) is usually called the multiplication theorem for conditional probability (or multiplication theorem for probability). It has great advantages as the following applications will show :

Example 1

Two cards are drawn at random from a pack of 52 cards. Find the probability that both are kings.

Let A be the event that the first card is a king and let B be the event that the second card is a king. We are thus interested in finding the value of $P(A \cap B) = P(B \cap A)$

$= P(A) . P(B \mid A)$.

Now $P(A) = \frac{4}{52}$ since there are 52 ways of drawing a card and there are just 4 kings. Let us now compute $P(B \mid A)$, that is, the probability that the second card is a king subject to the condition that the first card is already a king. Since there are 51 cards left of which three are kings, $P(B \mid A) = \frac{3}{51}$.

Thus $P(A \cap B) = \frac{4}{52} \cdot \frac{3}{51} = \frac{1}{221}$.

Note :

As said earlier, the multiplication theorem has great advantages. The problem solved above could have been attemped directly : The sample space in this case has ${}^{52}C_2$ points whereas we can draw 2 kings in ${}^{4}C_2$ ways; so the required probability is

$$\frac{^4C_2}{^{52}C_2} = \frac{4}{52} \cdot \frac{3}{51} = \frac{1}{221}.$$

Remark

The multiplication theorem for probability can be further generalized as follows :

If A, B, C are events then $P(C \cap B \cap A) = P(A). P(B \mid A) . P(C \mid A \cap B)$, **(B)**

(Provided $B \cap A \neq \phi$, i.e. $P(B \cap A) > 0$)

Proof : If we let $D = B \cap A$, then

$P(C \cap D) = P(D) . P(C \mid D) = P(A \cap B). P(C \mid A \cap B) = P(A) . P(B \mid A) . P(C \mid A \cap B)$,

Which is also written as $P(ABC)$ or $P(CBA) = P(A) P(B|A) P(C|AB)$;

where AB means $A \cap B$.

Example 2

Three cards are drawn at random from a pack of 52 cards. Find the probability that all three are queens.

Let A, B, C be the events that the first, second and the third cards are queens (respectively). Then we want the value of $P(A \cap B \cap C)$. Employing the formula (B), we have

$$P(A \cap B \cap C) = P(A). P(B \mid A). P(C \mid A \cap B) = \frac{4}{52} \cdot \frac{3}{51} \cdot \frac{2}{50} = \frac{1}{5525}$$

Example 3

A person draws five cards at random from a pack of 52 cards. Find the probability that all the five cards are spades.

The required probability, using the generalized multiplication rule, is

$$\frac{13}{52} \cdot \frac{12}{51} \cdot \frac{11}{50} \cdot \frac{10}{49} \cdot \frac{9}{48}.$$

Note that $\frac{13}{52}$ is the probability that the first card is a spade; $\frac{12}{51}$ is the probability that the second card is a spade under the condition that first card is already a spade etc.

Example 4

A person has four pairs of socks of four different colours. If four socks are picked at random what is the probability that there is at least one pair among them ?

Let A be the event that there is at least one pair among the four socks. Let us now compute the probability of the event A^c (= the event that all four socks are of different colours). Now, applying the multiplication rule, we have $P(A^c) = \frac{8}{8} \cdot \frac{6}{7} \cdot \frac{4}{6} \cdot \frac{2}{5} = \frac{8}{35}$, hence $P(A) = 1 - P(A^c) = \frac{27}{35}$.

Note : You are advised to compute the probability of A directly and see for yourself how lengthy and difficult it can be !

Generalised multiplication theorem for probability

(Generalised multiplication rule)

For events $A_1, A_2, \ldots, A_n$ we have

$$P(A_1A_2\ldots A_n) = P(A_1).P(A_2|A_1).\,P(A_3|A_1A_2)\ldots P(A_n|\,A_1A_2\ldots A_{n-1}), \qquad \text{(C)}$$

provided $P(A_1A_2\ldots A_{n-1}) > 0$.

($A_1A_2\ldots A_n$ means $A_1 \cap A_2 \cap \ldots \cap A_n$)

Proof:

$A_1 \supseteq A_1A_2 \supseteq A_1A_2A_3 \supseteq \ldots \supseteq A_1A_2\ldots A_{n-1}$

$\Rightarrow P(A_1) \geq P(A_1A_2) \geq P(A_1A_2A_3) \geq \ldots \geq P(A_1A_2\ldots A_{n-1})$

$P(A_1A_2\ldots A_{n-1}) > 0 \Rightarrow P(A_1A_2\ldots A_{n-1}) \neq 0$

and hence $A_1A_2\ldots A_{n-1} \neq \phi$

$\Rightarrow A_1, A_1A_2, A_1A_2A_3\ldots$ are all nonempty sets with positive probability.

Therefore all the conditional probabilities on RHS of (C) are well-defined.

$\therefore P(A_1).P(A_2 \mid A_1).P(A_3 \mid A_1A_2)\ldots P(A_n \mid A_1A_2\ldots A_{n-1})$

$$= P(A_1).\frac{P(A_1A_2)}{P(A_1)}.\frac{P(A_1A_2A_3)}{P(A_1A_2)}\ldots\frac{P(A_1A_2\ldots A_n)}{P(A_1A_2\ldots A_{n-1})}$$

$= P(A_1A_2\ldots A_n)$ (after cancellation)

This completes the proof.

Example 5

Two dice are thrown simultaneously and the sum of the numbers appearing is observed to be 7. What is the probability that the number 3 has appeared atleast once.

Solution:

The sample space S has $6^2 = 36$ sample points.

The event E : the sum of the numbers appearing is 7

$\therefore E = \{(1,6), (6,1), (2,5), (5,2), (3,4), (4,3)\}$ and

F : the number 3 has appeared at least once ;

$\therefore F = \{(3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (1,3), \ldots(6,3)\}$

Now $|E| = 6$, $|F| = 11$, also $|E \cap F| = |\{(3,4), (4,3)\}| = 2$

$\therefore$ The required probability $= P(F|E)$

$$= \frac{P(F \cap E)}{P(E)} = \frac{2/36}{6/36} = \frac{1}{3}$$

Example 6

A family has two children. What is the probability that both the children are girls, given that at least one of them is a girl.

Solution:

Let 'b' and 'g' stand for boy and girl respectively.

Then the sample space S = {gg, gb, bg, bb}

Let E and F stand for the events

E : both the children are girls,

F : at least one of them is a girl.

Then E= {gg}, F={gg,gb,bg}, $E\cap F$={gg}

So that $|E|=1$, $|F|=3$, $|E\cap F|=1$ and

$P(E\cap F)=\frac{|E\cap F|}{|S|}=\frac{1}{4}$, $P(F)=\frac{|F|}{|S|}=\frac{3}{4}$

So the required probability $P(E|F)=\frac{P(E\cap F)}{P(F)}=\frac{1}{3}$.

N.B. We can also conclude

(i) P(at least one boy)

$=1-\frac{1}{3}=\frac{2}{3}$

(ii) By similer reasoning as in the example, P (both are boys) = $\frac{1}{3}$.

6.2 Independent Events

Definition : Two events A and B are said to be **independent** if $P(A\cap B) = P(A).P(B)$.

Note that if A, B are independent, then $P(A\mid B) = \frac{P(A\cap B)}{P(B)} = \frac{P(B).P(A)}{P(B)} = P(A)$.

The first one shows that $P(A) = P(A \mid B)$; that is, the probability of A is same as the conditional probability of A subject to B. In other words, A is not influenced by B. Similarly $P(B \mid A) = P(B)$ shows that B is not influenced by A.

Example 7

Let a coin be tossed three times. Let A be the event that the first toss is head and let B be the event that the second toss is head. Are A and B independent ?

As shown earlier, the sample space has 8 points. we also have

A = {*hhh*, *hht*, *hth*, *htt*}, B = {*hhh*, *hht*, *thh*, *tht*}, $A\cap B$ = {*hhh*, *hht*}.

$P(A)=\frac{4}{8}=\frac{1}{2}$, $P(B)=\frac{4}{8}=\frac{1}{2}$, $P(A\cap B)=\frac{2}{8}=\frac{1}{4}$.

Thus $P(A\cap B) = P(A) . P(B)$; hence A and B are independent.

Example 8

There are 100 tickets in a bag numbered 1 through 100 and a ticket is picked at random. Let A be the event that the number on the ticket is divisible by 2 and let B be the event that number on the ticket is divisible by 5. Show that A and B are independent.

Clearly $P(A) = \frac{50}{100} = \frac{1}{2}$ and $P(B) = \frac{20}{100} = \frac{1}{5}$. $A \cap B$ is the event that the number on the ticket is divisible by both 2 and 5, hence divisible by 10. Therefore

$$P(A \cap B) = \frac{10}{100} = \frac{1}{10} = P(A) \cdot P(B).$$

Example 9

Let A and B be defined as in example 23 above where a ticket is picked at random from a bag containing 65 tickets numbered 1 through 65. In this case

$$P(A) = \frac{32}{65}, P(B) = \frac{13}{65} = \frac{1}{5}.$$

On the other hand, $P(A \cap B) = \frac{6}{65} \neq P(A). P(B)$.

Thus A and B are not independent.

Example 10

When two fair coins are tossed all the four out comes, HH, HT, TH, TT are equally likely. If E is the event that the first coin shows head and F, the event that the second shows tail, then

E and F are independent, since $P(EF) = P(\{H,T\}) = \frac{1}{4}$;

$$P(E) = P(\{H,H\},\{H,T\}) = \frac{1}{2}, P(F) = P(\{H,T\},\{T,T\}) = \frac{1}{2}.$$

Example 11

When tossing two fair dice, let E denote the event that the sum is 6 and F denote the event that the first die shows 4.

Here $|S| = 36$ where S is sample space.

$E = \{(1,5), (5,1), (2,4), (4,2), (3,3)\}$

$F = \{(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)\}$

Now, $P(E) = \frac{5}{36}$

$$P(F) = \frac{6}{36} = \frac{1}{6}$$

$$P(E).P(F) = \frac{5}{216}$$

Where as $P(EF) = P(\{4,2\}) = \frac{1}{36} \neq P(E).P(F)$

So E and F are not independent.

Example 12

A speaks the truth in 80% of cases and B in 70% of the cases. In what percentage of cases they are likely to contradict each other in reporting the same fact.

Solution :

Let A, B, $\overline{A}$, $\overline{B}$ respectively stand for

A : A speaks the truth

B : B speaks the truth

$\overline{A}$: A tells a lie

$\overline{B}$: B tells a lie

Let E be the event that A and B contradict each other.

Obviously $E = (A \cap \overline{B}) \cup (\overline{A} \cap B)$

So $P(E) = P(A \cap \overline{B}) + P(\overline{A} \cap B)$ ($\because A \cap \overline{B}$ and $\overline{A} \cap B$ are disjoint)

Obviously A, $\overline{B}$ and $\overline{A}$, B are independent events

$\therefore P(E) = P(A)P(\overline{B}) + P(\overline{A})P(B)$

$$= \frac{8}{10} \times \frac{3}{10} + \frac{2}{10} \times \frac{7}{10}$$

$$= \frac{38}{100}$$

or $P(E) = 38\%$

Extended Definition of Independent Events

Definition

Three events are said to be **independent** if

$P(EF) = P(E).P(F)$

$P(FG) = P(F).P(G)$

$P(GE) = P(G).P(E)$

$P(EFG) = P(E).P(F).P(G)$

The definition can be further extended to n number of events $E_1, E_2, ..., E_n$ under the condition :

$P(E_1 E_2 ... E_k) = P(E_1) P(E_2) ... P(E_k)$ for $2 \le k \le n$.

N.B. There can be extension up to infinite number of events, but that will take us outside the scope of the book.

EXERCISES 6 (a)

1. Two balls are drawn from a bag containing 5 white and 7 black balls. Find the probability of selecting 2 white balls if

   (i) the first ball is replaced before drawing the second.

   (ii) the first ball is not replaced before drawing the second.

2. Two cards are drawn from a pack of 52 cards; find the probability that

   (i) they are of different suits.

(ii) they are of different denominations.

3. Do both parts of problem 2 if 3 cards are drawn at random.

4. Do both parts of problem 2 if 4 cards are drawn at random.

5. A lot contains 15 items of which 5 are defective. If three items are drawn at random, find the probability that (i) all three are defective (ii) none of the three is defective. Do this problem directly.

6. A pair of dice is thrown. Find the probability of getting a sum of at least 9 if 5 appears on at least one of the dice.

Hints : Let A be the event of getting at least 9 points and B, the event that 5 appears on at least one of the dice. Clearly,

$B = \{(1, 5), (2, 5), (3, 5), (4, 5), (5, 5), (6, 5), (5, 1), (5, 2), (5, 3), (5, 4), (5, 6)\}$

whereas, $A \cap B = \{(4, 5)\ (5, 5), (6, 5), (5, 4), (5, 6)\}$.

Therefore $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{\frac{5}{36}}{\frac{11}{36}} = \dfrac{5}{11}$.

7. A pair of dice is thrown. If the two numbers appearing are different, find the probability that

   (i) the sum of points is 8. (ii) the sum of points exceeds 8.

   (iii) 6 appears on one die.

8. In a class 30% of the students fail in Mathematics, 20% of the students fail in English and 10% fail in both. A student is selected at random.

   (i) If he has failed in English, what is the probability that he has failed in Mathemetics ?

   (ii) If he has failed in Mathematics, what is the probability that he has failed in English ?

   (iii) What is the probability that he has failed in both ?

9. If A, B are two events such that $P(A) = 0.3$, $P(B) = 0.4$, $P(A \cup B) = 0.6$. Find

   (i) $P(A \mid B)$ (ii) $P(B \mid A)$

   (iii) $P(A \mid B^c)$ (iv) $P(B \mid A^c)$.

10. If A, B are events such that $P(A) = 0.6$ $P(B) = 0.4$ and $P(A \cap B) = 0.2$, then find

    (i) $P(A \mid B)$ (ii) $P(B \mid A)$

    (iii) $P(A \mid B^c)$ (iv) $P(B \mid A^c)$.

11. If A and B are independent events, show that (i) A^c and B^c are independent. (ii) A and B^c are independent (iii) A^c and B are independent.

12. Two different digits are selected at random from the digits 1 through 9.

    (i) If the sum is even, what is the probability that 3 is one of the digits selected ?

    (ii) 'If the sum is odd, what is the probability that 3 is one of the digits selected ?

(iii) If 3 is one of the digits selected, what is the probability that the sum is odd ?

(iv) If 3 is one of the digits selected, what is the probability that the sum is even ?

13. If $P(A) = 0.4$ $P(B \mid A) = 0.3$ and $P(B^c \mid A^c) = 0.2$. Find

(i) $P(A \mid B)$ (ii) $P(B \mid A^C)$ (iii) $P(B)$

(iv) $P(A^c)$ (v) $P(A \cup B)$.

14. If $P(A) = 0.6$, $P(B \mid A) = 0.5$; find $P(A \cup B)$ when A, B are independent.

15. Two cards are drawn in succession from a deck of 52 cards. What is the probability that both cards are of denomination greater than 2 and less than 9 ?

*16. From a bag containing 5 black and 7 white balls, 3 balls are drawn in succession. Find the probability that

(i) all three are of the same colour.

(ii) each colour is represented.

17. A die is rolled until a 6 is obtained. What is the probability that

(i) you end up in the second roll.

(ii) you end up in the third roll.

18. A person takes 3 tests in succession. The probability of his (her) passing the first test is 0.8. The probability of passing each successive test is 0.8 or 0.5 according as he passes or fails the preceding test. Find the probability of his (her) passing at least 2 tests.

Hints : We have the following mutually exclusive cases where S = success in a test and F = failure in a test

event	Probability
S S S	$.8 \times .8 \times .8$
S S F	$.8 \times .8 \times .2$
S F S	$.8 \times .2 \times .5$
F S S	$.2 \times .5 \times .8$.

*19. A person takes 4 tests in succession. The probability of his passing the first test is p, that of his passing each succeeding test is p or $\frac{p}{2}$ depending on his passing or failing the preceding test. Find the probability of his passing (i) at least three tests (ii) just three tests.

20. Given that all three faces are different in a throw of three dice, find the probability that

(i) at least one is a six, (ii) the sum is 9.

Hints for part (i) : Let A be the event that at least one (of the three results) is a six. Let B be the event that all three faces are different. The size of the sample space S is 216; the size of B is $^6C_3.3!$. Let us compute $P(A^c \mid B) = \dfrac{P(A^c \cap B)}{P(B)}$.

$A^c \cap B$ is the event that all three faces are different and 6 does not occur. Thus the size of

$A^c \cap B$ is ${}^5C_3.3!$. Thus $P(A^c \mid B) = \dfrac{{}^5C_3 \,.\, 3!/216}{{}^6C_3 .\, 3!/216} = \dfrac{1}{2}$.

21. From the set of all families having three children, a family is picked at random.
    (i) If the eldest child happens to be a girl, find the probability that she has two brothers.
    (ii) If one child of the family is a son, find the probability that he has two sisters.

22. Three persons hit a target with probabilities $\frac{1}{2}$, $\frac{1}{3}$ and $\frac{1}{4}$ respectively. If each one shoots at the target once, (i) find the probability that exactly one of them hits the target (ii) if only one of them hits the target what is the probability that it was the first person ?

Special Application of Multiplication Rule

(Compound events)

The multiplication rule comes in handy while dealing with the so called compound events.

Suppose that a bag has 4 white and 6 black balls and that a second bag contains 5 white and 7 black balls. Suppose that a ball is randomly picked from the first bag and put in the second bag (without knowing its colour). If you now pick a ball from the second bag what is the probability that it is white ?

To solve this problem, let

W_1 = the event that the first draw is white (which is transferred to the second bag)

B_1 = the event that the first draw is black

W_2 = the event that the second draw is white

B_2 = the event that the second draw is black.

We thus have the following four situations.

(i) $W_1 \cap W_2$ (ii) $W_1 \cap B_2$ (iii) $B_1 \cap W_2$ and (iv) $B_1 \cap B_2$.

The probability of each of these events can be calculated by using the multiplication rule. For example,

$$P(W_1 \cap W_2) = P(W_1) \,.\, P(W_2 \mid W_1) = \frac{4}{10} \cdot \frac{6}{13} = \frac{12}{65}.$$

$$P(W_1 \cap B_2) = P(W_1) \,.\, P(B_2 \mid W_1) = \frac{4}{10} \cdot \frac{7}{13} = \frac{14}{65}.$$

$$P(B_1 \cap W_2) = P(B_1) \,.\, P(W_2 \mid B_1) = \frac{6}{10} \cdot \frac{5}{13} = \frac{15}{65}.$$

$$P(B_1 \cap B_2) = P(B_1) \,.\, P(B_2 \mid B_1) = \frac{6}{10} \cdot \frac{8}{13} = \frac{24}{65}.$$

Thus the probability of drawing a white ball from the second bag is given by

$$P(W_1 \cap W_2) + P(B_1 \cap W_2) = \frac{12}{65} + \frac{15}{65} = \frac{27}{65}.$$

Example 13

From a bag containing 6 black and 4 white balls, two balls are drawn at random. Find the probability that both are of the same colour.

We have two mutually exclusive cases; either both are white or both are black.

P (both are white) = $\frac{{}^{4}C_2}{{}^{10}C_2} = \frac{2}{15}$.

P (both are black) = $\frac{{}^{6}C_2}{{}^{10}C_2} = \frac{1}{3}$.

P (both are of the same colour)

= P (both are white or both are black)

= P (both are white) + P (both are black)

$= \frac{2}{15} + \frac{1}{3} = \frac{7}{15}$.

Example 14

From a bag containing 3 white, 4 black and 5 red balls, three balls are drawn at random. Find the probability that all three are of the same colour.

We have three mutually exclusive cases whose probabilities we have to find :

P (all three are white) = $\frac{{}^{3}C_3}{{}^{12}C_3} = \frac{1}{220}$.

P (all three are black) = $\frac{{}^{4}C_3}{{}^{12}C_3} = \frac{4}{220}$.

P (all three are red) = $\frac{{}^{5}C_3}{{}^{12}C_3} = \frac{10}{220}$.

Therefore, P (all three are of the same colour) = $\frac{1+4+10}{220} = \frac{3}{44}$.

Example 15

A bag contains 4 white and 5 black balls; a second bag has 3 white and 4 black balls; a third bag has 5 white and 4 black balls. A ball is randomly picked from the first bag and put in the second; then another ball is picked at random from the second bag and put in the third. If now ball is randomly picked from the third bag, what is the probability that it is white ?

Let B = black and W = white. We are thus interested in the following cases :

	First transfer	Second transfer	Third pick	Probability
(i)	W	B	W	$\frac{4}{9} \cdot \frac{4}{8} \cdot \frac{5}{10}$
(ii)	B	W	W	$\frac{5}{9} \cdot \frac{3}{8} \cdot \frac{6}{10}$
(iii)	B	B	W	$\frac{5}{9} \cdot \frac{5}{8} \cdot \frac{5}{10}$
(iv)	W	W	W	$\frac{4}{9} \cdot \frac{4}{8} \cdot \frac{6}{10}$

Thus the total probability = $\frac{391}{720}$.

Example 16

A purse contains ten one-rupee coins and one two-rupee coin; a second purse contains eleven one-rupee coins. Ten coins are transferred from the first purse to the second and then ten coins are transfered from the second purse to the first at random. Find the probability that the two rupee coin is still in the first purse.

We have two mutually exclusive cases : (i) the two - rupee coin has not been transferred at all or (ii) the two- rupee coin has been transferred from purse 1 to purse 2 and then again from purse 2 to purse 1. For case (i), the probability p_1 is given by

$$p_1 = 1 - \frac{1 \times {}^{10}C_9}{{}^{11}C_{10}} = 1 - \frac{10}{11} = \frac{1}{11}.$$

For case (ii), the probability p_2 is given by $p_2 = \frac{1 \times {}^{10}C_9}{{}^{11}C_{10}} \times \frac{1 \times {}^{20}C_9}{{}^{21}C_{10}} = \frac{10}{11} \times \frac{10}{21}$

$$p_1 + p_2 = \frac{11}{21}.$$

EXERCISES 6 (b)

1. A bag contains 5 white and 3 black marbles and a second bag contains 3 white and 4 black marbles. A bag is selected at random and a marble is drawn from it. Find the probability that it is white. Assume that either bag can be chosen with the same probability.

Hint : There are two mutually exclusive cases : (i) you are selecting bag 1 and then drawing a white; call this event 1 W (ii) You are selecting bag 2 and then drawing a white, call this event 2 W.

Now P(1W) = P(1).P(W/1) = $\frac{1}{2} \cdot \frac{5}{8}$ and P(2W) = P(2) . P(W/2) = $\frac{1}{2} \cdot \frac{3}{7}$; hence the required probability is $\frac{5}{16} + \frac{3}{14} = \frac{59}{112}$.

2. A bag contains 5 white and 3 black balls; a second bag contains 4 white and 5 black balls; a third bag contains 3 white and 6 black balls. A bag is selected at random and a ball is drawn.

Find the probability that the ball is black.

(i) Do the problem assuming that the probability of choosing each bag is same

(ii) Do the problem assuming that the probability of choosing the first bag is twice as much as choosing the second bag, which is twice as much as choosing the third bag.

3. A and B play a game by alternately throwing a pair of dice. One who throws 8 wins the game. If A starts the game, find their chances of winning.

Hint : 8 can be obtained with a pair of dice in the following five ways :

(6, 2), (5, 3), (4, 4), (3, 5) and (2, 6).

Thus $P(8) = \frac{5}{36}$. (Remember the sample spece, when a pair of dice are rolled !);

$P(\text{not } 8) = \frac{31}{36}$. Since A starts the game, A can win in the following situations :

(i) A throws 8.

(ii) A does not throw 8, B does not throw 8, A throws 8.

(iii) A does not throw 8, B does not throw 8, A does not throw 8, B does not throw 8, A throws 8 etc.

In case (i) probability $= \frac{5}{36}$.

In case (ii) probability $= \frac{31}{36} \cdot \frac{31}{36} \cdot \frac{5}{36} = \frac{5}{36} \cdot \left(\frac{31}{36}\right)^2$.

In case (iii) probability $= \frac{31}{36} \cdot \frac{31}{36} \cdot \frac{31}{36} \cdot \frac{31}{36} \cdot \frac{5}{36} = \frac{5}{36}\left(\frac{31}{36}\right)^4$ etc.

Thus $P(\text{A wins}) = \frac{5}{36}\left\{1+\left(\frac{31}{36}\right)^2+\left(\frac{31}{36}\right)^4+\ldots\right\}$

$$= \frac{\frac{5}{36}}{\left[1-\left(\frac{31}{36}\right)^2\right]} = \frac{36}{67}.$$

Hence, $P(B) = 1 - \frac{36}{67} = \frac{31}{67}$.

4. A, B,C play a game by throwing a pair of dice in that order. One who gets 8 wins the game. If A starts the game, find their chances of winning.

5. There are 6 white and 4 black balls in a bag. If four are drawn successively (and not replaced), find the probability that they are alternately of different colour.

Hint : There are two mutually exclusive cases, WBWB and BWBW.

6. Five boys and four girls randomly stand in a line. Find the probability that no two girls come together.

7. If you throw a pair of dice n time, find the probability of getting at least one doublet.

[When you get indentical members you call it a doublet. You can get a double in six ways : (1, 1), (2, 2), (3, 3), (4, 4), (5, 5) and (6, 6); thus the probability of getting a doublet is $\frac{6}{36} = \frac{1}{6}$, so that the probability of not getting a doublet in one throw is $\frac{5}{6}$]

8. Suppose that the probability that your alarm goes off in the morning is 0.9. If the alarm goes off, the probability is 0.8 that you attend your 8 a.m class. If the alarm does not go off, the probability that you make your 8 a.m. class is 0.5. Find the probability that you make your 8 a.m. class.

9. If a fair coin is tossed 6 times, find the probability that you get just one head.

Hint : There are 6 mutually exclusive cases :

HTTTTT, THTTTT, TTHTTT, TTTHTT, TTTTHT, TTTTTH.

10. Can you generalize this situation ? If a fair coin is tossed six times, find the probability of getting exactly 2 heads.

Hint : First find out the number of ways you can write a sequence of six heads and tails with 2 heads and 4 tails. Argue this way : Fill any two places with H and the remaining four places by T. Since the two places can be chosen arbitrarily we can do so in ${}^6C_2$ ways. For example HHTTTT is one such. Again each of these results has a probability equal to $\frac{1}{2^6}$.

A remark

We now know that there are sample spaces which are equiprobable (or uniform) and also sample spaces which are not equiprobable. If a fair coin is tossed twice (or two fair coins are tossed once), we know that we get an equiprobable sample space

$$S = \{hh, ht, th, tt\}$$

so that each elementary event has probability equal to $\frac{1}{4}$. In particular, the probability of getting one head and one tail (not necessarily in that order), is equal to $\frac{1}{2}$. Even such a distinguished mathematician as D'Alembert (1717 -83) argued that there are three possible outcomes in all, namely,

(i) 2 heads and no tail

(ii) 2 tails and no head

(iii) 1 head and 1 tail.

He therefore argued that each elementary event should have a probability equal to $\frac{1}{3}$. His surmise, of course, didn't match with experimental results. It was observed that if two coins were tossed a large number of times then one head and one tail appeared approximately half of the times and not one-third of the times as D'Alembert claimed. So the conclusion was that even if we take the sample space S to be

S = {2 heads, one head and one tail, 2 tails}, this is not an equiprobable space.

In describing the rules of probability for a finite sample space S, we had this rule:

If A_1 and A_2 are mutually exclusive then $P(A_1 \cup A_2) = P(A) + P(A_2)$

From this rule, one can easily derive the following : If $A_1, A_2, \ldots, A_n$ are mutually exclusive events in a sample space S, that is $A_i \cap A_j = \phi$ whenever $i \neq j$ then

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_2).$$

However, if the sample space is not finite, for example if S′ is countably infinite, we will need a stronger axiom : If $\{A_n\}$, $n = 1, 2 \ldots, \infty$, is a sequence of mutually exclusive events then

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

When S is uncountable (for example, the set of reals and the set of irrationals), every subset of S need not be an **event**; only a special collection of subsets of S (called *a* σ algebra) will be events. These facts are beyond the scope of this book.

In the beginning of this chapter, we defined the probability of an event A (in a finite sample S) as

$$P(A) = \frac{\text{size of A}}{\text{size of S}}.$$

In all the examples and problems given until now, size simply meant the number of elements of the concerned set. We used the word size simply because there are situations in which **size** may mean other things. To justify our claim you may consider the following experiment. Imagine that you have been invited to open a science exhibition in your school where you have to cut a ribbon of length 1 meter tied across the door. If you do it at random, then every point of the entire length (of one metre) is a point of the sample space. Thus, the sample space is S = [0, 1].

Suppose now that the middle one-third of the ribbon is of red colour and the two sides are of green colour. Let A be the event that you cut the ribbon at a point in the middle one third. What is then the probability of A assuming that all points of S are equally likely to get you scissor ?

$$A = \left[\frac{1}{3}, \frac{2}{3}\right]$$

is of length $\frac{1}{3}$ and S is of length 1. Our intuition tells us that $P(A) = \frac{1/3}{1} = \frac{1}{3}$. If you look at your definition $P(A) = \frac{\text{size of A}}{\text{size of S}}$,

then it would be obvious that size in this case simply means **length.**

Take yet another example; imagine yourself standing in front of a wall of dimension 4*m*

$\times 3m$. In the centre of the wall is a circular area (painted red), the radius of the circle being $1m$. If you are asked to fire a pistol at random against the wall, what is the probability that you will hit a point in the circular area ? You may assume that all points of the wall are equally likely to be hit by your shot. Once again, it is intuitively clear that since the area of the wall is $12m^2$ and the area of the painted circular area is πm^2, the probability of your hitting a red spot is $\frac{\pi}{12}$. You can therefore conclude that **size** in this case is simply the **area**.

Until now, we have dealt with three different cases where

(i) size means number of elements,

(ii) size means length of an interval,

(iii) size means the area.

All these are particular cases of a more general mathematical concept (associated with a particular family of sets) called **measure.**

To give you a sample of how these ideas are used in computing probabilities of events, consider the following example. Suppose that you and your friend decide to meet each other at the railway station on a particular Sunday between 5 p.m. and 6 p.m. You further decide that whoever comes first will wait for the other for just fifteen minutes. If you and your friend come to the station at random during that 60- minute period find the probability that you will meet your friend.

To solve this problem, let us make one simplification. Let us fix the origin of time at 5 p.m. and also take- the unit of time as minute. Let

x = time of arrival of your friend

y = time of your arrival

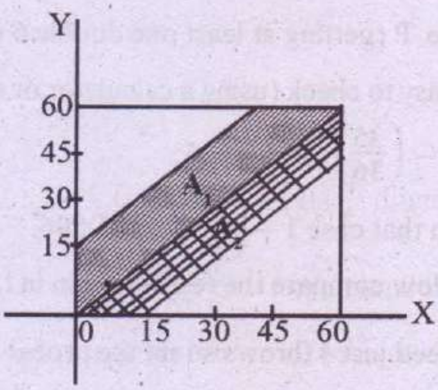
Since $x \in [0, 60]$ and $y \in [0, 60]$,

therefore the sample space S= [0, 60] × [0, 60], which is a square of side 60 units. Let A be the event that you meet your friend. This is possible only when

Case 1 : $x - y \le 15$; this means that your friend arrives first and you arrive within fifteen minutes of his arrival

Case 2 : $y - x \le 15$; this means that you arrive first and your friend arrives within fifteen minutes of your arrival.

Let us denote these mutually exclusive events by A_1 and A_2 respectively; clearly $A = A_1 \cup A_2$ and $P(A) = P(A_1) + P(A_2)$. Using ordinary coordinate geometry of dimension 2, it is obvious that the two straight lines $y = x$ and $y = x + 15$ enclose the area A_1 inside the square which has been shaded.



Therefore,

$$P(A_1) = \frac{\text{area of } A_1}{\text{area of S}} = \frac{\frac{1}{2}\times 60 \times 60 - \frac{1}{2}\times 45\times 45}{60 \times 60} = \frac{7}{32}.$$

One can similarly deduce that $P(A_2) = \frac{7}{32}$; hence $P(A) = \frac{7}{16}$.

We have already mentioned about the exchange of letters between Pascal and Fermat (1654). You can get some insight into the early stages of development of probability theory from the contents of one of those letters which dealt with a simple question.

(a) Suppose you throw a perfect die, then what is the minimum number of throws necessary so that the probability of getting at least one six exceeds $\frac{1}{2}$.

Well, this is an easy question to answer. Suppose you need n throws. Since the probability of not getting a 6 in one throw is $\frac{5}{6}$,

P (not getting 6 in all n throws) = $\left(\frac{5}{6}\right)^n$.

Hence, P(getting at least one 6 in n throws) = $1 - \left(\frac{5}{6}\right)^n$.

Since, $1 - \left(\frac{5}{6}\right)^n > \frac{1}{2}$, it is easy to deduce that $n = 4$. Thus,

P (getting at least one 6 in 4 throws) = $1 - \left(\frac{5}{6}\right)^4 \cong 0.518$.

(b) Suppose that you are now throwing a pair of dice; then what is the minimum number of throws necessary so that the probability of getting at least one double 6 exceeds $\frac{1}{2}$.

The answer to this question is equally simple. For each throw of a pair of dice, the probability of getting a double 6 is $\frac{1}{36}$, so that

P (not getting a double 6) = $\frac{35}{36}$.

P (not getting a double 6 in n throws) = $\left(\frac{35}{36}\right)^n$.

Hence, P (getting at least one double 6 in n throws) = $1 - \left(\frac{35}{36}\right)^n$.

It is easy to check (using a calculator or a logarithmic table) that the least value of n for which

$1 - \left(\frac{35}{36}\right)^n > \frac{1}{2}$ is 25;

and in that case $1 - \left(\frac{35}{36}\right)^{25} \cong 0.506$.

(c) Now compare the results given in (a) and (b). In (a), you are throwing one die and you need just 4 throws so tht the probability of getting at least one six exceeds $\frac{1}{2}$. In (b), you are throwing a pair of dice and you need 25 throws so that the probability of getting a double six exceeds $\frac{1}{2}$. On the surface of it isn't there an apparent contradiction ? At least, so it looked in 1654 A.D. !

6.3 Total Probability and Bayes' Theorem

We first introduce a fundamental concept, the partition of a sample space.

Definition :

Partition of a sample space

A set of events $\{E_1, E_2, ...E_n\}$ is said to form a partition of a sample space S, if

(i) $S = \bigcup_{i=1}^{n} E_i$

(ii) $E_i \neq \phi$, for i = 1, 2, ..., n

(iii) $E_i \cap E_j = \phi,\ i \neq j$

Theorem of Total Probability

Let the events E_i, i=1,2,...,n form a partition of the sample space S.

Then for $A \subseteq S$, we have $\mathbf{P(A) = \sum_{i=1}^{n} P(E_i).P(A \mid E_i)}$

Proof :

The second condition implies that $P(E_i) > 0$, so that

$P(A|E_i)$ is well defined for every i.

$$S = \bigcup_{i=1}^{n} E_i \Rightarrow A \cap S = A \cap \left(\bigcup_{i=1}^{n} E_i\right) = \bigcup_{i=1}^{n} (A \cap E_i) \quad \text{(by distributive law)}$$

From (iii) it follows that $P(A) = P\left(\bigcup_{i=1}^{n} (A \cap E_i)\right) = \sum_{i=1}^{n} P(A \cap E_i)$

By multiplication rule $P(A \cap E_i) = P(E_i).P(A \mid E_i)$

So $P(A) = \sum_{i=1}^{n} P(E_i).P(A \mid E_i)$

This complete the proof.

Example 17

There are two bags B_1 and B_2. B_1 contains 4 red and 3 black balls and B_2 contains 2 red and 4 black balls. A bag is selected at random and a ball is drawn from it.

Find the probability that the ball drawn is red.

Solution :

We have to determine the probability of drawing a red ball which we can do in the mutually exclusive ways, either from B_1 or from B_2.

Let us write :

E_1 = the event of selecting B_1

E_2 = the event of selecting B_2

A = event of drawing a red ball.

(Obviously the sample space S comprises B_1 and B_2 with all their balls which form a disjoint partition)

Now $P(A|E_1)$ = probability that the red ball is drawn from $B_1 = \frac{4}{7}$

Similarly $P(A|E_2) = \frac{2}{6} = \frac{1}{3}$

By the theorem of total probability,

$P(A) = P(E_1)P(A|E_1) + P(E_2)\,P(A|E_2)$

$= \frac{1}{2}.\frac{4}{7}+\frac{1}{2}.\frac{1}{3}=\frac{19}{42}$

Note on the Theorem of Total Probability

The expression $P(A)=\sum_{i=1}^{n}P(E_i).P(A\mid E_i)$ signifies that the occurrence of the events A is contingent upon the occurrence of the events E_1, E_2, ..., E_n which form a partition of the sample space.

Thus, if A is termed as an effect, then E_1, E_2, ..., E_n are its causes. The theorem of total probability gives the probability of an effect in terms of the probability of its causes.

The next theorem takes us to a reverse situation. Given that a particular phenomenon A (effect) has already occurred, the theorem gives us an estimate $P(E_i|A)$ (for $1 \le i \le n$), which is the probability of its occurrence due to a particular cause.

Bayes' Theorem

If a set of events $\{E_i | i=1, 2, ..., n\}$ form a partition of a sample space S,

$$\mathbf{P(E_i\mid A)=\frac{P(E_i)P(A\mid E_i)}{\sum_{j=1}^{n}P(E_j)P(A\mid E_j)}}\text{; provided } P(A)>0.$$

Proof :

$$P(E_i|A) = \frac{P(E_i \cap A)}{P(A)} \text{ (By definition of conditional probability)}$$

$$= \frac{P(E_i)P(A | E_i)}{P(A)} \text{ (By multiplication rule)}$$

$$= \frac{P(E_i)P(A | E_i)}{\sum_{j=1}^{n} P(E_j)P(A | E_j)} \text{ (By the theorem of total probability).}$$

N.B. The subscript 'j' under the summation sign is a dummy suffix in the sense that it can be replaced by any other suffix, even 'i'.

A Note on Bayes' Theorem

In contrast to the theorem of total probability, Bayes' theorem gives us an estimate of the probability of a particular cause having known that the event (effect) has already occurred.

Usually, in real life situations the probabilities of the causes E_i are not so easy to determine. They are mostly estimated depending upon experiments and experience. Hence $P(E_i)$ is known as 'a priori' probability and $P(E_i|A)$ is called 'a posteriori' probability of the cause E_i.

Example 18 (Coming back to example-17, reworded)

Threre are two bags B_1 and B_2 containing 4 red, 3 black balls and 2 red, 4 black balls respectively. If the ball drawn from a bag, selected at random, is red, find the probability that the ball is drawn from the bag B_1.

Solution :

With E_1, E_2 and A having the same meaning as before, applying Bayees' theorem.

Probability that the red ball is drawn from B_1 is given by

$$P(E_1|A) = \frac{P(E_1)P(A | E_1)}{P(E_1)P(A | E_1) + P(E_2)P(A | E_2)}$$

$$= \frac{\frac{1}{2} \times \frac{4}{7}}{\frac{1}{2} \times \frac{4}{7} + \frac{1}{2} \times \frac{2}{6}} = \frac{12}{19}$$

N.B. You can similarly calculate $P(E_2|A) = \frac{7}{19}$

Thus $P(E_1|A) + P(E_2|A) = 1$ which turns out to be a certainty !

But, infact, probability of drawing a red ball from either of the bags is $\frac{6}{13}$.

How do you account for this ?

Note that, in this problem, drawing of a red ball is not a nondeterministic event. It is given that a red ball has already been drawn, which is consequently a certainty.

The case is like this :

The event, say an epidemic has already broken out. We are simply ascertaining the probable reasons !

Further note that in problems as such there are multiple random selections; as in this case– a bag is randomly selected and then a ball is selected from the bag at random. This has to be borne in mind while working out problems.

EXERCISES - 6(c)

1. There are 3 bags B_1, B_2 and B_3 having respectively 4 white, 5 black; 3 white, 5 black and 5 white, 2 black balls. A bag is chosen at random and a ball is drawn from it. Find the probability that the ball is white.

2. There are 25 girls and 15 boys in class XI and 30 boys and 20 girls in class XII. If a student chosen from a class, selected at random, happens to be a boy, find the probability that he has been chosen from class XII.

3. Out of the adult population in a village 50% are farmers, 30% do business and 20% are service holders. It is known that 10% of the farmers, 20% of the business holders and 50% of service holders are above poverty line. What is the probability that a member chosen from any one of the adult population, selected at random, is above poverty line ?

4. Take the data of question number 3. If a member from any one of the adult population of the village, chosen at random, happens to be above poverty line, then estimate the probability that he is a farmer.

5. From a survey conducted in a cancer hospital it is found that 10% of the patients were alcoholics, 30% chew gutka and 40% have no specific carcinogenic habits. If cancer strikes 80% of the smokers, 70% of alcoholics, 50% of gutka chewers and 10% of the non specific, then estimate the probability that a cancer patient chosen from any one of the above types, selected at random,

   (i) is a smoker

   (ii) is alcoholic

   (iii) chews gutka

   (iv) has no specific carcinogenic habits.

6.4 Random Variable:

A random variable is a rule that assigns a numerical value to each possible outcome of an experiment. The term 'random variable' is actually a misnomer since a random variable X is a function whose domain is the sample space S and whose range is a subset of the set of real numbers.

Definition: A random variable X on a sample space S is a function $X : S \to R$. If the range of X is countable, i.e., in one to one correspondence with a subset of the set N of natural numbers, then the random variable X is a discrete random variable. In this chapter we shall be concerned only with discrete random variables.

Consider a random experiment consisting of three independent trials with outcomes either a success (S) or a failure (F) in each trial.

Let us denote a success by 1 and a failure by 0.

Then the sample space

$$S = \{(1, 1, 1), (1, 1, 0), (1, 0, 1), (0, 1, 1), (1, 0, 0), (0, 1, 0), (0, 0, 1), (0, 0, 0)\}$$

Any number of random variables can be defined on S. For example, we may consider the total number of successes as a random variable because it takes values either 3 or 2 or 1 or 0 each with an assigned probability.

Here $X(1, 1, 1) = 3, X(1, 1, 0) = 2, X(1, 0, 1) = 2$

$X(0, 1, 1) = 2, X(1, 0, 0) = 1, X(0, 1, 0) = 1$

$X(0, 0, 1) = 1$, and $X(0, 0, 0) = 0$.

Thus X takes values 3 or 2 or 1 or 0.

A probability can be assigned to each of these values of X.

Since each point of S is equally likely to occur their probabilities are equal and each is equal to $\frac{1}{8}$.

So $P(X=3)=\frac{1}{8}$, $P(X=2)=\frac{3}{8}$, $P(X=1)=\frac{3}{8}$, $P(X=0)=\frac{1}{8}$.

We observe that

$$\frac{1}{8}+\frac{3}{8}+\frac{3}{8}+\frac{1}{8}=1.$$

6.5 Probability Mass Function:

Let X be a discrete random variables which takes the possible values $x_1, x_2, \ldots\ldots, x_n$.

With each x_i we associate a real number

$p_i = P(X = x_i), i = 1,2,3,\ldots\ldots,n$

which is called the probability of $X = x_i$ satisfying the following conditions:

(i) $p_i \geq 0$ for each i,

(ii) $\sum_{i=1}^{n} p_i = p_1 + p_2 + ... + p_n = 1$

i.e. the total probability is 1.

To be more precise, if X is a discrete random variable we define

$$p(x) = P(X = x)$$

Such that $p(x) \geq 0$ and $\Sigma p(x) = 1$ summation being taken over all values of the variable. The function $p(x) = p(X = x)$ is called the **probability mass function** or simply the probability function of the random variable X. The set of all ordered pairs $(x, p(x))$ is called the **probability distribution of X.**

Example : A random variable X has the following distribution:

x:	0	1	2	3	4
$p(x)$:	0.1	0.3	k	0.2	3k

Determine the value of k so that it represents a probability distribution.

Solution: For the given distribution to be a probability distribution we must have

$$\Sigma p_i = 1.$$

This implies that

$$0.1 + 0.3 + k + 0.2 + 3k = 1$$

$$\Rightarrow 4k = 0.4$$

$$\Rightarrow k = 0.1$$

6.6 Mean and Variance of Random Variable

Let X be a random variable having the following probability distribution:

x :	x_1	x_2...............	x_3	x_n
$p(x)$:	p_1	p_2...............	p_3	p_n.

Then the **mean** of the random variable X dentoed by $\overline{x}$ is defined as

$$\overline{x} = \sum_{i=1}^{n} x_i p(x_i)$$

The **variance** of X dentoed by σ^2 is defined as

$$\sigma^2 = \sum_{i=1}^{n} (x_i - \overline{x})^2 p(x_i)$$

An alternative form of this can be obtained as follows:

We have

$$\sum_{i=1}^{n} (x_i - \overline{x})^2 p(x_i) = \sum_{i=1}^{n} (x_i^2 - 2\overline{x}x_i + \overline{x}^2) p(x_i)$$

$$= \sum_{i=1}^{n} x_i^2 p(x_i) - 2\overline{x} \sum_{i=1}^{n} x_i p(x_i) + \overline{x}^2 \sum_{i=1}^{n} p(x_i)$$

$$= \sum_{i=1}^{n} x_i^2 p(x_i) - 2\overline{x}^2 + \overline{x}^2$$

$$=\sum_{i=1}^{n} x_i^2 p(x_i) - \overline{x}^2$$

Thus $\sigma^2 = \sum_{i=1}^{n} x_i^2 p(x_i) - \overline{x}^2$

The positive square root of the variance of X is called the **standard deviation** of X and is denoted by σ.

The mean of a random variable X is also called the expectation of X and is denoted by E(X).

Example : Two cards are drawn successively with replacement from a well shuffled pack of 52 cards. Find the probability distribution of the number of kings. Also determine the mean and the variance of the number of kings.

Solution:

If X denote the number of kings in a successive draw of two cards with replacement from a deck of 52 cards then X is a random variable which takes values 0 or 1 or 2.

Since we draw the cards with replacement the two draws are independent.

So P(X=0) = P (no king and no king)

= P(no king) $\times$ P(no king)

$$=\frac{48}{52}\times\frac{48}{52}=\frac{144}{169}$$

P(X =1) = (a King and no King or no King and a King)

= P(a King and no King) + P (no King or a King)

$$=\frac{4}{52}\times\frac{48}{52}+\frac{48}{52}\times\frac{4}{52}$$

$$=\frac{24}{169}$$

P(X=2) = P(a King and a King)

= P(a King) $\times$ P(a King)

$$=\frac{4}{52}\times\frac{4}{52}=\frac{1}{169}$$

Thus the required probability distribution is

$X=x$	0	1	2
$p(x)$	$\frac{144}{169}$	$\frac{24}{169}$	$\frac{1}{169}$

$$\overline{x}=0\times\frac{144}{169}+1\times\frac{24}{169}+2\times\frac{1}{169}$$

$$= 0 + \frac{24}{169} + \frac{2}{169} = \frac{26}{169} = \frac{2}{13}$$

Variance $\sigma^2 = \sum_{i=1}^{3} x_i^2 p(x_i) - \overline{x}^2$

$$= 0^2 \times \frac{144}{169} + 1^2 \times \frac{24}{169} + 2^2 \times \frac{1}{169} - \left(\frac{2}{13}\right)^2$$

$$= 0 + \frac{24}{169} + \frac{4}{169} - \frac{4}{169}$$

$$= \frac{24}{169}$$

Example : If a pair of dice is thrown trice then find the mean and the variance of the number of doublets.

Solution: We denote the number of doublets by the random variable X.

Here X can take the values 0 or 1 or 2 or 3.

The possible doublets are

(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6) which are 6 in number.

So the probability of getting a doublet $= \frac{6}{36} = \frac{1}{6}$

and not getting a doublet = $= 1 - \frac{1}{6} = \frac{5}{6}$

Hence $P(X = 0) = = \frac{5}{6} \times \frac{5}{6} \times \frac{5}{6} = \frac{125}{216}$

$$P(X = 1) = \frac{1}{6} \times \frac{5}{6} \times \frac{5}{6} + \frac{5}{6} \times \frac{1}{6} \times \frac{5}{6} + \frac{5}{6} \times \frac{5}{6} \times \frac{1}{6}$$

$$= 3 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^2 = \frac{75}{216}$$

$$P(X = 2) = \frac{1}{6} \times \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \frac{5}{6} \times \frac{1}{6} + \frac{5}{6} \times \frac{1}{6} \times \frac{1}{6}$$

$$= \frac{15}{216}$$

$$P(X = 3) = \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \frac{1}{216}$$

Thus the required probability distribution is

$X = x$	0	1	2	3
$p(x)$	$\frac{125}{216}$	$\frac{75}{216}$	$\frac{15}{216}$	$\frac{1}{216}$

Mean $\overline{x} = \sum_{i=1}^{4} x_i p(x_i)$

$$= 0 \times \frac{125}{216} + 1 \times \frac{75}{216} + 2 \times \frac{15}{216} + 3 \times \frac{1}{216}$$

$$= 0 + \frac{75}{216} + \frac{30}{216} + \frac{3}{216}$$

$$= \frac{108}{216} = \frac{1}{2}$$

Variance $\sigma^2 = \sum_{i=1}^{4} x_i^2 p(x_i) - \overline{x}^2$

$$= 0^2 \times \frac{125}{216} + 1^2 \times \frac{75}{216} + 2^2 \times \frac{15}{216} + 3^2 \times \frac{1}{216} - \left(\frac{1}{2}\right)^2$$

$$= 0 + \frac{75}{216} + \frac{60}{216} + \frac{9}{216} - \frac{1}{4}$$

$$= \frac{144}{216} - \frac{1}{4} = \frac{2}{3} - \frac{1}{4} = \frac{5}{12}.$$

Example : A box containing 20 electric bulbs includes 5 defective bulbs. Four bulbs are drawn at random with replacement. Find the probability destribution of the number of non- defective bulbs. Calculate also the mean and the variance.

Solution: Let X be the random variable of the number of non-defective bulbs.

Clearly X takes values 0 or 1 or 2 or 3 or 4.

Number of non-defective bulbs in the box = 20 – 5 = 15

So the probability of getting a non-defective bulbs.

$$= \frac{15}{20} = \frac{3}{4}$$

Hence the probability of getting a defective bulb

$$= 1 - \frac{3}{4} = \frac{1}{4}$$

Consequently $P(X = 0) = \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} = \frac{1}{256}$

$$P(X = 1) = \frac{3}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} + \frac{1}{4} \times \frac{3}{4} \times \frac{1}{4} \times \frac{1}{4} + \frac{1}{4} \times \frac{1}{4} \times \frac{3}{4} \times \frac{1}{4} + \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{3}{4}$$

$$= 4 \times \frac{3}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} = \frac{12}{256}$$

$$P(X=2)=\frac{3}{4}\times\frac{3}{4}\times\frac{1}{4}\times\frac{1}{4}\times\frac{1}{4}\times\frac{3}{4}\times\frac{3}{4}\times\frac{1}{4}+\frac{1}{4}\times\frac{1}{4}\times\frac{3}{4}\times\frac{3}{4}+\frac{3}{4}\times\frac{1}{4}\times\frac{3}{4}\times\frac{1}{4}$$

$$+\frac{3}{4}\times\frac{1}{4}\times\frac{1}{4}\times\frac{3}{4}+\frac{1}{4}\times\frac{3}{4}\times\frac{1}{4}\times\frac{3}{4}$$

$$=6\times\frac{3}{4}\times\frac{3}{4}\times\frac{1}{4}\times\frac{1}{4}=\frac{54}{256}$$

$$P(X=3)=\frac{3}{4}\times\frac{3}{4}\times\frac{3}{4}\times\frac{1}{4}+\frac{1}{4}\times\frac{3}{4}\times\frac{3}{4}\times\frac{3}{4}+\frac{3}{4}\times\frac{1}{4}\times\frac{3}{4}\times\frac{3}{4}+\frac{3}{4}\times\frac{3}{4}\times\frac{1}{4}\times\frac{3}{4}$$

$$=4\times\frac{3}{4}\times\frac{3}{4}\times\frac{3}{4}\times\frac{1}{4}=\frac{108}{256}$$

$$P(X=4)=\frac{3}{4}\times\frac{3}{4}\times\frac{3}{4}\times\frac{3}{4}=\frac{81}{256}$$

Thus, the required probability distribution is

$X=x$	0	1	2	3	4
$p(x)$	$\frac{1}{256}$	$\frac{12}{256}$	$\frac{54}{256}$	$\frac{108}{256}$	$\frac{81}{256}$

Mean $\overline{x}=\sum_{i=1}^{5}x_i p(x_i)=0\times\frac{1}{256}+1\times\frac{12}{256}+2\times\frac{54}{256}+3\times\frac{108}{256}+4\times\frac{81}{256}$

$$=0+\frac{12}{256}+\frac{108}{256}+\frac{324}{256}+\frac{324}{256}$$

$$=\frac{768}{256}=3$$

Variance $=\sum_{i=1}^{5}x_i^2 p(x_i)-\overline{x}^2$

$$=0^2\times\frac{1}{256}+1^2\times\frac{12}{256}+2^2\times\frac{54}{256}+3^2\times\frac{108}{256}+4^2\times\frac{81}{256}-3^2$$

$$=0+\frac{12}{256}+\frac{216}{256}+\frac{972}{256}+\frac{1296}{256}-9$$

$$=\frac{2496}{256}-9=\frac{3}{4}$$

6.7 Bernoulli Trials :

Very often experiments are performed in which there are only two possible mutually exclusive outcomes. For example, in tossing a coin the outcomes are head or tail, in rolling a cubic die the possible outcomes are an even number or an odd number, in selecting an article the outcomes may be defective or non-defective. Under this situation one of the outcomes is called a success and other a failure. For example, in selecting an article if the outcome is non-detective we may call it a success and if the outcome is detective then we call it a failure.

Each time an experiment if performed is a trial.

Definition: Trials of a random experiment are called Bernoulli trials if the following conditions are satisfied:

(1) The number of trials is finite,

(2) Trials are independent

(3) The outcomes are dichotomous (success or failure)

(4) The probability of success (or failure) in each trial is constant.

Examples; Tossing a coin five times is an example of a Bernoulli trial.

Drawing 3 balls successively from a bag containing 5 red and 6 blue balls with replacement is also an example of a Bernoulli trial.

6.8 Binomial Distribution :

Consider an experiment resulting in two outcomes, a success (S) and a failure (F). Assume that the probability of a success is p and that of a failure is q so that p + q = 1. Suppose that the experiment is repeated n times under identical conditions so that the probability of a success in each trial is a constant, i.e., equal to p. In order to find the probability of x successes and hence $(n-x)$ failures in n independent trials we proceed as follows:

Assume that the first x trials result in success and the remaining $(n-x)$ trials result in failure. Then the sequence of outcomes will be

$$\underbrace{S\,S\,S\ldots\ldots S}_{x\ times}\ \underbrace{F\,F\,F\ldots\ldots\ldots F}_{(n-x)\ times}.$$

So $$\underbrace{P(S\,S\,S\ldots\ldots S}_{x\ times}\ \underbrace{F\,F\,F\ldots\ldots\ldots F)}_{(n-x)\ times}$$

= {(P(S) P(S) P(S) ..... to x factors} {P(F) P(F) P(F) ....P(F) .... to $(n-x)$ factors}

= $p^x q^{n-x}$ as the trials are independent.

But x successes in n independent trials can occur in ^nC_x different ways,

P(x successes) = $^nC_x p^x q^{n-x}, x = 01, 2, \ldots, n$

Since $^nC_x p^x q^{n-x}$ is the $(x+1)$th term in the expansion of $(q+p)^n$ the probability distribution of number of successes in an experiment consisting of n trials are the terms in the expansion of $(q+p)^n$.

This distribution can be expressed as

$X=x$	0	1	2	.	.	.	x	–	–	n
$p(X=x)$	$^nC_0q^n$	$^nC_1q^{n-1}p$	$^nC_2q^{n-2}p^2$				$^nC_xq^{n-x}p^x$	–	–	$^nC_np^n$

and is called the binomial distribution with parameters n and p. Here p is the probability function of the distribution.

This binomial distribution is usually denoted by the symbol B(n, p).

Example : If a fair coins is tossed 5 times then find the probability of getting

(i) Exactly three heads

(ii) at least three heads

(iii) at most three heads

Solution: This is a case of Bernoulli trials where getting a head in each trial is a success, we know that if

the number of trials = n

the probability of a success in each trial = p

and the probability of a failure in each trial = q,

then P(x successes) = $^nC_x q^{n-x} p^x$.

Here $n=5, p=\frac{1}{2}$ and $q=1-\frac{1}{2}=\frac{1}{2}$

So (i) P(exactly three heads)

= P(three successes)

$$= {}^5C_3\left(\frac{1}{2}\right)^{5-3}\left(\frac{1}{2}\right)^3$$

$$= {}^5C_3\frac{1}{2^5}=\frac{5!}{3!2!}.\frac{1}{32}=\frac{10}{32}=\frac{5}{16}$$

(ii) P(at least three heads)

= P(three successes) + P (four successes) + P (five successes)

$$= {}^5C_3\left(\frac{1}{2}\right)^{5-3}\left(\frac{1}{2}\right)^3+{}^5C_4\left(\frac{1}{2}\right)^{5-4}\left(\frac{1}{2}\right)^4+{}^5C_5\left(\frac{1}{2}\right)^{5-5}\left(\frac{1}{2}\right)^5$$

$$= {}^5C_3\frac{1}{2^5}+{}^5C_4.\frac{1}{2^5}+{}^5C_5.\frac{1}{2^5}$$

$$=\frac{5!}{3!2!}.\frac{1}{32}+\frac{5!}{4!1!}.\frac{1}{32}+1.\frac{1}{32}$$

$$=\frac{10+5+1}{32}=\frac{1}{2}$$

(iii) P(at most three heads)

= P(at most three successes)

= P(no success) + P(one success) + P(two successes) + P(three successes)

$$= {}^5C_0\left(\frac{1}{2}\right)^5+{}^5C_1\left(\frac{1}{2}\right)^{5-1}.\frac{1}{2}+{}^5C_2\left(\frac{1}{2}\right)^{5-2}\left(\frac{1}{2}\right)^2+{}^5C_3\left(\frac{1}{2}\right)^{5-3}\left(\frac{1}{2}\right)^3$$

$$=\frac{1}{2^5}+\frac{5}{2^5}+\frac{10}{2^5}+\frac{10}{2^5}$$

$$=\frac{1+5+10+10}{32}=\frac{26}{32}=\frac{13}{16}$$

Example : Four cards are drawn successively from a well-shuffled pack of 52 cards with replacement after each draw. Find the probability that

(i) all four cards are diamonds

(ii) only two cards are diamonds,

(ii) none is a diamond

Solution: This is a case of Bernoulli trials where getting a diamond in a draw is considered a success.

We know that if the number of trials = n,

the probability of a success in each trial = p

and the probability of a failure in each trial = q,

then P (x successes) = ${}^nC_x q^{n-x} p^x$

Here $n=4$, $p=\frac{13}{52}=\frac{1}{4}$ and $q=1-p=1-\frac{1}{4}=\frac{3}{4}$

So (i) P (four successes) = ${}^4C_4\left(\frac{3}{4}\right)^{4-4}\left(\frac{1}{4}\right)^4=\frac{1}{256}$

(ii) P (two successes) = ${}^4C_2\left(\frac{3}{4}\right)^{4-2}\left(\frac{1}{4}\right)^2$

$$=6.\frac{9}{16}.\frac{1}{16}=\frac{27}{128}$$

(iii) P(no success) = ${}^4C_0\left(\frac{3}{4}\right)^{4-0}\left(\frac{1}{4}\right)^0$

$$=1.\frac{81}{256}.1=\frac{81}{256}$$

Example : The probability of a shooter hitting a target is $\frac{4}{5}$. Find the minimum number of times he must fire so that the probability of hitting the target at least once is greater than 0.999.

Solution: Let the shooter fires n times. for each time the probability of the shooter hitting the target is considered as a success.

If p = the probability of hitting the target and

q = the probability of not hitting the target,

then $p=\frac{4}{5}$ and $q=1-\frac{4}{5}=\frac{1}{5}$.

We have P(x successes) = ${}^nC_x q^{n-x} p^x$

So P(at least one success)

$$=1-P\text{ (no success)}$$

$$=1-{}^{n}C_0\left(\frac{1}{5}\right)^n$$

$$=1-\frac{1}{5^n}>0.999$$

This implies that $\frac{1}{5^n}<0.001$

$$\Rightarrow 5^n>\frac{1}{0.001}=1000$$

$$\Rightarrow n\geq 5$$

Thus the shooter must fire minimum 5 times to hit the target with probability greater than 0.999.

6.9 Mean and Variance of Binomial Distribution

If the **mean** of the binomial distribution is denoted by $\overline{x}$, then

$$\overline{x}=\sum_{x=0}^{n}x\ {}^{n}C_x p^x q^{(n-x)}$$

$$=n_{C_1}pq^{(n-1)}+2\,{}^{n}C_2p^2q^{(n-2)}+3\,{}^{n}C_3p^3q^{(n-3)}+\dots+n\,{}^{n}C_np^n$$

$$=npq^{(n-1)}+n(n-1)p^2q^{(n-2)}+\frac{n(n-1)(n-2)}{\underline{|2}}p^3q^{(n-3)}+\dots+n\,{}^{n}C_np^n$$

$$=np\left[q^{n-1}+{}^{(n-1)}C_1pq^{n-2}+{}^{(n-1)}C_2p^2q^{n-3}+\dots+{}^{(n-1)}C_{(n-1)}p^{(n-1)}\right]$$

$$=np(q+p)^{n-1}$$

$$=np \qquad (\because\ q+p=1)$$

Variance of the Binomial Distribution

We have

$$\sum_{x=0}^{n}x^2p(x)=\sum_{x=0}^{n}\{x(x-1)+x\}\,p(x)$$

$$=\sum_{x=2}^{n}x(x-1)\,{}^{n}C_xp^xq^{(n-x)}+\sum_{x=1}^{n}x\,{}^{n}C_xp^xq^{(n-x)}$$

$$=2\,{}^{n}C_2p^2q^{n-2}+3.2\,{}^{n}C_3p^3q^{n-3}+\dots+n(n-1)\,{}^{n}C_np^n+np$$

$$=n(n-1)p^2q^{(n-2)}+\frac{n(n-1)(n-2)}{\underline{|1}}p^3q^{(n-3)}$$

$$+\frac{n(n-1)(n-2)(n-3)}{\underline{|2}}p^4q^{(n-4)}+\dots+n(n-1)p^n+np$$

$$= n(n-1)p^2[q^{(n-2)} + {}^{(n-2)}C_1 pq^{n-3} + {}^{(n-2)}C_2 p^2 q^{(n-4)} + ... + {}^{(n-2)}C_{(n-2)} p^{(n-2)}] + np$$

$$= n(n-1)p^2(q+p)^{n-2} + np$$

$$= n(n-1)p^2 + np$$

Variance $\sigma^2 = \sum_{x=0}^{n} x^2 p(x) - \bar{x}^2$

$$= n(n-1)p^2 + np - n^2p^2$$

$$= -np^2 + np$$

$$= np(1-p)$$

$$= npq$$

Example : For a binomial distribution the mean is 6 and the variance is 4. Obtain the probability distribution and hence find the

(i) probability of no success

(ii) Probability of at least one success

(iii) probability of at most two successes.

Solution: Given that $np = 6$ and $npq = 4$

where the number of Bernoulli trials = n

the probability of a success in one trial = p

and the probability of a failure in one trial = q

So $q = \frac{npq}{np} = \frac{4}{6} = \frac{2}{3}$

Hence $p = 1 - q = 1 - \frac{2}{3} = \frac{1}{3}$

$$np = 6 \Rightarrow \frac{n}{3} = 6$$

$$\Rightarrow n = 18$$

So the required binomial distribution is given by

$$p(x) = {}^{18}C_x \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{n-x}, \ x = 0,1,2,3,.....,18$$

Clearly (i) P(no success) $= {}^{18}C_0 \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^{18-0} = \left(\frac{2}{3}\right)^{18}$.

(ii) P(at least one success) $= 1 - p$ (no success)

$$= 1 - \left(\frac{2}{3}\right)^{18}$$

(iii) P(at most two successes)

$= \text{P(no success)} + \text{P(one success)} + \text{P(two successes)}$

$$= \left(\frac{2}{3}\right)^{18} + {}^{18}C_1 \frac{1}{3}\left(\frac{2}{3}\right)^{17} + {}^{18}C_2\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^{16}$$

$$= \left[\left(\frac{2}{3}\right)^2 + {}^{18}C_1 \frac{1}{3}.\frac{2}{3} + {}^{18}C_2.\frac{1}{9}\right]\left(\frac{2}{3}\right)^{16}$$

$$= \left(\frac{4}{9} + 4 + 17\right)\left(\frac{2}{3}\right)^{16}$$

$$= \frac{193}{9}\left(\frac{2}{3}\right)^{16}$$

Exercise - 6 (d)

1. State which of the following is the probability distribution of a random variable X with reasons to your answer.

(a)

$X = x$	0	1	2	3	4
$p(x)$	0.1	0.2	0.3	0.4	0.1

(b)

$X = x$	0	1	2	3
$p(x)$	0.15	0.35	0.25	0.2

(c)

$X = x$	0	1	2	3	4	5
$p(x)$	0.4	R	0.6	R^2	0.7	0.3

2. Find the probability distribution of number of doublets in four throws of a pair of dice. Find also the mean and the variance of the number of doublets.
3. Four cards are drawn successively with replacement from a well shuffled pack of 52 cards. Find the probability distribution of the number of aces. Calculate the mean and variance of the number of aces.
4. Find the probability distribution of
   (a) number of heads in three tosses of a coin
   (b) number of heads in simultaneous tosses of four coins.
5. A biased coin where the head is twice as likely to occur as the tail is, tossed thrice. Find the probability distribution of number of heads.
6. Find the probability distribution of the number of aces in question no.3 if the cards are drawn successively without replacement.
7. From a box containing 32 bulbs out of which 8 are defective 4 bulbs are drawn at random successively one after another with replacement. Find the probability distribution of the number of defective bulbs.

8. A random variable X has the following probability distribution.

$X=x$	0	1	2	3	4	5
$p(x)$	0	R	$2R$	$3R$	$3R$	R

Determine (a) R (b) P(X<4) (c) $P(X \geq 2)$ (d) $P(2 \leq X \leq 5)$

9. Find the mean and the variance of the number obtained on a throw of an unbiased coin.
10. A pair of coins is tossed 7 times. Find the probability of getting
    (i) exactly five tails
    (ii) at least five tails
    (iii) at most five tails
11. If a pair of dice is thrown 5 times then find the probability of getting three doublets.
12. Four cards are drawn successively with replacement from a well-shuffled pack of 52 cards. What is the probability that
    (i) all the four cards are diamonds
    (ii) only two cards are diamonds
    (ii) none of the cards is a diamond.
13. In an examination there are twenty multiple choice questions each of which is supplied with four possible answers. What is the probability that a candidate would score 80% or more in the answers to these questions?
14. A bag contains 7 balls of different colours. If five balls are drawn successively with replacement then what is the probability that none of the balls drawn is white?
15. Find the probability of throwing at least 3 sixes in 5 throws of a die.
16. The probability that a student securing first division in an examination is $\frac{1}{10}$. What is the probability that out of 100 students twenty pass in first diyision?
17. Sita and Gita throw a die alternatively till one of them gets a 6 to win the game. Find their respective probability of winning if Sita starts first.
18. If a random variable X has a binomial distribution $B\left(8, \frac{1}{2}\right)$ then find X for which the outcome is the most likely.
    Hint: Find X = x for which P(X = x) is the maximum, x = 0, 1, 2, 3, ...,8

●

CHAPTER - 7

Continuity and Differentiability

The infinite! No other question has ever moved so profoundly the spirit of man.

- David Hilbert

7.1 Continuity

Let us measure the distance covered by a vehicle from a given point. Let the function $f: \Re \rightarrow \Re$ be such that $f(t)$ is the distance covered by the vehicle by the time t.

Let the graph of the function f be as follows :

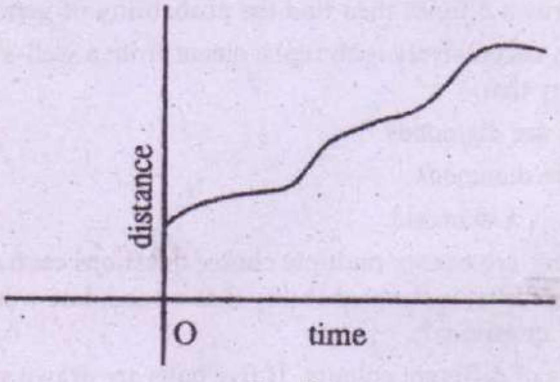


Fig. 7.1

Take another example :

If $P(w)$ stands for the postage necessary to send an ordinary letter with weight w then $P: \Re_+ \rightarrow \Re$ is a function whose values as of 2002 are as follows :

$$P(w) = \begin{cases} 5 \text{ for } 0 < w \leq 20 \\ 10 \text{ for } 20 < w \leq 40 \\ 15 \text{ for } 40 < w \leq 60. \end{cases}$$

which says that the postage is Rs. 5/- for first twenty grams or part there of and Rs.5/- extra for every extra 20 grams or part thereof.

How will its graph look like ?

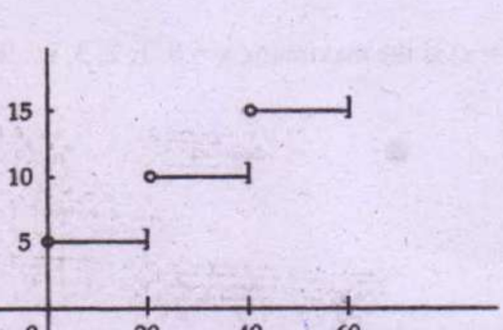


Fig. 7.2

What difference do we see in the graphs of the two functions described ? In the first case the change was **gradual** at every point whereas in the second one the change is **abrupt** at certain points. We try to describe what these words gradual and abrupt mean in this context. A change in the value of a function will be called gradual if small changes in the value of the function $f(t)$ can be brought about by a small change in the value of the argument t. What it really means is that

for every $\varepsilon > 0$ we can find a $\delta > 0$ such that

$$|f(t) - f(t_0)| < \varepsilon \text{ for } |t - t_0| < \delta$$

This point t_0 is called the point of continuity of the function f at $t=t_0$.

In fig.7.2, if $20 < t_0 < 40$, then $f(t)-f(t_0) = 0$ as it is a point of continuity. If $t_0 = 40$, then the above equation breaks down (verify). Here t_0 is a point of discontinuity.

This establishes an obvious connection between limit and continuity :

$f : [a, b] \to \Re$ is continuous at $c \in [a, b]$ if

$\lim_{x \to c} f(x)$ exists and is equal to $f(c)$;

in other words, if $\lim_{x \to c} f(x) = f(c)$.

If for every $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|f(x) - f(c)| < \varepsilon \text{ for } c \le x < c + \delta,$$

then we say f is **right continuous at** c.

On the other hand, if for every $\varepsilon > 0$, $\exists\ \delta > 0$ such that

$$|f(x) - f(c)| < \varepsilon \text{ for } c - \delta < x \le c,$$

then we say that f is **left continuous at** c.

To illustrate this point let us take an example. Suppose we have an artillery gun with which we want to hit a target. We know when the cannon is pointed in the direction of the target making an angle θ with the horizontal, as long as the gun and the shell being fired is fixed the distance of the point is dependent only on θ. We can say that the range at the angle θ is $f(\theta)$. In other words the range is a function of θ. Suppose we know that the target is at distance t from the gun and $f(\theta_0) = t$. Then the best result is achieved if the gun is trained to make an angle θ_0 with horizontal. Our experience tells that while fixing the angle there is likely to be little deviation in fixing the gun making an angle θ_0 with the horizontal. This would obviously mean that our shell will miss the target. But we know that the shell, when hits a point explodes causing damage to a limited region around the point. Suppose ε is the distance with which the shell is effective in the sense that we get the desired result if the target is within a distance ε from the point of hit. Now the question is how much of play should be allowed in training our gun to get an effective hit. That is what is the maximum deviation of θ allowed for the gun from the angle θ_0 so that

$$|f(\theta) - f(\theta_0)| < \varepsilon .$$

So if f is a continuous function of θ, then for the given $\varepsilon > 0$ we can find a $\delta > 0$ such that

$$|f(\theta) - f(\theta_0)| < \varepsilon \text{ whenever } |\theta - \theta_0| < \delta,$$

that is, if the angle θ is not allowed to deviate from θ_0 by more than δ then our desired effectiveness of the fire would be achieved.

Thus many situations are encountered in experience as continuous processes such as flow of wind, flow of time, flow of liquid etc. The formulation of the concept of continuity enables us to give a mathematical structure for these situations. So it is necessary to study the notion of continuity of a function in a more formal way.

Definition :

A function f is said to be **continuous** at a point $a \in D_f$ if

(i) $f(x)$ has definite value $f(a)$ at $x = a$,

(ii) $\lim_{x \to a} f(x)$ exists,

(iii) $\lim_{x \to a} f(x) = f(a)$.

If one or more of the above conditions fail, the function f is said to be **discontinuous** at $x = a$.

The above definition of continuity of a function at a point can also be formulated as follows :

A function f is said to be continuous at $x = a$ if

(i) holds and for a given $\in > 0$, there exists a $\delta > 0$ depending on $\in$ such that

$| x - a | < \delta \Rightarrow | f(x) - f(a) | < \in$.

A function f is continuous on an interval if it is continuous at every point of the interval.

If the interval is a closed interval $[a, b]$ the function f is continuous on $[a, b]$ if it is continuous on (a, b), $\lim_{x \to a+} f(x) = f(a)$ and $\lim_{x \to b-} f(x) = f(b)$.

Note : The expression $\lim_{x \to a} f(x) = f(a)$ can be written in the form $\lim_{h \to 0} f(a + h) = f(a)$ by putting $x = a + h$.

Let us consider a continuous curve C which is the graph of a continuous function $y = f(x)$ in the interval $[a, b]$. We use the term "Continuous Curve" in an intuitive sense to mean that it can be drawn with a continuous motion of a pencil without lifting it from the paper.

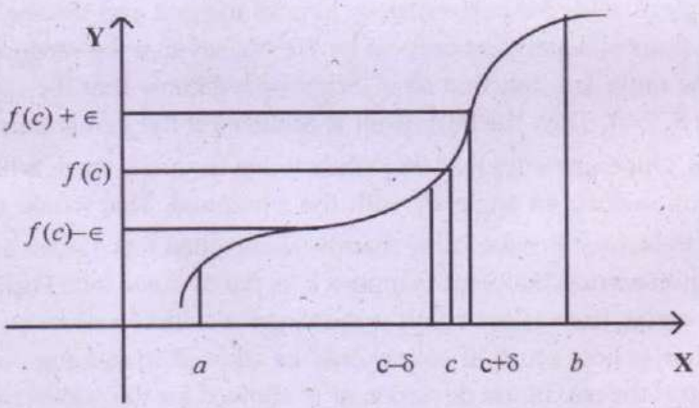


Fig. 7.3

The above diagram is a pictorial representation of the statement :

$| x - c | < \delta \Rightarrow | f(x) - f(c) | < \in$,

i.e. $c - \delta < x < c + \delta \Rightarrow f(c) - \in < f(x) < f(c) + \in$ exhibiting the continuity of the function f at $x = a$.

Thus the definition of continuity of a function is in agreement with the intuitive idea of a continuous curve.

If a function is discontinuous at a point, then intuitively we mean that there is a gap or jump in the graph of the function at that point.

We consider the following examples :

7.2 Continuity of some real valued functions :

Real valued functions along with their graphs have already been discussed in section 3.4, Vol-I. We now discuss their continuity.

Example 1 :

Examine the continuity of $|x|$.

Solution :

$$|x| = \begin{cases} x \text{ if } x \geq 0 \\ -x \text{ if } x < 0 \end{cases}$$

So $\lim\limits_{x \to 0-} |x| = \lim\limits_{x \to 0-} (-x) = 0$

and $\lim\limits_{x \to 0+} |x| = \lim\limits_{x \to 0+} x = 0.$

Since $\lim\limits_{x \to 0-} |x| = \lim\limits_{x \to 0+} |x| = 0$, $\lim\limits_{x \to 0} |x| = 0.$

Thus $\lim\limits_{x \to 0} |x| = 0.$

So $|x|$ is continuous at $x = 0$.

The graph of the function is given in fig 7.4.

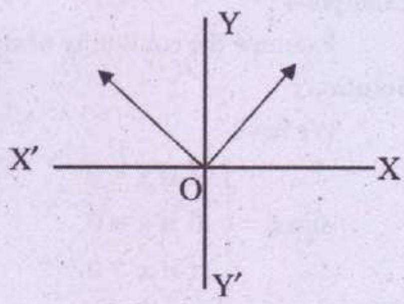


Fig. 7.4

Example 2 :

Examine the continuity of the function f defined by

$$f(x) = \begin{cases} 2, & x \leq 0 \\ x^2 + 2, & 0 < x < 1 \\ 3, & x \geq 1 \end{cases}$$

at x = 0.

Solution :

$\lim\limits_{x \to 0^-} f(x) = 2$

$\lim\limits_{x \to 0^+} f(x) = \lim\limits_{x \to 0^+} x^2 + 2 = 2$

and $f(0) = 2$.

So f(x) is continuous at x = 0.

The graph of this function is given in fig 7.5

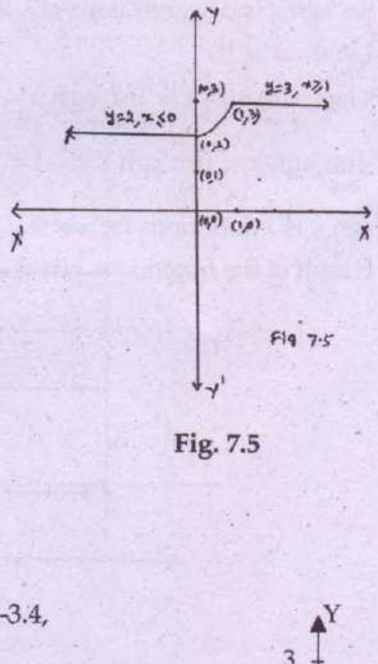


Fig. 7.5

Example- 3 :

Discuss the continuity of f(x) = [x].

(Recall greatest integer function, Example-22, Section-3.4, Vol.-I)

Solution :

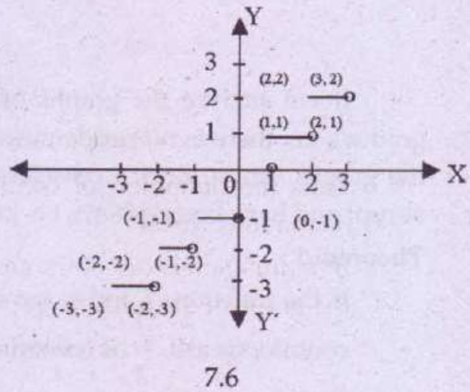
Clearly the function is defined for every real number.

Case-1 Suppose x = n, an integer.

In example – 10 of section 14.3 (Vol.-I), we have seen that $\lim\limits_{x \to n}$ [x] does not exist for any integer n.

So [x] is discontinuous at every integer n.

Case-2 Let $x = a$ be a real number which is not an integer.



7.6

Then $\lim_{x \to a} f(x) = \lim_{x \to a} [x] = [a]$

Also f(a) = [a]

Thus $\lim_{x \to a} f(x) = f(a)$

So f(x) = [x] is continuous everywhere except at integer-points.

Graph of this function is given in Fig.7.6.

Example-4 :

Examine the continuity of signum function f(x) = sgn x.

Solution :

We have

$$\operatorname{sgn} x = \begin{cases} -1 \text{ if } x < 0 \\ \ \ 0 \text{ if } x = 0 \\ \ \ 1 \text{ if } x > 0 \end{cases}$$

(Compare with the definition given in example-19 of Section 3.4 Vol.-I)

First we observe that f is defined for all real numbers.

Case-1 Let x = 0

In example – 9 of section 14.3 Vol.-I, we have seen that $\lim_{x \to 0}$ sgn x does not exist.

So sgn x is discontinuous at x = 0.

Case-2 Let $x = a \neq 0$

Then $\lim_{x \to a^-} \operatorname{sgn} x = \lim_{x \to a^+} \operatorname{sgn} x = 1 = f(a) \text{ for } a > 0$

and $\lim_{x \to a^-} \operatorname{sgn} x = \lim_{x \to a^+} \operatorname{sgn} x = -1 = f(a) \text{ for } a < 0$

So sgn x is continuous for $x \neq 0$.

Graph of the function is given in fig 7.7.

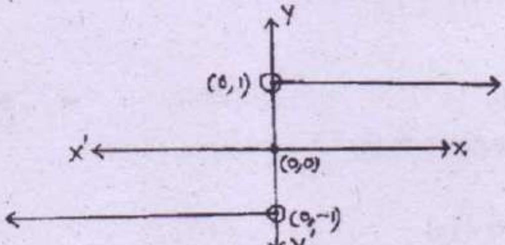


Fig.-7.7

If we analyze the graphs of the functions. In examples-1 and 2, we observe that they are all gradual, i.e. there is no suddenness in them. They do not have break-points or jumps. In other words, the graphs can be drawn ceaselessly without lifting the pencil. The graphs in examples-3 and 4 are abrupt and have break-points, i.e. jumps.

Theorem 1 :

If the functions f and g are continuous at a point a, then $f + g$, $f - g$, cf (c is a constant), fg are continuous at a. $\frac{f}{g}$ is continuous at a if $g(a) \neq 0$.

Proof :

Since f and g are continuous at a, $\lim_{x \to a} f(x) = f(a)$ and $\lim_{x \to a} g(x) = g(a)$.

$$\text{So } \lim_{x \to a} (f+g)(x) = \lim_{x \to a} (f(x) + g(x))$$
$$= \lim_{x \to a} f(x) + \lim_{x \to a} g(x)$$
$$= f(a) + g(a)$$
$$= (f+g)(a).$$

Hence $f + g$ is continuous at a.

We leave the proofs of the other parts as an exercise for the reader.

Theorem 2 :

If a function f is continuous at a point a and a function g is continuous at the point $b = f(a)$, then the composite function gof is continuous at a.

In this case $\lim_{x \to a} (gof)(x) = \lim_{f(x) \to f(a)} g(f(x))$

$= g(f(a))$

$= (g \text{ of })(a)$.

The proof of this theorem is the direct application of the definition of limit.

Note: If a function f is obtained from several functions by a finite number of arithmetical operations and the operations of forming a function of a function (composite function) the verification of continuity of f at a given point can be done by the successive applications of the last two theorems provided that these theorems are (used a finite number of times) applicable.

Theorem 3 :

If f is continuous at a, then $|f|$ is continuous at a, for all real values of a, but not conversely.

Proof :

Let f be continuous at a.

By definition, for a given $\varepsilon > 0$, $\exists$ a $\delta > 0$, such that

$|f(x) - f(a)| < \varepsilon$ whenever $|x-a| < \delta$

$\| f(x)| - |f(a)\| \le |f(x) - f(a)| < \in$, whenever $|x-a| < \delta$

$\left(\because \| a| - |b\| \le |a-b|\right)$.

So $|f|$ is continuous at any $a \in R$.

To show that the converse is not true, we consider an example :

$$f(x) = \begin{cases} 1, & if\ x \ge 0 \\ -1, & if\ x < 0 \end{cases}$$

so that $|f|(x) = |f(x)| = 1\ \forall x \in R$.

Then $\lim_{x \to 0} |f|(x) = 1 = |f|(0)$

$\Rightarrow$ $|f|$ is continuous at 0.

However, $\lim_{x \to 0} f(x)$ does not exist as $\lim_{x \to 0^-} f(x) = -1$ and $\lim_{x \to 0^+} f(x) = 1$

Hence f is not continuous at 0.

Continuity of trigonometric functions

Continuity of sin$x$ and cos$x$

Each of sin$x$ and cos$x$ is a continuous function of x for all real x. We prove below the continuity of the sine function.

Let us consider the continuity of sin$x$ at an arbitrary point $x = a \in R$.

We have
$$|\sin x - \sin a| = \left|2\cos\frac{x+a}{2}\sin\frac{x-a}{2}\right|$$
$$= 2\left|\cos\frac{x+a}{2}\right|\left|\sin\frac{x-a}{2}\right|$$
$$\le 2\left|\frac{x-a}{2}\right|$$
$$= |x-a|.$$

($\because |\cos\theta| \le 1$ and $|\sin\theta| \le |\theta|$; refer theorem-6, section-14.6, Vol-I)

Let $\in$ be an arbitrarily small positive number. Take $\delta = \in$

Then $|x-a| < \delta \Rightarrow |\sin x - \sin a| < \in$.

So $\sin x \to \sin a$ as $x \to a$,

i.e. $\lim\limits_{x \to a} \sin x = \sin a$.

Hence sin$x$ is continuous at $x = a$.

Since a is an arbitrary real number, it follows that sin$x$ is continuous for every real x.

The proof of continuity of cosine function is left to the reader as an exercise.

Continuity of tanx

We have $\tan x = \dfrac{\sin x}{\cos x}$. It is defined for all real values of x, such that $\cos x \neq 0$ for which we must have $x \neq (2n+1)\dfrac{\pi}{2}$, $n \in Z$.

Therefore tanx is continuous for all real x except $x=(2n+1)\dfrac{\pi}{2}$, $n \in Z$.

The continuity of secx, cosecx and cotx are left to the reader as exercises.

Example-5 :

Examine the continuity of sin(sinx).

Solution :

We observe that the function is defined for every real number x and it is a composition fog where f(x) = sinx and g(x) = sinx. Since both f and g are continuous functions (by theorem-2),

$(fog)(x) = f(g(x)) = \sin(\sin x)$ is a continuous function.

Continuity of an exponential function

Recall that the exponential function f is defined by $f(x) = a^x \ \forall \ x \in \mathrm{R}$ where $a > 0$ and $a \neq 1$ (ref. Section 3.4, Vol-I, Example-20). a is called the base of the exponential function.

We state without proof an important result :

a^x is continuous for every real x ($a > 0, a \neq 1$).

Corollary : e^x is continuous $\forall \ x \in \mathrm{R}$.

Continuity of Logarithmic Function :

Continuity of $\log_a x$, ($a > 0$ and $a \neq 1$) and $x > 0$

The logarithmic function has been discussed in example-21, Section-3.4, Vol-I.

For the proof of the continuity of the logarithmic function we quote below a theorem whose proof appears in higher mathematics.

Theorem 4 :

If a function defined by $y = f(x)$ is continuous and strictly increasing over an open interval (a, b), then the inverse function defined by $x = f^{-1}(y)$ is strictly increasing and continuous on the range of f.

Remark : The above theorem remains valid if "increasing" is replaced by "decreasing". The theorem also holds if the intervals are infinite intervals.

If $a > 1$, then $y = a^x$ is a continuous and strictly increasing function defined on $\mathrm{R} = (-\infty, \infty)$.

The range of this function is $(0, \infty)$. The logarithm function defined by $x = \log_a y$ is the inverse of the function $y = a^x$ with domain $(0, \infty)$ and range $(-\infty, \infty)$. So the logarithm function is strictly increasing and continuous on $(0, \infty)$ if the base a is greater than 1.

If $0 < a < 1$, the argument is quite similar with the word "increasing" replaced by "decreasing".

Thus $\log_a x$ is a strictly increasing and continuous function of x on $(0, \infty)$ by the use of the preceding theorem.

So $\lim\limits_{x \to x_0} \log_a x = \log_a x_0 \ \forall \ x_0 \in (0, \infty)$.

The Continuity of power function x^α

Let us consider the function f defined by $f(x) = x^\alpha$, where α is a real constant. For any given $\alpha \in \mathrm{R}$ this function is defined and is positive for $x > 0$.

We have $x^\alpha = e^{\alpha \log_e x}$, where $x > 0$.

Since exponential and logarithm functions are continuous functions it follows that x^α is a continuous function of $x \ \forall \ x > 0$.

If $\alpha > 0$, the function x^α is continuous from the right at $x = 0$.

Since x^α is a continuous function of x, $\forall \ x > 0$, where α is a real constant

$\lim\limits_{x \to a} x^\alpha = a^\alpha \ \forall \ a > 0$.

Continuity of Polynomial function

We recall that a polynomial function is defined by

$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$ where n is a non negative integer, $a_0, a_1, \ldots, a_n \in \mathrm{R}$, $a_n \neq 0$

(Refer Example -16, section 3:4, Vol-I)

Since power function is continuous, it follows from theorem-1 that polynomial function is continuous.

Continuity of Rational Function

A rational function is defined by

$$f(x)=\frac{p(x)}{q(x)},\ q(x)\neq 0$$

Where $p(x)$ and $q(x)$ are polynomial functions.

Clearly the domain of f is the set of all real numbers except those for which $q(x)=0$.

It follows from continuity of polynomial functions that f is continuous in its domain.

Example 6 :

Examine continuity of the function f defined by

$$f(x)=\begin{cases}\dfrac{\sin 2x}{x} & \text{if } x\neq 0\\ 1 & \text{if } x=0\end{cases}$$

at $x = 0$.

Solution :

We have $f(o) = 1$

$$\lim_{x\to 0} f(x)=\lim_{x\to 0}\frac{\sin 2x}{x}=2\lim_{x\to 0}\frac{\sin 2x}{2x}=2$$

Since $\lim_{x\to 0} f(x)\neq f(o)$, the given function is discontinuous at $x = 0$.

Example 7 :

Examine the continuity of the function f defined by

$$f(x)=\begin{cases}2x+1 & \text{if } x<1\\ 0 & \text{if } x=1\\ x^2-1 & \text{if } x>1\end{cases}$$

Solution :

From the definition of the function f it is easy to see that f is continuous for $x < 1$ as well as for $x > 1$.

The point $x = 1$ is a possible point of discontinuity of the function f.

We have $f(1) = 0$.

$$\lim_{x\to 1-} f(x)=\lim_{x\to 1-}(2x+1)=3.$$

$$\lim_{x\to 1+} f(x)=\lim_{x\to 1+}(x^2-1)=0.$$

Since $\lim_{x\to 1-} f(x)\neq\lim_{x\to 1-} f(x)$, $\lim_{x\to 1} f(x)$, does not exist.

So the given function is discontinuous at $x = 1$.

Some facts based on continuity of functions

We state below, without proof, some important facts which can be derived as properties of continuous functions. They can also be realized by taking into consideration the graphs of the functions concerned.

Facts :

1. If f is continuous at $x = a$ and $f(a) \neq 0$, then there exists an open interval $(a - \delta, a+\delta)$, such that $\forall x \in (a - \delta, a+\delta)$, f(x) has the same sign as $f(a)$.

2. If f is a continuous function defined on $[a, b]$, such that $f(a)\, f(b) < 0$, then there exists at least one solution of the equation $f(x) = 0$ in the open interval (a, b).

3. If f is a continuous function defined on $[a, b]$ and k is any real number between f(a) and f(b), then there exists at least one solution of the equation $f(x) = k$ in the open interval (a, b).

4. If a function f is continuous on a closed interval $[a, b]$, then it is also bounded on $[a, b]$, i.e. there is a real number k, such that $|f(x)| \leq k$.

5. If f is continuous at every point of a closed interval I, then it assumes both an absolute maximum value M and an absolute minimum value m somewhere in I, i.e. there are numbers x_1 and x_2 in I with $f(x_1) = m$, $f(x_2)=M$, and $m \leq f(x) \leq M$ for all x in I .

Some exercises using the above properties :

Example-8

Prove that $x = \cos x$ for some $x \in (0, \pi/2)$.

Solution :

Let $f(x) = x - \cos x$

$f(0) = -1$, $f(\pi/2) = \pi/2 - \cos \pi/2 = \pi/2$

$\Rightarrow f(0)f(\pi/2) < 0$

$\Rightarrow \exists\, x \in (0, \pi/2)$ such that $f(x) = 0$, i.e. $x - \cos x = 0$ (ref. fact-2)

So $x = \cos x$ for some $x \in (0, \pi/2)$.

Example-9

Prove that the expression 2^x+x^2 attains the value 2 for some value of x between 0 and 1.

Solution :

Let $f(x) = 2^x+x^2$

f is a continuous function for all x in R, hence, necessarily in [0,1].

$f(0) = 1$ and $f(1) = 3$

$f(0) < 2 < f(1)$

$\Rightarrow \exists x \in (0,1)$ such that $f(x) = 2$ (by fact-3).

This proves the result.

EXERCISE 7(a)

1. Examine the continuity of the following functions at indicated points.

(i) $f(x)=\begin{cases}\dfrac{x^2-a^2}{x-a} & \text{if } x\neq a\\ a & \text{if } x=a\end{cases}$ at $x=a$

(ii) $f(x)=\begin{cases}\dfrac{\sin 2x}{x} & \text{if } x\neq 0\\ 2 & \text{if } x=0\end{cases}$ at $x=0$

(iii) $f(x)=\begin{cases}(1+2x)^{\frac{1}{x}} & \text{if } x\neq a\\ e^2 & \text{if } x=0\end{cases}$ at $x=0$

(iv) $f(x)=\begin{cases}x\sin\dfrac{1}{x} & \text{if } x\neq a\\ 0 & \text{if } x=0\end{cases}$ at $x=0$

(v) $f(x)=\begin{cases}\dfrac{x^2-1}{x-1} & \text{if } x\neq a\\ 0 & \text{if } x=0\end{cases}$ at $x=0$

(vi) $f(x)=\begin{cases}\sin\dfrac{1}{x} & \text{if } x\neq a\\ 0 & \text{if } x=0\end{cases}$ at $x=0$

(vii) $f(x)=[3x+11]$ at $x=-\dfrac{11}{3}$

(viii) $f(x)=\begin{cases}\dfrac{e^{\frac{1}{x}}-1}{e^{\frac{1}{x}}+1} & \text{if } x\neq a\\ 0 & \text{if } x=0\end{cases}$ at $x=0$

(ix) $f(x)=\begin{cases}\dfrac{1}{x+[x]} & \text{if } x<a\\ -1 & \text{if } x\geq 0\end{cases}$ at $x=0$

(x) $f(x)=\begin{cases}\dfrac{|x|}{x} & \text{if } x\neq a\\ 0 & \text{if } x=0\end{cases}$ at $x=0$

(xi) $f(x)=\begin{cases} 2x+1 & if\ x \le 0 \\ x & if\ 0<x<1 \\ 2x-1 & if\ x \ge 1 \end{cases}$ at $x=0,\ 1$

(xii) $f(x)=\begin{cases} \dfrac{1}{e^{\frac{1}{x}}-1} & if\ x>0 \\ 0 & if\ x \le 0 \end{cases}$ at $x=0$

(xiii) $f(x)=\text{Sin}\dfrac{\pi\,[x]}{2}$ at $x=0$

(xiv) $f(x)=\dfrac{g(x)-g(1)}{x-1}$ at $x=1$, where $g(x)=|x-1|$

2. If a function f is continuous at $x=a$, then find

(i) $\lim_{h\to 0}+\dfrac{1}{2}\{f(a+h)+f(a-h)\}$

(ii) $\lim_{h\to 0}+\dfrac{1}{2}\{f(a+h)-f(a-h)\}$

3. Find the value of a such that the function f defined by

$$f(x)=\begin{cases} \dfrac{\sin ax}{\sin x} & if\ x\neq 0 \\ \dfrac{1}{a} & if\ x=0 \end{cases}$$

Is continuous at x = 0.

4. If $f(x)=\begin{cases} ax^2+b & if\ x<1 \\ 1 & if\ x=1 \\ 2ax-b & if\ x>1 \end{cases}$

Is continuous at x=1, then find a and b.

5. Show that sin x is continuous for every real x.

6. Show that the function f defined by

$$f(x)=\begin{cases} 1\ if\ x\ is\ rational \\ 0\ if\ x\ is\ irrational \end{cases}$$

is discontinuous $\forall \neq 0 \in \text{R}$.

7. Show that the function f defined by

$$f(x)=\begin{cases} x & \text{if } x \text{ is rational} \\ -x & \text{if } x \text{ is irrational} \end{cases}$$

Is continuous at x=0 and discontinuous $\forall x \neq 0 \in \mathbb{R}$.

8. Show that the function *f* defined by

$$f(x) = \begin{cases} x & \text{if } x \text{ is rational} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$$

Is discontinuous everywhere except at $x = 0$.

9. Show that $f(x) = \begin{cases} x\sin\frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$

is continuous at $x = 0$.

10. Prove that $e^x - 2 = 0$ has a solution between 0 and 1

(Hints : Use continuity of e^x-2 and fact-2)

11. So that $x^5+x+1 = 0$ for some value of x between -1 and 0.

7.3 Differentiability

In the chapter on limits and derivatives of Vol.-I, we have already defined the derivative of a function. We have learnt to differentiate certain functions. Before going further into derivatives and their applications, let us recapitulate the following ideas which have already been discussed in Vol.-I.

Derivative :

A function is said to be derivable or differentiable at x ∈ (*a*, b) if $\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$ exists and this limit is denoted as f′(x) or $\frac{df}{dx}$ or Df. D or $\frac{d}{dx}$ is called the differential operator. The derivative of $y = f(x)$ is also denoted by $\frac{dy}{dx}$ or y' or D*y*.

The process of finding derivative is called differentiation or derivation.

Using the definition of derivative, we have derived the following standard results.

(Ref. Sec.-14.12, Vol.-I) :

$\frac{d}{dx}x^{\alpha} = \alpha x^{\alpha-1}$, α ∈ R

$\frac{d}{dx}\sin x = \cos x$

$\frac{d}{dx}\cos x = -\sin x$

$\frac{d}{dx}\tan x = \sec^2 x$

$\frac{d}{dx}\cot x = -\operatorname{cosec}^2 x$

$\frac{d}{dx}\sec x = \sec x \tan x$

$\frac{d}{dx}\operatorname{cosec} x = -\operatorname{cosec} x \cot x$

The following rules were also established as a part of the algebra of derivatives (Ref. Sec.-14.13, Vol.-I)

If u and v be two derivable functions of x, then

(i) $(u+v)' = u'+v'$

(ii) $(u-v)' = u' - v'$

(iii) $(uv)' = u'v+uv'$

(iv) $\left(\frac{u}{v}\right)' = \frac{u'v-uv'}{v^2}$

7.4 Derivatives of Exponential and Logarithmic Functions

Exponential Function

Example-10

If $y=a^x$, $a > 0$ then $\frac{dy}{dx} = a^x \, In \, a$.

Proof :

$$\frac{dy}{dx} = \lim_{\delta x \to 0} \frac{\delta y}{\delta x}$$

$$= \lim_{\delta x \to 0} \frac{a^{x+\delta x} - a^x}{\delta x}$$

$$= \lim_{\delta x \to 0} \frac{a^x(a^{\delta x} - 1)}{\delta x}$$

$$= a^x \lim_{\delta x \to 0} \frac{a^{\delta x} - 1}{\delta x}$$

$$= a^x \text{ in } a.$$

$\left(\because \lim_{x \to 0} \frac{a^x - 1}{x} = In \, a\right.$, see section-14.7, Example-23, Vol-I)

Corollary :

If $y = e^x$ then $\frac{dy}{dx} = e^x$

It follows from the above result on substituting a by e and observing that $ln \, e = 1$.

Example 11 :

If $y = \log_a x$

$(x > 0, a > 0 \; a \neq 1)$,

then $\frac{dy}{dx} = \frac{1}{x \ln a}$

Proof :

$$\frac{dy}{dx} = \lim_{\delta x \to 0} \frac{\delta y}{\delta x}$$

$$= \lim_{\delta x \to 0} \frac{\log_a(x+\delta x) - \log_a x}{\delta x}$$

$$= \lim_{\delta x \to 0} \frac{\log_a\left(\frac{x+\delta x}{x}\right)}{\delta x} \qquad (\because \log m - \log n = \log\frac{m}{n})$$

$$= \lim_{\delta x \to 0} \frac{1}{x}.\frac{x}{\delta x}.\log_a\left(1+\frac{\delta x}{x}\right)$$

$$= \frac{1}{x}\lim_{\frac{\delta x}{x} \to 0} \log_a\left(1+\frac{\delta x}{x}\right)^{x/\delta x}$$

$(\because \ n \log_a m = \log_a m^n$ and $\delta x \to 0 \Leftrightarrow \frac{\delta x}{x} \to 0)$

$$= \frac{1}{x}\log_a e \quad \left(\because \lim_{h\to 0}(1+h)^{1/h} = e\right)$$

$$= \frac{1}{xlna} \quad \left(\because \log_a e = \frac{1}{\log_e a}\right)$$

Corollary :

If $y = \ln x$ then $\frac{dy}{dx} = \frac{1}{x}$.

Example 12 :

Differentiate $e^{\sqrt{x}}$ from definition.

Solution :

Put $u=\sqrt{x}$ so that $y = e^u$.

Let δx, δy and δu be changes in x, y and u respectively.

Now, $u = \sqrt{x} \Rightarrow u + \delta u = \sqrt{x+\delta x}$.

Hence $y + \delta y = e^{\sqrt{x+\delta x}} = e^{u+\delta u}$

$$\therefore \frac{\delta y}{\delta x} = \frac{e^{u+\delta u} - e^u}{\delta x} = \frac{e^{u+\delta u} - e^u}{\delta u}.\frac{\delta u}{\delta x}$$

$$= e^u.\frac{e^{\delta u} - 1}{\delta u}.\frac{\delta u}{\delta x}$$

$$\therefore \frac{dy}{dx} = \lim_{\delta x \to 0}\frac{\delta y}{\delta x} = e^u.\lim_{\delta u \to 0}\frac{e^{\delta u}-1}{\delta u}.\lim_{\delta x \to 0}\frac{\delta u}{\delta x} \qquad \dots \quad \dots \quad (A)$$

We observe that as $\delta x \to 0$, δu also tends to 0.

Hence taking limit in (A),

$$\frac{dy}{dx} = e^u.\text{ In e.}\lim_{\delta x\to 0}\frac{\left(\sqrt{x+\delta x}-\sqrt{x}\right)}{\delta x}$$

$$= e^u.\lim_{\delta x\to 0}\frac{\left(\sqrt{x+\delta x}-\sqrt{x}\right)\left(\sqrt{x+\delta x}+\sqrt{x}\right)}{\delta x\left(\sqrt{x+\delta x}+\sqrt{x}\right)}$$

$$= e^u.\lim_{\delta x\to 0}\frac{1}{\sqrt{x+\delta x}+\sqrt{x}}$$

$$= e^u.\frac{1}{2\sqrt{x}} = \frac{e^{\sqrt{x}}}{2\sqrt{x}}.$$

EXERCISE 7 (b)

1. **Differentiate from definition.**

(i) e^{3x} (ii) 2^{x^2}

(iii) $ln\,(3x+1)$ (iv) $\log x^5$ (Hint : $\log x^5 = \frac{ln 5}{ln x}$

(v) $ln \sin x$ (vi) $x^2 a^{2x}$

7.5 Derivative of a Composite Function (The Chain Rule) :

Let $y = f(u)$ be a differentiable function of u and $u = g(x)$ be a differentiable function of x so that $y = \text{fog}(x)$ is a composite function of x. For example if $y = \cos u$ and $u = x^2$ then $y = \cos x^2$ is a composite function of x. The derivative of y w.r.t. x is not $-\sin x^2$ as one might think considering that derivative of $\cos u$ is $-\sin u$. It is very important to observe here that the derivative of y i.e. $\cos u$ is $-\sin u$ **w.r.t. u** but we require the derivative w.r.t.x.

Now let δx be a small increment in x and δu, δy be corresponding increments in u and y respectively.

Since $u = g(x)$ is a differentiable function of x it is continuous (Ref. 7.11) and hence $\delta u \to 0$ as $\delta x \to 0$.

[This has been proved later while discussing 'differentiability and continuity'. However you can read that article and understand the contents right now !]

Now $\frac{\delta y}{\delta x} = \frac{\delta y}{\delta u}\cdot\frac{\delta u}{\delta x}$ $(\delta u \neq 0)$

$$\Rightarrow \frac{dy}{dx} = \lim_{\delta x\to 0}\frac{\delta y}{\delta x} = \lim_{\delta u\to 0}\frac{\delta y}{\delta u}\cdot\lim_{\delta x\to 0}\frac{\delta u}{\delta x}$$

$$= \frac{dy}{du}\cdot\frac{du}{dx}.$$

$\left[\ \delta u = 0 \Rightarrow u \text{ is constant} \Rightarrow y \text{ is constant} \Rightarrow \frac{dy}{dx} = 0\ \right]$

The result

$$\frac{dy}{dx} = \frac{dy}{du}\,\frac{du}{dx}$$

is called the **chain rule** of differentiation. It can be generalized as follows :

If y is a differentiable function of u, u is a differentiable function of v, ... and finally t is a differentiable function of x, then the derivative of y w.r.t. x is given by the rule

$$\frac{dy}{dx} = \frac{dy}{du}\,\frac{du}{dv}\,\cdots\,\frac{dt}{dx}.$$

It is important to learn to detect the cases where the chain rule is to be applied.

Example 13 :

Find $\frac{dy}{dx}$ when

(i) $y = (x^2 + 2x - 1)^5$ (ii) $y = \cot^3 x$, (iii) $y = \cos (lnx)^2$.

Solution :

(i) Let $u = x^2 + 2x - 1$.

Then $y = u^5$, $\therefore \frac{dy}{du} = \frac{du^5}{du} = 5u^4$

and $\frac{du}{dx} = \frac{d}{dx}(x^2 + 2x - 1) = 2x + 2$.

Hence $\frac{dy}{dx} = \frac{dy}{du}\,\frac{du}{dx} = 5u^4 (2x + 2) = 10 (x + 1) (x^2 + 2x - 1)^4$.

(ii) Let $u = \cot x$, Then $y = u^3$.

$\therefore \frac{dy}{du} = \frac{du^3}{du} = 3u^2$ and $\frac{du}{dx} = \frac{d}{dx} \cot x = -\operatorname{cosec}^2 x$

Hence $\frac{dy}{dx} = 3u^2 (-\operatorname{cosec}^2 x) = -3 \cot^2 x \operatorname{cosec}^2 x$.

(iii) Let $u = (ln\, x)^2$. Then $y = \cos u$.

Again let $\vartheta = lnx$. Then $u = \vartheta^2$.

Now y is a function of u and u is a function of ϑ.

Bt the chain rule,

$$\frac{dy}{dx} = \frac{dy}{du}\,\frac{du}{d\vartheta}\,\frac{d\vartheta}{dx},$$

$$\frac{dy}{du} = \frac{d(\cos u)}{du} = -\sin u,$$

$$\frac{du}{d\vartheta} = \frac{d\vartheta^2}{d\vartheta} = 2\vartheta,$$

$$\frac{d\vartheta}{dx} = \frac{d}{dx} lnx = \frac{1}{x}.$$

On substitution in (2) we get

$$\frac{dy}{dx} = -2\sin (lnx)^2\, lnx/x.$$

We can use the chain rule to obtain the following general results where a and b are constants.

$$\frac{d}{dx}(ax + b)^{\alpha} = \alpha\,(ax + b)^{\alpha - 1} a,\ \alpha \in \mathrm{R}$$

$$\frac{d}{dx}\sin(ax+b) = a\cos(ax+b),$$

$$\frac{d}{dx}\ln(ax+b) = a/(ax+b),$$

$$\frac{d}{dx}a^{bx} = b\,a^{bx}\ln a,$$

$$\frac{d}{dx}e^{ax} = a\,e^{ax}.$$

More generally if u is a function of x we have the following formulae.

$$\frac{du^{\alpha}}{dx} = \alpha u^{\alpha-1}\frac{du}{dx},\ \alpha \in R$$

$$\frac{d}{dx}\sin u = \cos u\frac{du}{dx},\ \frac{d}{dx}\cos u = -\sin u\frac{du}{dx}.$$

$$\frac{d}{dx}\tan u = \sec^2 u\frac{du}{dx}$$ (Similarly for other trigonometric functions)

$$\frac{d}{dx}\ln u = \frac{1}{u}\frac{du}{dx}$$

$$\frac{d}{dx}a^{u} = a^{u}\ln a\frac{du}{dx}$$

$$\frac{d}{dx}e^{u} = e^{u}\frac{du}{dx}.$$

For practice, the students may begin by actually making substitutions for applying the chain rule as in the Example 13. At a later stage the substitutions need not be explicitly stated. Instead, one can make use of the generalized formulae to shorten the procedure. For instance, in case of example 13 (iii), we proceed briefly as follows.

$$\frac{d}{dx}\cos(\ln x)^2 = -\sin(\ln x)^2\frac{d}{dx}(\ln x)^2$$

$$= -\sin(\ln x)^2\, 2(\ln x)\frac{d}{dx}(\ln x)$$

$$= -\sin(\ln x)^2\, 2\ln x\,.\,\frac{1}{x}.$$

EXERCISE 7 (c)

Find derivatives of the following functions.

1. $(x^2+5)^8$
2. $\dfrac{1}{(x^3+\sin x)^2}$
3. $ln\left(\sqrt{x}+1\right)$
4. $\sin 5x + \cos 7x$
5. $e^{\sin t}$
6. $\sqrt{ax^2+bx+c}$
7. $\left(\dfrac{x+1}{x^2+3}\right)^{-3}$
8. $\sec(\tan\theta)$
9. $\sin\left(\dfrac{1-x^2}{1+x^2}\right)$
10. $\sqrt{\tan(3z)}$

11. $\tan^3 x$
12. $\sin^4 x$
13. $\sin^2 x \cos^2 x$
14. $\sin 5x \cos 7x$
15. $\tan x \cot 2x$
16. $\sqrt{\sin \sqrt{x}}$
17. $\sqrt{\sec (2x + 1)}$
18. $\operatorname{cosec} (ax + b)^2$
19. $a^{\ln x}$
20. $a^{x^2} b^{x^3}$
21. $\ln \tan x$
22. $5^{\sin x^2}$
23. $\ln \tan \left(\frac{\pi}{4} + \frac{x}{2}\right)$
24. $\sqrt{\left(a^{\sqrt{x}}\right)}$
25. $\ln \left(e^{nx} + e^{-nx}\right)$
26. $e^{\sqrt{ax}}$
27. $\sqrt{\log x}$
28. $e^{\sin x} - a^{\cos x}$
29. $\dfrac{e^{3x^2}}{\ln \sin x}$
30. Prove that $\dfrac{d}{dx}\left[\dfrac{1 - \tan x}{1 + \tan x}\right]^{\frac{1}{2}} = -1 / \sqrt{\cos 2x}\ (\cos x + \sin x)$

7.6 Derivatives of Inverse Functions :

Theorem 5 :

Let f be a differentiable function of x which admits of an inverse function f^{-1}.

Then $\dfrac{df^{-1}}{dy} = \dfrac{1}{\left(\dfrac{df}{dx}\right)}$ provided $\dfrac{df}{dx} \neq 0$,

Or equivalently $\dfrac{dx}{dy} = \dfrac{1}{\left(\dfrac{dy}{dx}\right)}$, provided $\dfrac{dy}{dx} \neq 0$. .... (1)

Proof :

Let $y = f(x)$.. Let δx be a small increment in x and δy the corresponding increment in y.

Hence $y + \delta y = f(x + \delta x)$

$$\Leftrightarrow x + \delta x = f^{-1}(y + \delta y)$$

$$\Rightarrow \frac{d}{dy} f^{-1}(y) = \lim_{\delta y \to 0} \frac{f^{-1}(y + \delta y) - f^{-1}(y)}{\delta y}$$

$$= \lim_{\delta x \to 0} \frac{(x + \delta x) - x}{f(x + \delta x) - f(x)}$$

$$= \lim_{\delta x \to 0} \frac{\delta x}{f(x + \delta x) - f(x)}$$

$$= \lim_{\delta x \to 0} \frac{1}{\dfrac{f(x + \delta x) - f(x)}{\delta x}}$$

$$= \frac{1}{\left(\dfrac{df}{dx}\right)}, \text{ proved } \frac{df}{dx} \neq 0.$$

Example 14 :

Find $\frac{dy}{dx}$ if $x = \frac{\sqrt{y}}{\sqrt{y} + 1}$.

Here $x = \frac{\sqrt{y}}{\sqrt{y} + 1} = 1 - \frac{1}{\sqrt{y} + 1}$

$$\therefore \frac{dx}{dy} = -\frac{d}{dy}\left(\sqrt{y} + 1\right)^{-1}$$

$$= -(-1)\left(\sqrt{y} + 1\right)^{-2} \frac{d}{dy}\left(\sqrt{y} + 1\right) \qquad \text{(by the chain rule)}$$

$$= \frac{y^{-\frac{1}{2}}}{2\left(\sqrt{y} + 1\right)^2} = \frac{1}{2\sqrt{y}\left(\sqrt{y} + 1\right)^2}$$

$$\Rightarrow \frac{dy}{dx} = 2\sqrt{y}\left(\sqrt{y} + 1\right)^2.$$

Derivative of inverse Trigonometric Functions :

We recollect that $\sin : \mathrm{R} \to [-1, 1]$ is a trigonometric function and by restricting the domain to $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, we obtain a bijective function which possesses an inverse function.

$$\sin^{-1} : [-1, 1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

Thus $y = \sin^{-1}x$, $x \in [-1, 1]$

$\Leftrightarrow x = \sin y$, $y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

Similarly the other inverse functions viz. $\cos^{-1}$, $\tan^{-1}$, $\cot^{-1}$, cosec^{-1}, are defined. Their derivatives can be obtained by using (1).. Two examples are given below. Derivatives of rest of the inverse trigonometric functions can be obtained in a similar way.

Example 15 :

Find derivative of

(i) $\sin^{-1}x, -1 < x < 1$

(ii) $\sec^{-1}x, |x| > 1$.

Solution :

(i) Let $y = \sin^{-1}x,\ x \in (-1, 1)$

$\Leftrightarrow x = \sin y$

Differentiating w.r.t. y we obtain

$$\frac{dx}{dy} = \cos y$$

$$\Rightarrow \frac{dy}{dx} = \frac{1}{\cos y} = \frac{1}{\pm\sqrt{1 - \sin^2 y}} = \frac{1}{\pm\sqrt{1 - x^2}}.$$

Since by definition of inverse function $y \in \left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right)$ whenever $x \in (-1, 1)$, $\cos y > 0$.

So $\dfrac{dy}{dx} = \dfrac{1}{\sqrt{1-x^2}}$. .... (2)

(ii) Let $y = \sec^{-1}x, |x|>1$, i.e. $x \in (-\infty, -1) \cup (1, \infty)$

$$\Leftrightarrow x = \sec y, y \in \left(0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right)$$

Differentiating w.r.t. y we obtain

$$\frac{dx}{dy} = \sec y \tan y$$

$$\Rightarrow \quad \frac{dy}{dx} = \frac{1}{\sec y \tan y}$$

$$= \frac{1}{\sec y\left(\pm \sqrt{\sec^2 y - 1}\right)}$$

$$= \frac{1}{\pm x\sqrt{x^2 - 1}}$$

Where the sign before the radical depends on the sign of $\tan y$. Now if $x \in (-\infty, -1)$ then $\sec y = x < -1$ which implies that $y \in \left(\dfrac{\pi}{2}, \pi\right)$. But then $\tan y < 0$. Similarly if $x \in (1, \infty)$, then $y \in \left(0, \dfrac{\pi}{2}\right)$ where $\tan y > 0$.

$$\text{Thus } \frac{dy}{dx} = \begin{cases} \dfrac{1}{-x\sqrt{x^2 - 1}}, & \text{if } x < -1 \\ \dfrac{1}{x\sqrt{x^2 - 1}}, & \text{if } x > 1. \end{cases}$$

Combining the two cases we may write

$$\frac{dy}{dx} = \frac{d}{dx}\sec^{-1}x = \frac{1}{|x|\sqrt{x^2 - 1}}. \qquad \text{........ (3)}$$

Similarly we can find

$$\frac{d}{dx}\cos^{-1}x = \frac{-1}{\sqrt{1 - x^2}}, \ -1 < x < 1 \qquad \text{......... (4)}$$

$$\frac{d}{dx}\tan^{-1}x = \frac{1}{1 + x^2}, \ x \in \mathrm{R} \qquad \text{......... (5)}$$

$$\frac{d}{dx}\cot^{-1}x = \frac{-1}{1 + x^2}, \ x \in \mathrm{R} \qquad \text{......... (6)}$$

$$\frac{d}{dx}\mathrm{cosec}^{-1}x = \frac{-1}{|x|\sqrt{x^2 - 1}}, \ |x| > 1. \qquad \text{......... (7)}$$

Example 16 :

Find the derivative of $\tan^{-1}x$ from definition

Let $y = \tan^{-1}x,\ x \in \mathbb{R}$

$$\Leftrightarrow x = \tan y,\ y \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Let δx be a small increment in x and δy the corresponding increment in y. Then $x + \delta x = \tan(y + \delta y)$

$\therefore \delta x = \tan(y + \delta y) - \tan y$

$$\Rightarrow \frac{\delta y}{\delta x} = \frac{\delta y}{\tan(y + \delta y) - \tan y}$$

$$= \frac{\delta y \cos(y + \delta y)\cos y}{\sin(y + \delta y)\cos y - \sin y \cos(y + \delta y)}$$

$$= \frac{\delta y \cos(y + \delta y)\cos y}{\sin(y + \delta y - y)}$$

$$= \frac{\delta y}{\sin\delta y} . \cos(y + \delta y)\cos y.$$

$$\Rightarrow \frac{dy}{dx} = \lim_{\delta x \to 0} \frac{\delta y}{\delta x} = \lim_{\delta y \to 0} \frac{1}{\frac{\sin \delta y}{\delta y}} \cdot \lim_{\delta y \to 0} \cos(y + \delta y)\cos y.$$

$$= 1 . \cos^2 y = \frac{1}{\sec^2 y} = \frac{1}{1 + x^2}.$$

Example 17 :

Find derivative of $\tan^{-1}(\sin^2 x)$.

Solution :

Let $y = \tan^{-1}(\sin^2 x) = \tan^{-1}u$, where $u = \sin^2 x$.

Then $\frac{dy}{du} = \frac{d}{du}\tan^{-1}u = \frac{1}{1 + u^2}$,

$$\frac{du}{dx} = \frac{d}{dx}\sin^2 x = 2\sin x \frac{d}{dx}\sin x$$

$$= 2\sin x \cos x = \sin 2x.$$

Hence $\frac{dy}{dx} = \frac{dy}{du}\ \frac{du}{dx} = \frac{1}{1 + u^2}\sin 2x = \frac{\sin 2x}{1 + \sin^4 x}$.

Example 18 :

Differentiate $\tan^{-1}\left(\frac{1}{1 - x + x^2}\right)$.

Solution :

We have

$$\frac{d}{dx}\tan^{-1}\left(\frac{x+y}{1-xy}\right)$$

$$= \frac{d}{dx}\tan^{-1}x + \frac{d}{dx}\tan^{-1}y.$$

Hence writing $1-x$ for y, we get

$$\frac{d}{dx}\tan^{-1}\left(\frac{1}{1-x+x^2}\right)$$

$$= \frac{d}{dx}\tan^{-1}x + \frac{d}{dx}\tan^{-1}(1-x)$$

$$= \frac{1}{1+x^2} - \frac{1}{1+(1-x)^2}.$$

Otherwise, $\frac{d}{dx}\tan^{-1}\left(\frac{1}{1-x+x^2}\right)$

$$= \frac{(1-x+x^2)^2}{1+(1-x+x^2)^2}\ \frac{d}{dx}\left(\frac{1}{1-x+x^2}\right)$$

$$= \frac{(1-x+x^2)^2}{(1-x+x^2)^2+1}\cdot\frac{(-1)}{\left(1-x+x^2\right)^2}.(-1+2x)$$

$$= \frac{1-2x}{1+(1-x+x^2)^2}$$

which is equal to the earlier result.

EXERCISE 7(d)

1. Prove the formulas (4) to (7)

Find derivatives of the following functions.

2. $\sin^{-1}2x$
3. $\cot^{-1}\sqrt{x}$
4. $\sec^{-1}(2x+1)$
5. $\cos^{-1}\sqrt{\dfrac{1+x}{2}}$
6. $\cos^{-1}\left(\dfrac{x-\frac{1}{x}}{x+\frac{1}{x}}\right)$
7. $\tan^{-1}(\cos\sqrt{x})$
8. $x^2\,\text{cosec}^{-1}\left(\dfrac{1}{\ln x}\right)$
9. $\cot^{-1}\dfrac{\sqrt{1-x^2}}{x}$
10. $(x\sin^{-1}x)^{15}$
11. $\sin^{-1}\sqrt{\dfrac{1-x}{1+x}}$.

7.7 Methods of Differentiation :

Now that we have learnt how to find derivatives of standard functions and some rules for differentiating sum, product, quotient etc. of functions we are in a position to differentiate any given function (provided, of course, it is differentiable). But again, to find the derivative in the simplest possible way we have to adopt various methods depending on the nature of the given function and this we illustrate below through examples.

Differentiation By Substitution :

Sometimes with proper substitution we can transform the given function to a simpler function in the new variable so that the process of differentiation w.r.t. the new variable becomes easier. Finally we apply the chain rule to obtain the derivative w.r.t. the original variable.

Example 19 :

Differentiate $\cos^{-1}(4x^3 - 3x)$.

Solution :

Let $y = \cos^{-1}(4x^3 - 3x)$.

Let $x = \cos\theta$ so that $\theta = \cos^{-1}x$.

Then $y = \cos^{-1}(4\cos^3\theta - 3\cos\theta)$.

$= \cos^{-1}(\cos 3\theta)$.

$= 3\theta = 3\cos^{-1}x$.

Hence $\dfrac{dy}{dx} = \dfrac{-3}{\sqrt{1 - x^2}}$.

Example 20 :

Differentiate $\tan^{-1}\left(\dfrac{\sqrt{x} - x}{1 + x^{3/2}}\right)$.

Solution :

Let $y = \tan^{-1}\left(\dfrac{\sqrt{x} - x}{1 + x^{3/2}}\right)$.

Put $\sqrt{x} = \tan\alpha$, $x = \tan\beta$

so that $\alpha = \tan^{-1}\sqrt{x}$ and $\beta = \tan^{-1}x$ .... (1)

Then $\left(\dfrac{\sqrt{x} - x}{1 + x^{3/2}}\right) = \dfrac{\tan\alpha - \tan\beta}{1 + \tan\alpha\tan\beta} = \tan(\alpha - \beta)$

$\Rightarrow y = \tan^{-1}\tan[\alpha - \beta]$

$= \alpha - \beta$

$= \tan^{-1}\sqrt{x} - \tan^{-1}x$, from (1)

Hence $\dfrac{dy}{dx} = \dfrac{d}{dx}\tan^{-1}\sqrt{x} - \dfrac{d}{dx}\tan^{-1}x$

$= \dfrac{1}{1 + \left(\sqrt{x}\right)^2}\dfrac{d\left(\sqrt{x}\right)}{dx} - \dfrac{1}{1 + x^2}$

$= \dfrac{1}{2\sqrt{x}(1 + x)} - \dfrac{1}{1 + x^2}$.

Example 21 :

Find $\frac{d\theta}{dt}$ if $\tan\theta = \frac{1-t}{1+t}$.

Solution :

Let $t = \tan\alpha$ so that $\alpha = \tan^{-1}t$.

Hence $\tan\theta = \frac{1-\tan\alpha}{1+\tan\alpha} = \tan\left(\frac{\pi}{4} - \alpha\right)$

$$\Rightarrow \sec^2\theta\frac{d\theta}{dt} = \sec^2\left(\frac{\pi}{4} - \alpha\right)(-)\frac{d\alpha}{dt}.$$

$$\Rightarrow \frac{d\theta}{dt} = -\frac{d\alpha}{dt} = -\frac{d}{dt}\tan^{-1}t = -\frac{1}{1+t^2}. \quad \left(\textit{Since } \sec^2\theta = \sec^2\left(\frac{\pi}{4} - \alpha\right)\right)$$

EXERCISE 7(e)

Differentiate the following functions by proper substitution.

1. $\sin^{-1} 2x\sqrt{1-x^2}$
2. $\tan^{-1}\frac{2x}{1-x^2}$
3. $\tan^{-1}\sqrt{\frac{1-t}{1+t}}$
4. $\left[\left(\frac{1+t^2}{1-t^2}\right)^2 - 1\right]^{\frac{1}{2}}$
5. $\tan^{-1}\left(\frac{\sqrt{x}+\sqrt{a}}{1-\sqrt{xa}}\right)$
6. $\sin^{-1}\left(\frac{2x}{1+x^2}\right)$
7. $\sec^{-1}\left(\frac{\sqrt{a^2+x^2}}{a}\right)$
8. $\sin^{-1}\left(\frac{2\sqrt{t^2-1}}{t^2}\right)$
9. $\cos^{-1}\left(\frac{1-t^2}{1+t^2}\right)$
10. $\cos^{-1}(2t^2-1)$.

Differentiation using Logarithms :

When a function appears as an exponent of another function we make use of logarithms.

Examples 22 :

Differentiate $(\sin x)^{\tan x}$.

Solution :

Let $y = (\sin x)^{\tan x}$.

Taking logarithms of both sides we have

$$\ln y = \ln(\sin x)^{\tan x}$$
$$= \tan x \ln \sin x.$$

Differentiating both sides w.r.t. x we obtain

$$\frac{1}{y}\frac{dy}{dx} = \frac{d}{dx}\tan x \, ln \sin x + \tan x \frac{d}{dx} ln \sin x$$

$$= \sec^2 x \, ln \sin x + \tan x . \frac{\cos x}{\sin x}. \text{ (by chain rule)}$$

$$= \sec^2 x \, ln \sin x + 1$$

$$\Rightarrow \frac{dy}{dx} = y(\sec^2 x \, ln \sin x + 1) = (\sin x)^{\tan x}(\sec^2 x \, ln \sin x + 1).$$

(ii) When a given function is expressed as a product of several functions we use logarithmic differentiation.

Example 23 :

Differentiate

$$y = \frac{(x-1)^2\sqrt{3x-1}}{x^7(6-7x^2)^{3/2}}$$

Solution :

Taking logarithms of both sides we get

$$ln\, y = ln\,(x-1)^2 + ln\,(3x-1)^{1/2} - ln\, x^7 - ln\,(6-7x^2)^{3/2}$$

$$= 2ln\,(x-1) + \frac{1}{2}\, ln\,(3x-1) - 7\ln x - \frac{3}{2}\, ln\,(6-7x^2).$$

Now differentiating both sides w.r.t. x and using the chain rule for differentiation on the r.h.s. we have,

$$\frac{1}{y}\frac{dy}{dx} = 2.\frac{1}{x-1}\frac{d}{dx}(x-1) + \frac{1}{2}\frac{1}{3x-1}\frac{d}{dx}(3x-1) - 7.\frac{1}{x} - \frac{3}{2}.\frac{1}{6-7x^2}.\frac{d}{dx}(6-7x^2)$$

$$= \frac{2}{x-1} + \frac{3}{2(3x-1)} - \frac{7}{x} + \frac{21x}{6-7x^2}$$

$$\Rightarrow \frac{dy}{dx} = y\left[\frac{2}{x-1} + \frac{3}{2(3x-1)} - \frac{7}{x} + \frac{21x}{6-7x^2}\right]$$

$$= \frac{(x-1)^2(3x-1)^{\frac{1}{2}}}{x^7(6-7x^2)^{\frac{3}{2}}}\left[\frac{2}{x-1} + \frac{3}{2(3x-1)} - \frac{7}{x} + \frac{21x}{6-7x^2}\right].$$

EXERCISE 7(f)

Find derivatives of the following functions in their natural domains :

1. x^x
2. $\left(1+\frac{1}{x}\right)^x$
3. $x^{\sin x}$
4. $(\log x)^{\tan x}$
5. $2^{(2^x)}$
6. $(1+\sqrt{x})^{x^2}$
7. $(\sin^{-1}x)^{\sqrt{1-x^2}}$
8. $(\tan x)^{\log x^3}$

9. $x^{1/x} + (\sin x)^x$

10. $(\cos x)^x + x^{\cos x}$

11. $(x^2+1)^{2/3}(3x+1)^{1/4}\sqrt{x}$

12. $\dfrac{(x+1)(x+2)^2(x+3)^3}{(x-1)(x-2)^2(x-3)^3}$

13. $(\sin x)^x\sqrt{\sin x}\,(1+x^2)^{\frac{1}{2}+x}$

14. $(\sec x + \tan x)^{\cot x}$

15. $\left(2^{\sqrt{x}}\right)^{1+\sqrt{x}}$

N.B. When we write logx without mentioning the base in cases as above it is to be understood that the base is e and logx actually means lnx.

7.8 Differentiation of Implicit Functions :

Let f be a real function. It we take x as an independent variable and y as a dependent variable in the equation $y = f(x)$ (e.g. $y = \sin x$, $y = x^2 + 1$ etc.) we observe that the function is given by an equation in which y is expressed explicitly in terms of x. But in case of some equations (e.g. $x^3y^3 + x + y - 1 = 0$) the dependent variable y may not be expressible explicitly in terms of the independent variable x. Even if we can separate the two variables it may not uniquely determine y for a given value of x. For example consider the equation $x^2 + y^2 - a^2 = 0$ which implies a relation and not a function. We have two distinct functions $y = +\sqrt{a^2 - x^2}$ and $y = -\sqrt{a^2 - x^2}$ determined from it. These two functions are said to be implicity defined by the equation $x^2 + y^2 - a^2 = 0$.

In contrast, the equation $2x + 3y - 1 = 0$ determines a unique value of y in terms of x such as $y = \frac{1}{3}(1 - 2x)$. So the equation $2x + 3y - 1 = 0$ implicity defines the function given by $y = \frac{1}{3}(1 - 2x)$. Thus an equation $F(x, y) = 0$ in two variables in which x and y are the independent and the dependent variables respectively may determine one or more functions. Any such function as determined from $F(x, y) = 0$ is known as an implicit function. It may be noted however that every equation of the type $F(x, y) = 0$ does not determine y as an implicit function of x. The study of the conditions under which $F(x, y) = 0$ determines an implicit function is beyond the scope of the book. For our purpose, we assume that every equation we consider in two varibles x and y does determine y as an implicit function of x. Our aim is to find dy/dx in such cases. This process is known as implicit differentiation.

The advantage of this process is that it gives the derivative of every differentiable function that is implicity defined by the given equation and we do not have to express y explicity in terms of x for each implicit function. This is illustrated in the following example.

Example 24 :

Find $\frac{dy}{dx}$ if $x^2 + y^2 - a^2 = 0$.

Solution :

Given $x^2 + y^2 - a^2 = 0$ .... (1)

Differentiating both sides w.r.t. x we have

$$2x + 2y \frac{dy}{dx} = 0 \Rightarrow \frac{dy}{dx} = -\frac{x}{y}. \qquad \text{......... (2)}$$

We note that the equation (1) implicitly defines two functions.

(i) $y_1 = \sqrt{a^2 - x^2}$

and (ii) $y_2 = -\sqrt{a^2 - x^2}$

Their derivatives are given by

$$\frac{dy_1}{dx} = \frac{1}{2}(a^2 - x^2)^{-\frac{1}{2}}(-2x) = -\frac{x}{y_1}$$

and $$\frac{dy_2}{dx} = -\frac{1}{2}(a^2 - x^2)^{-\frac{1}{2}}(-2x) = -\frac{x}{y_2}$$

which are already contained in (2).

Example 25 :

Find $\frac{dy}{dx}$ if $y^3 + 3x^2y - 2x = 10$.

Solution :

Differentiating both sides w.r.t. x we have

$$\frac{dy^3}{dx} + \frac{d}{dx}(3x^2y) - \frac{d}{dx}(2x) = 0$$

$$\Rightarrow \quad 3y^2 \frac{dy}{dx} + 3\left(2xy + x^2 \frac{dy}{dx}\right) - 2 = 0$$

$$\Rightarrow \quad (3y^2 + 3x^2)\frac{dy}{dx} + 6xy - 2 = 0$$

$$\Rightarrow \quad \frac{dy}{dx} = \frac{2(1 - 3xy)}{3(x^2 + y^2)}$$

We note that in case of implicit differentiation the derivative usually contains both the variables x and y.

Interchanging the role of x and y we may regard x as a dependent variable and y as an independent variable and the above procedure will yield dx/dy without expressing x explicity in terms of y.

EXERCISE 7(g)

Find $\frac{dy}{dx}$.

1. $xy^2 + x^2y + 1 = 0$
2. $x^{\frac{1}{2}}y^{-\frac{1}{2}} + x^{\frac{3}{2}}y^{-\frac{3}{2}} = 0.$
3. $x^2 + 3y^2 = 5$
4. $y^2 \cot x = x^2 \cot y$
5. $y = \tan xy$
6. $x = y\, ln\,(xy)$
7. $e^{xy} + y \sin x = 1$
8. $ln\sqrt{x^2 + y^2} = \tan^{-1}\frac{y}{x}$

9. $y^x = x^{\sin y}$

10. If $\sin(x + y) = y\cos(x + y)$ then prove that $\dfrac{dy}{dx} = -\dfrac{1 + y^2}{y^2}$.

11. If $\sqrt{1 - x^4} + \sqrt{1 - y^4} = k(x^2 - y^2)$ then show that $\dfrac{dy}{dx} = \dfrac{x\sqrt{1 - y^4}}{y\sqrt{1 - x^4}}$.

(**Hints :** Put $x^2 = \cos\theta$, $y^2 = \cos\phi$)

7.9 Differentiation of Parametric Functions :

For a real–valued function f the set of points given by the ordered pairs (x, y) where $y = f(x)$, determines a curve in R^2. Sometimes the variables x and y are given as functions of another single variable, say t. For example, any point (x, y) on the circle $x^2 + y^2 = r^2$ can he given by

$$x = r\cos t,\ y = r\sin t. \quad \text{......... (1)}$$

The variable quantity t is called **parameter.** [The term 'parameter' is also used to mean a quantity which is invariable for a given curve but changes when we move from one curve of a given type to another. For example, the quantity m in the equation $y - k = m(x - h)$ of a straight line passing through (h, k) which varies as we move from one straight line to another all passing through the same given point (h, k)].

The equations (1) are called the parametric equations of the circle $x^2 + y^2 = r^2$

Suppose in general

$$x = \phi(t) \text{ and } y = \Psi(t).$$

where ϕ and Ψ are two differentiable functions of t. By eliminating t between these equations we may obtain a relation of the type $y = f(x)$ from which we may obtain $\dfrac{dy}{dx}$. But the process of elimination of the parameter t is not always convenient. So we apply the chain rule to obtain $\dfrac{dy}{dx}$ as follows :

$$\frac{dy}{dx} = \frac{dy}{dt}\,\frac{dt}{dx} = \frac{dy}{dt} \Big/ \frac{dx}{dt} = \frac{\psi'(t)}{\phi'(t)}. \quad \text{......... (2)}$$

Example 26 :

Find $\dfrac{dy}{dx}$ if $x = a(\cos t + t\sin t)$

and $y = a(\sin t - t\cos t)$.

$\dfrac{dx}{dt} = a(-\sin t + \sin t + t\cos t) = at\cos t$

and $\dfrac{dy}{dt} = a\left(\cos t - \cos t - t(-\sin t)\right) = at\sin t$

Hence $\frac{dy}{dx} = \frac{dy}{dt} \Big/ \frac{dx}{dt} = \frac{at\sin t}{at\cos t} = \tan t$.

EXERCISE–7(h)

Find $\frac{dy}{dx}$.

1. $x = a\cos\theta, y = a\sin\theta$
2. $x = at^2,\ y = 2at$ at $t = \frac{1}{2}$
3. $x = a\cos^3 t, y = a\sin^3 t$ at $t = \frac{\pi}{4}$
4. $\sin x = \frac{2t}{1+t^2}, \tan y = \frac{2t}{1-t^2}$
5. $x = 3\cos t - 2\cos^3 t,\ y = 3\sin t - 2\sin^3 t$.

7.10 Differentiation with respect to a function :

Suppose we have two differentiable functions given by $y = f(x)$ and $z = g(x)$. To find derivative of y w.r.t. z we regard x as a parameter and find $f'(x) = \frac{dy}{dx}$ and $g'(x) = \frac{dz}{dx}$. Then as in 12.6

$$\frac{dy}{dz} = \frac{dy}{dx} \Big/ \frac{dz}{dx} = \frac{f'(x)}{g'(x)}.$$

Example 27 :

Differentiate $\tan^{-1}x$ w.r.t. $\cos^{-1}x$.

Solution :

Let $y = \tan^{-1}x$ and $z = \cos^{-1}x$.

We have to find $\frac{dy}{dz}$. Now

$$\frac{dy}{dx} = \frac{1}{1+x^2} \text{ and } \frac{dz}{dx} = \frac{-1}{\sqrt{1-x^2}}$$

$$\therefore \quad \frac{dy}{dz} = \frac{dy}{dx} \Big/ \frac{dz}{dx} = -\frac{\sqrt{1-x^2}}{1+x^2}.$$

EXERCISE 7(i)

Differentiate.

1. $\sqrt{x}$ w.r.t. x^2
2. Sin$x$ w.r.t. cot$x$
3. $\frac{1-\cos x}{1+\cos x}$ w.r.t. $\frac{1-\sin x}{1+\sin x}$
4. $\tan^{-1}x$ w.r.t. $\tan^{-1}\sqrt{1+x^2}$
5. $\sin^{-1}\left(\frac{2x}{1+x^2}\right)$ w.r.t. $\cos^{-1}\left(\frac{1-x^2}{1+x^2}\right)$.

7.11 Differentiability and Continuity

Let us now make a short break to investigate a very important relation between differentiability and continuity. In higher mathematics continuity of a function plays a vital role. The following theorem gives us a helping hand in this connection.

Theorem 6 :

If a function f is differentiable at a point $x = c$, then it is continuous at that point. But the converse is not necessarily true.

Proof :

Let a function f be differentiable at $x = c$. Then by definition,

$$\lim_{h \to 0} \frac{f(c+h) - f(c)}{h} \text{ exists finitely and is equal to } f'(c).$$

$$\text{Now } \lim_{h \to 0} [f(c+h) - f(c)] = \lim_{h \to 0} \left[\frac{f(c+h) - f(c)}{h} . h\right]$$

$$= \lim_{h \to 0} \frac{f(c+h) - f(c)}{h} . \lim_{h \to 0} h$$

$$= f'(c) . 0 = 0.$$

$$\Rightarrow \lim_{h \to 0} f(c+h) = f(c)$$

$$\Rightarrow \lim_{x \to c} f(x) = f(c). \quad \text{(Pulling } x = c + h\text{)}$$

So f is continuous at $x = c$.

But the converse of the above theorem is not necessarily true, i.e. continuity at a point does not imply differentiability at that point. This is shown in the following example.

Example 28 :

Show that $|x|$ is continuous at the origin but it is not differentiable there.

Solution :

Let $f(x) = |x|$.

We know that $|x| = \begin{cases} x, & x \geq 0 \\ -x, & x < 0 \end{cases}$

At $x = 0, f(0) = 0$.

$$\lim_{x \to 0+} |x| = \lim_{h \to 0} |h|, h > 0$$

$$= \lim_{h \to 0} h = 0.$$

$$\lim_{x \to 0-} |x| = \lim_{h \to 0} |-h|, \ h > 0$$

$$= \lim_{h \to 0} h = 0.$$

Thus $\lim_{x \to 0} |x| = 0 = f(0)$.

$\Rightarrow$ $|x|$ is continuous at $x = 0$.

Now $f'(0+) = \lim_{h \to 0} \frac{|0 + h| - 0}{h}, h > 0$

$= \lim_{h \to 0} \frac{h}{h} = 1.$

$f'(0-) = \lim_{h \to 0} \frac{|0 - h| - 0}{-h}, h > 0$

$= \lim_{h \to 0} \frac{-(-h)}{-h} = -1.$

Since $f'(0+) \neq f'(0-)$ it follows that $f(x) = |x|$ is not differentiable at $x = 0$.

The student may verify the following.

(i) $f(x) = [x]$ is not continuous at $x = m, m \in Z$. and hence not differentiable at these points.

(ii) $f(x) = |x - 2|$ is not differentiable at $x = 2$ although continuous at that point.

(iii) $f(x) = \begin{cases} x + \frac{1}{2}, & x \leq \frac{1}{2} \\ \frac{3}{2} - x, & x > \frac{1}{2} \end{cases}$

is not differentiable at $x = \frac{1}{2}$. What about continuity at $x = \frac{1}{2}$?

Thus we observe that if a function is differentiable in a domain it must be continuous there and if it is not continuous in some domain then it is definitely not derivable there. But if it is continuous in some domain, then only from this fact we cannot conclude about its derivability. It will not be out of place to mention here that there exist functions (we shall not search for them) which are continuous everywhere on R but differentiable nowhere.

EXERCISE 7(j)

Test differentiability and continuity of the following functions.

1. $\left|1 - \frac{1}{x}\right|$ at $x = 1$.
2. $x^2 |x|$ at $x = 0$.
3. $f(x) = \tan x$ at $x = \frac{\pi}{2}$
4. $f(x) = \cot x$ at $x = \frac{\pi}{2}$
5. $f(x) = |\sin x|$. at $x = \pi$..
6. $f(x) = \frac{x}{1 + |x|}$ at $x = 0$.
7. $f(x) = \begin{cases} x \sin \frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$ at $x = 0$.

8. $f(x) = \begin{cases} \dfrac{1 - e^{-x}}{x}, & x \neq 0 \\ 1, & x = 0 \end{cases}$ at $x = 0$.

7.12 Miscellaneous Examples :

Example 29 :

Differentiate $y = \dfrac{\sqrt{a + x} + \sqrt{a - x}}{\sqrt{a + x} - \sqrt{a - x}}$.

Solution :

By rationalization,

$$y = \frac{\sqrt{a + x} + \sqrt{a - x}}{\sqrt{a + x} - \sqrt{a - x}} \cdot \frac{\sqrt{a + x} + \sqrt{a - x}}{\sqrt{a + x} + \sqrt{a - x}}$$

$$= \frac{(a + x) + 2\sqrt{a + x}\,\sqrt{a - x} + (a - x)}{(a + x) - (a - x)}$$

$$= \frac{a + \sqrt{a^2 - x^2}}{x}.$$

$$\therefore \frac{dy}{dx} = \left[x \cdot \frac{1}{2} (a^2 - x^2)^{-\frac{1}{2}} (-2x) - a - \sqrt{a^2 - x^2} \right] / x^2$$

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

$$= -\frac{a^2 + a\sqrt{a^2 - x^2}}{x^2 \sqrt{(a^2 - x^2)}} \quad \text{on simplification.}$$

Example 30 :

Differentiate $\sin x^0$.

Solution :

Converting degree to radian

$$y = \sin x^0 = \sin\left(\frac{\pi x}{180}\right)$$

$$= \sin\theta \text{ (say)}$$

where θ is in radians and $\theta = \dfrac{\pi x}{180}$.

$$\therefore \frac{dy}{dx} = \cos\theta \frac{d\theta}{dx} = \cos\theta \cdot \frac{\pi}{180} = \frac{\pi}{180} \cos x^0.$$

Example 31 :

Find $\dfrac{dy}{dx}$ if $y = \sqrt{x + \sqrt{x + \sqrt{x}}}$.

Solution :

Here $y^2 = x + \sqrt{x + \sqrt{x}}$

$\Rightarrow \quad (y^2 - x)^2 = x + \sqrt{x}$

$\Rightarrow \quad y^4 - 2xy^2 + x^2 = x + \sqrt{x}$

Differentiating implicitly w.r.t. x we get

$$4y^3 \frac{dy}{dx} - \left(2y^2 + 2x \,.\, 2y \frac{dy}{dx}\right) + 2x = 1 + \frac{1}{2\sqrt{x}}$$

$$\Rightarrow \quad \frac{dy}{dx} = \frac{1 + \frac{1}{2\sqrt{x}} - 2x + 2y^2}{4y^3 - 4xy}$$

$$= \frac{2\sqrt{x} + 1 - 4x\sqrt{x} + 4\sqrt{x}y^2}{4y\,(y^2 - x)\,2\sqrt{x}}$$

$$= \frac{2\sqrt{x} + 1 - 4x\sqrt{x} + 4\sqrt{x}\left(x + \sqrt{x + \sqrt{x}}\right)}{8y\left(\sqrt{x + \sqrt{x}}\right)\sqrt{x}}$$

$$= \frac{2\sqrt{x} + 4\sqrt{x} \,.\, \sqrt{x + \sqrt{x}} + 1}{8\sqrt{x + \sqrt{x + \sqrt{x}}} \,.\, \sqrt{x + \sqrt{x}}\sqrt{x}}$$

$$= \frac{2\sqrt{x}\left(1 + 2\sqrt{x + \sqrt{x}}\right) + 1}{8\sqrt{x}\sqrt{x + \sqrt{x}}\sqrt{x + \sqrt{x + \sqrt{x}}}}.$$

Example 32 :

Find $f'(x)$ where

$$f(x) = \frac{|x - 1|}{x^2 + 1}, \qquad x \neq 1.$$

Solution :

$$\text{Here } f(x) = \begin{cases} \dfrac{1 - x}{x^2 + 1}, & x < 1 \\[2ex] \dfrac{x - 1}{x^2 + 1}, & x > 1 \end{cases}$$

So when $x < 1$, $f'(x) = \dfrac{d}{dx}\left(\dfrac{1-x}{x^2+1}\right)$

$$= \frac{(-1)\,(x^2+1) - (1-x)2x}{(x^2+1)^2}$$

$$= \frac{x^2 - 2x - 1}{(x^2 + 1)^2}$$

and similarly, when $x > 1$, $f'(x) = \dfrac{1 + 2x - x^2}{(x^2 + 1)^2}$.

Example 33 :

Find $\frac{dy}{dx}$ if $x^y = y^x$.

Solution :

Taking logarithm,

$y\, lnx = x\, lny.$

Differentiating implicitly w.r.t. x

$$\frac{dy}{dx}\, lnx + y \cdot \frac{1}{x} = 1 \cdot lny + x \cdot \frac{1}{y} \cdot \frac{dy}{dx}$$

$$\Rightarrow \quad \frac{dy}{dx}\left(lnx - \frac{x}{y}\right) = lny - \frac{y}{x}$$

$$\Rightarrow \quad \frac{dy}{dx} = \frac{lny - \frac{y}{x}}{lnx - \frac{x}{y}}.$$

EXERCISE 7(k)

1. State True (T) or False (F).

(i) There is no function whose derivative is $\log\pi$.

(ii) There is no function which is its own derivative.

(iii) A function f is not differentiable at $x = c \Rightarrow f$ is not continuous at $x = c$.

(iv) $[x^2]$ is differentiable on $(-1, 1)$

(v) $|x + 2|$ is not differentiable at $x = 2$.

(vi) Derivative of $e^{3\log x}$ w.r.t. x is $3x^2$.

(vii) The derivative of a non constant even function is always an odd function.

(viii) If f and g are not derivable at x_0 then $f + g$ is not derivable at x_0.

2. Fill in the gaps by using the correct answer.

(i) If u is a constant and ϑ is a variable then $\frac{du^{\vartheta}}{d\vartheta}$ = - $(u^{\vartheta}ln\vartheta,\ \vartheta u^{\vartheta-1},\ u^{\vartheta}lnu,\ u\vartheta^{u-1})$

(ii) If $t = e^a$ then $\frac{d}{dx}\, x^t$ = - $(tx^{t-1},\ x^t,\ x^t lna,\ t\,x^t)$

(iii) If $u = t^2$ and $\vartheta = \sin t^2$, then $\frac{d\vartheta}{du}$ = - $\left(\cos^2 t,\ \frac{\sin t}{t},\ \sec t^2,\ \cos t^2\right)$

(iv) The tangent to the curve $y = (1 + x^2)^2$ at $x = -1$ has slope -- $(4, -4, 8, -8)$

(v) If $y = (gof)\ (x)$ then $\frac{dy}{dx}$ = - $\left(\frac{dg}{dx}\frac{dx}{df},\ \frac{dg}{df}\frac{df}{dx},\ \frac{df}{dx}\frac{dx}{dg},\ \frac{df}{dg}\frac{dg}{dx}\right)$.

(vi) If $y = \sec^{-1}\frac{\sqrt{x}+1}{\sqrt{x}} + \sin^{-1}\frac{\sqrt{x}}{\sqrt{x}+1}$ then $\frac{dy}{dx}$ = - $\left(0,\ \text{undefined},\ \frac{\pi}{2},\ 1\right)$

(vii) If $f(x) = \sqrt{x^2 - 2x + 1}$, $x \in [0, 2]$ then at $x = 1$, $f'(x)$ = - (1, 0, – 1, does not exist)

(viii) If $f(x) = [x^2]$ then $f'\left(\frac{3}{2}\right) =$ - (0, 2, 3, does not exist)

3. Differentiate from first principles

(i) e^{2x} (ii) $\sin^2 x$ (iii) $\cos x^2$

(iv) e^{x^2} (v) $\sqrt{\tan x}$ (vi) $x^2 \sin x$

(vii) $ln \sin x$ (viii) $\sin\sqrt{x}$ (ix) $\cos lnx$

(**Hints :** write $y = \cos u$ where $u = lnx$. Then find limit of $\delta y/\delta x = \frac{\delta y}{\delta u}(\delta u/\delta x)$.)

4. Test differentiability of the following functions at the indicated points.

(i) $f(x) = [x^2 + 1]$ at $x = -\frac{1}{2}$

(ii) $f(x) = \begin{cases} 1 - 2x, & x \le \frac{1}{2} \\ x - \frac{1}{2}, & x > \frac{1}{2} \end{cases}$ at $x = \frac{1}{2}$

(iii) $f(x) = x + |\cos x|$ at $x = \frac{\pi}{2}$

Hence onwards domain of a function is to be understood to be its natural domain unless stated otherwise.

5. Differentiate.

(i) $\frac{1}{\ln\left(x\sqrt{x+1}\right)}$ (ii) $\frac{\ln x}{e^x \sin x}$

(iii) $e^x(\tan x - \cot x)$ (iv) $\left(\sqrt{x} + \frac{1}{\sqrt{x}}\right) x \tan x$

(v) $\frac{\cos 3x - \cos x}{\cos 5x - \cos 3x}$ (vi) $x^2 e^x \operatorname{cosec} x$

(vii) $\frac{(x+1)\, lnx}{\sqrt{x+2}}$ (viii) $(x^3 - 1)^9 \sec^2 x$

(ix) $\sin^2(\cos^{-1}x)$ (x) $a^x\left(x + \frac{1}{x}\right)^{10}$

(xi) $ln\,\frac{\sqrt{x+4} - 2}{\sqrt{x+4} + 2}$ (xii) $ln\,\frac{4x^2(2x-7)^3}{\left(3x^2+7\right)^5}$

(xiii) $5^{\ln \sin x}$ (xiv) $\sqrt{\sin\sqrt{x}}$

(xv) $x^{\sin x} + (\tan x)^x$ (xvi) e^{e^x}

(xvii) $x^{\sqrt{x}}$ (xviii) $\sec^{-1}(e^x + x)$

(xix) $ln \cos e^x$ (xx) $a^{\sin^{-1}x^2}$

(xxi) $\cos^{-1}\left(\frac{x^4 - 1}{x^4 + 1}\right)$ (xxii) $(x^e)^{e^x} + (e^x)^{x^e}$

(xxiii) $x^{\left(x^{x}\right)}$ (xxiv) $\dfrac{(x+1)^2\sqrt{x-1}}{\left(x^2+3\right)^3 3^x}$

(xxv) $\left[5\,ln\,(x^3+1)-x^4\right]^{2/3}$ (xxvi) $\log_{10}\sin x+\log_x 10,\ 0<x<\pi$.

6. Differentiate

(i) $\sec^{-1}\left(\dfrac{x^2+1}{x^2-1}\right)$ (ii) $e^{\tan^{-1}x^2}$

(iii) $\dfrac{x\sin^{-1}x}{\sqrt{1+x^2}}$ (iv) $\tan^{-1}e^{2x}$

(v) $\tan^{-1}\dfrac{\cos x}{1+\sin x}$ (vi) $\tan^{-1}\left(\dfrac{\cos x-\sin x}{\cos x+\sin x}\right)$

(vii) $\tan^{-1}\dfrac{7ax}{a^2-12x^2}$ (viii) $\tan^{-1}\dfrac{\sqrt{1+x^2}+\sqrt{1-x^2}}{\sqrt{1+x^2}-\sqrt{1-x^2}}$ (put $x^2=\cos\theta$)

(ix) $x^2\cos^{-1}\dfrac{\sqrt{x}-1}{\sqrt{x}+1}+x^2\ \mathrm{cosec}^{-1}\dfrac{\sqrt{x}+1}{\sqrt{x}-1}$ (x) $\tan^{-1}\dfrac{x}{1+\sqrt{1-x^2}}$

(xi) $\tan^{-1}\left(\dfrac{x\sin\alpha}{1-x\cos\alpha}\right)$

7. Find $\dfrac{dy}{dx}$ if

(i) $x^3+y^3=12xy$ (ii) $\left(\dfrac{x}{a}\right)^{2/3}+\left(\dfrac{y}{b}\right)^{2/3}=1$

(iii) $x^y=c$ (iv) $y^x=c$

(v) $x\cot y+y\,\mathrm{cosec}x=0$ (vi) $y^2+x^2=ln\,(xy)+1$

(vii) $(\cos x)^y=\sin y$ (viii) $y^2=a^{\sqrt{x}}$

(ix) $x^m y^n=\left(\dfrac{x}{y}\right)^{m+n}$ (x) $y=x\cot^{-1}\left(\dfrac{x}{y}\right)$

(xi) $y=(\sin y)^{\sin 2x}$ (xii) $y^2=x^y$

(xiii) $(x+y)^{\cos x}=e^{x+y}$. (xiv) $x\tan y+y\tan x=0$

(xv) $\sqrt{x^2+y^2}=k\tan^{-1}\left(\dfrac{y}{x}\right)$.

8. Differentiate

(i) $\tan^{-1}\dfrac{2x}{1-x^2}$ w.r.t. $\sin^{-1}\dfrac{2x}{1+x^2}$

(ii) $\sec^{-1}\left(\dfrac{1}{2x^2-1}\right)$ w.r.t. $\sqrt{1-x^2}$

(iii) $\tan^{-1}\left(\dfrac{1+\sin x}{1-\sin x}\right)^{\frac{1}{2}}$ w.r.t. $\log\left(\dfrac{1+\cos x}{1-\cos x}\right)$.

9. Find $\frac{dy}{dx}$ when

(i) $x = a\left[\cos t + \log\tan(t/2)\right],\ y = a\sin t.$

(ii) $\sin x = \frac{2t}{1+t^2}$, $\tan y = \frac{2t}{1-t^2}$

(iii) $\cos x = \sqrt{\frac{1}{1+t^2}}$ $\sin y = \frac{2t}{1+t^2}$

(iv) $x = \sqrt{\sin 2u}$, $y = \sqrt{\cos 2u}$

(v) $x = \frac{\cos^3 t}{\sqrt{\cos 2t}}$, $y = \frac{\sin^3 t}{\sqrt{\cos 2t}}$.

10. Assuming the validity of the operations on the r.h.s. find $\frac{dy}{dx}$

(i) $y = \sqrt{[\sin x + \sqrt{\{\sin x + \sqrt{(\sin x + \ldots\ldots\ldots)}\}}]}$

(ii) $y = 1 \div [x + 1 \div (x + 1 \div (x + 1 \div \ldots\ldots\ldots\ldots))]$

(iii) $y = ln\,[x + ln\,(x + ln\,(x + \ldots\ldots\ldots\ldots\ldots\ldots))]$

11. (i) If $\cos y = x\cos(a+y)$ then prove that

$$\frac{dy}{dx} = \frac{\cos^2(a+y)}{\sin a}$$

(ii) If $e^{\theta\phi} = c + 4\,\theta\phi$, show that $\phi + \theta\,\frac{d\phi}{d\theta} = 0$

12. Can you differentiate $\log\log|\sin x|$? Justify your answer.

13. If $F(x) = \begin{vmatrix} f_1(x) & g_1(x) & h_1(x) \\ f_2(x) & g_2(x) & h_2(x) \\ f_3(x) & g_3(x) & h_3(x) \end{vmatrix}$, then show that

$$F'(x) = \begin{vmatrix} f'_1(x) & g_1(x) & h_1(x) \\ f'_2(x) & g_2(x) & h_2(x) \\ f'_3(x) & g_3(x) & h_3(x) \end{vmatrix} + \begin{vmatrix} f_1(x) & g'_1(x) & h_1(x) \\ f_2(x) & g'_2(x) & h_2(x) \\ f_3(x) & g'_3(x) & h_3(x) \end{vmatrix} + \begin{vmatrix} f_1(x) & g_1(x) & h'_1(x) \\ f_2(x) & g_2(x) & h'_2(x) \\ f_3(x) & g_3(x) & h'_3(x) \end{vmatrix}.$$

[**Hints** : Prove a similar result for a determinant of order 2.]

14. If $x = \frac{1-\cos^2\theta}{\cos\theta}$, $y = \frac{1-\cos^{2n}\theta}{\cos^n\theta}$ then

show that $\left(\frac{dy}{dx}\right)^2 = n^2\left(\frac{y^2+4}{x^2+4}\right)$.

15. Show the $\frac{dy}{dx}$ is independent of t if

$x = \cos^{-1}\frac{1}{\sqrt{t^2+1}}$ and $y = \sin^{-1}\frac{t}{\sqrt{t^2+1}}$.

16. If $y\sqrt{x^2+1} = \log\left\{\sqrt{x^2+1} - x\right\}$ then prove that

$$(x^2+1)\frac{dy}{dx} + xy + 1 = 0.$$

17. If $e^{y/x} = \frac{x}{a + bx}$, then show that

$$x^3 \frac{d}{dx}\left(\frac{dy}{dx}\right) = \left(x\frac{dy}{dx} - y\right)^2.$$

18. Find the points where the following functions are not differentiable.

(i) $e^{|x|}$ (iii) $|x^2 - 4|$ (iii) $|x-1| + |x-2|$ (iv) $\sin|x|$.

7.13 Second order derivatives

If f is a differentiable function of x then the derivative of $f(x)$ may determine another differentiable function of x. The new function $f'(x)$ is called the first derivative of f. If $f'(x)$ is differentiable, we can find its derived function $f''(x)$ and call it the derived function of second order. The process can be successively continued to obtain derived functions of higher orders. These functions $f'(x), f''(x), f'''(x), \ldots\ldots\ldots f^{n}(x)$ are called respectively the first order, second order, third order .... nth order derivatives of $f(x)$. Other notations used for higher order derivatives are :

$$y', y'', y''', \ldots, y^{(n)}, \ldots$$

$$y_1, y_2, y_3, \ldots, y_n, \ldots$$

$$\frac{dy}{dx}, \frac{d^2y}{dx^2}, \frac{d^3y}{dx^3}, \ldots, \frac{d^ny}{dx^n}, \ldots$$

However, the scope of the book restricts our discussion up to derivatives of second order only.

Example 34 :

Find y_2 if $y = x^5 + 4x^3 - 2x^2 + 1$.

Solution :

Differentiating successively we get

$$y_1 = \frac{d}{dx}x^5 + 4\frac{dx^3}{dx} - 2\,\frac{3}{5} + \frac{d1}{dx}$$

$$= 5x^4 + 12x^2 - 4x$$

$$y_2 = 20x^3 + 24x - 4$$

Example 35 :

If $x = \sin t$, $y = \sin(pt)$ then show that

$$(1-x^2)\frac{d^2y}{dx^2} - x\frac{dy}{dx} + p^2y = 0.$$

Solution :

$y = \sin(pt) = \sin(p\sin^{-1}x)$

$$\Rightarrow \quad \frac{dy}{dx} = \cos(p\sin^{-1}x)\,.\,\frac{p}{\sqrt{1-x^2}}$$

$$\therefore \quad (1-x^2)\left(\frac{dy}{dx}\right)^2 = p^2\cos^2(p\sin^{-1}x)$$

$$= p^2 - p^2\sin^2(p\sin^{-1}x)$$

$$= p^2 - p^2y^2.$$

Differentiating once more,

$$(1-x^2)\,2\left(\frac{dy}{dx}\right).\left(\frac{d^2y}{dx^2}\right) - 2x\left(\frac{dy}{dx}\right)^2 = -p^2\,.\,2y\,\frac{dy}{dx}$$

Dividing throughout by $2\,\frac{dy}{dx}$,

$$(1-x^2)\frac{d^2y}{dx^2} - x\frac{dy}{dx} = -p^2y \quad \text{or} \quad (1-x^2)\frac{d^2y}{dx^2} - x\frac{dy}{dx} + p^2y = 0.$$

Example 36 :

Find y_2 if :

(i) $y = (ax+b)^m,\ m \in \mathrm{R}$

(ii) $y = e^{ax}$

(iii) $y = \ln(ax+b)$

(iv) $y = \sin(ax+b)$

(v) $y = \cos(ax+b)$

(vi) $y = e^{ax}\sin(bx+c)$

(vii) $y = e^{ax}\cos(bx+c)$

Solution :

(i) $y = (ax+b)^m,\ m \in \mathrm{R}$

$\Rightarrow \quad y_1 = m(ax+b)^{m-1}a$

$y_2 = m(m-1)(ax+b)^{m-2}a^2$

(ii) $y = e^{ax}$

$\Rightarrow \quad y_1 = a\,e^{ax}$

$y_2 = a^2e^{ax}$

(iii) $y = \ln(ax+b)$

$$\Rightarrow y_1 = \frac{a}{ax+b}$$

$$y_2 = a\,.\,\frac{d}{dx}\left(\frac{1}{ax+b}\right) = a\,.\,\frac{-a}{(ax+b)^2} = -a^2(ax+b)^{-2}$$

(iv) $y = \sin(ax+b)$

$\Rightarrow \quad y_1 = a\cos(ax+b)$

$y_2 = -a^2\sin(ax+b)$

(v) Similarly incase of $y = \cos(ax+b)$ we obtain

$y_1 = -a\sin(ax+b)$ and $y_2 = -a^2\cos(ax+b)$.

It is a routine matter to find y_2 in case of other trigonometric functions.

(vi) Let $y = e^{ax}\sin(bx+c)$

$$\therefore \quad y_1 = be^{ax}\cos(bx+c) + ae^{ax}\sin(bx+c)$$

$$= e^{ax}\{a\sin(bx+c) + b\cos(bx+c)\}$$

(taking $a = r\cos\varphi$, $b = r\sin\varphi$ so that $r = \sqrt{a^2+b^2}$ and $\varphi = \tan^{-1}\frac{b}{a}$)

$$= re^{ax}\sin(bx+c+\varphi)$$

Similarly

$y_2 = r^2e^{ax}\sin(bx+c+2\varphi)$ (workout yourself)

(vii) Let $y = e^{ax}\cos(bx+c)$

$$\therefore \quad y_1 = -be^{ax}\sin(bx+c) + ae^{ax}\cos(bx+c)$$

$$= -re^{ax}\sin(bx+c)\sin\varphi + re^{ax}\cos(bx+c)\cos\varphi$$

(taking $a = r\cos\varphi$, $b = r\sin\varphi$ so that $r = \sqrt{a^2+b^2}$ and $\varphi = \tan^{-1}\frac{b}{a}$)

$$= re^{ax}\{\cos(bx+c)\cos\varphi - \sin(bx+c)\sin\varphi\}$$

$$= re^{ax}\cos(bx+c+\varphi)$$

Similarly

$$y_2 = r^2e^{ax}\cos(bx+c+2\varphi)$$

EXERCISE 7(*I*)

1. If $y = \tan^{-1}x$ prove that

   $(1+x^2)y_2 + 2xy_1 = 0.$

2. If $2y = x\left(1 + \frac{dy}{dx}\right)$ show that y_2 is a constant.

3. If $y = ax\sin x$ show that

   $x^2y_2 - 2xy_1 + (x^2+2)y = 0.$

4. If $y = e^{m\cos^{-1}x}$ then show that

   $(1-x^2)y_2 - xy_1 = m^2y.$

5. If $x = \sin t$, $y = \sin 2t$ then prove that

   $(1-x^2)\frac{d^2y}{dx^2} - x\frac{dy}{dx} + 4y = 0.$

6. If $y = (\sin^{-1}x)^2$, prove that

   $(1-x^2)y_2 - xy_1 - 2 = 0.$

7. If $y = \tan^{-1}x$, prove that $(1+x^2)y_2 + 2xy_1 = 0$.

*7.13 (a) Successive Derivatives of some standard functions

(Additional topic for interested students, not for examination)

Exercise 37 :

$$y = (ax+b)^m,\ m \in R$$

$$\Rightarrow y_1 = m(ax+b)^{m-1}a$$

$$y_2 = m(m-1)(ax+b)^{m-2}a^2$$

$$y_3 = m(m-1)(m-2)(ax+b)^{m-3}a^3$$

Proceeding in like manner we obtain

$$y_n = m(m-1)(m-2)\ldots\ldots\ldots(m-n+1)(ax+b)^{m-n}a^n. \qquad \ldots\ldots\ldots (1)$$

[This does not constitute a proof of (1). For a rigorous proof we have to make use of the principle of mathematical induction.]

Corollary 1 :

If m is a positive integer,

then from (1)

$$\frac{d^n}{dx^n}(ax+b)^m = \begin{cases} \dfrac{m!\,a^n(ax+b)^{m-n}}{(m-n)!}, & n \le m \\ 0 & , n > m. \end{cases} \qquad \ldots\ldots\ldots (2)$$

In particular when m is a positive integer, $a = 1$ and $b = 0$, then

$$\frac{d^n}{dx^n}x^m = \begin{cases} \dfrac{m!}{(m-n)!}, x^{m-n} & n \le m \\ 0 & , n > m. \end{cases} \qquad \ldots\ldots\ldots (3)$$

Corollary 2 :

If m = – 1 then from (1) we have

(i) $$\frac{d^n}{dx^n}\left(\frac{1}{ax+b}\right) = \frac{(-1)^n\, n!\, a^n}{(ax+b)^{n+1}}. \qquad \ldots\ldots\ldots (4)$$

(ii) $$\frac{d^n}{dx^n}e^{ax} = e^{ax}a^n.$$

Let $y = e^{ax}$

$$\Rightarrow y_1 = e^{ax}a$$

$$y_2 = e^{ax}a^2$$

$$y_3 = e^{ax}a^3.$$

In like manner we obtain

$$y_n = \frac{d^n}{dx^n}e^{ax} = e^{ax}a^n. \qquad \ldots\ldots\ldots (5)$$

(iii) $\frac{d^n}{dx^n} \ln(ax+b) = (-1)^{n-1}(n-1)!\, a^n (ax+b)^{-n}$.

Let $y = \ln(ax+b)$

$\Rightarrow \quad y_1 = \frac{a}{ax+b}$

$y_2 = a \cdot \frac{d}{dx}\left(\frac{1}{ax+b}\right) = a \cdot \frac{-a}{(ax+b)^2} = -a^2 (ax+b)^{-2}$

$y_3 = -a^2 \cdot (-2)(ax+b)^{-3} \cdot a = (-1)^2\, 2!\, a^3 (ax+b)^{-3}$.

In like manner we obtain

$$y_n = \frac{d^n}{dx^n} \ln(ax+b) = (-1)^{n-1}(n-1)!\, a^n (ax+b)^{-n}. \qquad \text{......... (6)}$$

(iv) $\frac{d^n}{dx^n} \sin(ax+b) = a^n \sin\left(n\frac{\pi}{2} + ax + b\right)$

Let $y = \sin(ax+b)$

$\Rightarrow \quad y_1 = a\cos(ax+b) = a\sin\left(\frac{\pi}{2} + ax + b\right)$

$y_2 = a^2\cos\left(\frac{\pi}{2} + ax + b\right) = a^2\sin\left(2\frac{\pi}{2} + ax + b\right)$

$y_3 = a^3\cos\left(2\frac{\pi}{2} + ax + b\right) = a^3\sin\left(3\frac{\pi}{2} + ax + b\right)$.

In like manner we obtain

$$y_n = \frac{d^n}{dx^n} \sin(ax+b) = a^n \sin\left(n\frac{\pi}{2} + ax + b\right). \qquad \text{......... (7)}$$

Similarly we obtain

$$\frac{d^n}{dx^n} \cos(ax+b) = a^n \cos\left(n\frac{\pi}{2} + ax + b\right). \qquad \text{......... (8)}$$

Theorem 7 : (LEIBNITZ THEOREM)

If u and v are two differentiable functions having nth derivative, then the function uv is differentiable n times and the nth derivative is given by

$$(uv)_n = u_n v + c(n,1)\, u_{n-1} v_1 + c(n,2)\, u_{n-2} v_2 + \ldots\ldots + c(n,r)\, u_{n-r} v_r + \ldots\ldots + u v_n \qquad \text{.........(9)}$$

The proof is by the principle of mathematical induction. Hints for the proof are given below.

For $n = 1$, $(uv)_1 = u_1 v + u v_1$.

Assuming the statement (9) to be true for n = m and differentiating both sides we shall obtain

$(uv)_{m+1} = (u_{m+1} v + u_m v_1) + c(m,1)(u_m v_1 + u_{m-1} v_2) + \ldots\ldots + c(m,r)(u_{m-r+1} v_r + u_{m-r} v_{r+1}) + \ldots\ldots \ldots\ldots\ldots\ldots + u_1 v_m + u v_{m+1}$.

$= (u_{m+1} v + (c(m,0) + c(m,1)) u_m v_1 + (c(m,1) + c(m,2))\, u_{m-1} v_2 + \ldots\ldots + (c(m, r-1) + c(m,r))\, u_{m-r+1} v_r + \ldots\ldots + u v_{m+1}$.

$= u_{m+1}v + c(m+1, 1)\, u_m v_1 + c(m+1, 2)\, u_{m-1}v_2 + \ldots\ldots + c(m+1, r)\, u_{m+1-r}v_r + \ldots\ldots + u v_{m+1}$

and so the induction is complete.

Example 38 :

Find y_n it $y = x^2 \sin x$.

Solution :

Take $u = \sin x$, $\vartheta = x^2$.

Applying Leibnitz Theorem,

$$y_n = (\sin x\, x^2)_n$$

$$= (\sin x)_n\, x^2 + c(n, 1)\, (\sin x)_{n-1}\, 2x + c(n, 2)\, (\sin x)_{n-2} 2$$

$$= x^2 \sin\left(n\frac{\pi}{2} + x\right) + 2nx \sin\left\{(n-1)\frac{\pi}{2} + x\right\} + n(n-1)\sin\left\{(n-2)\frac{\pi}{2} + x\right\}.$$

It is worthwhile to note that while applying Leibnitz theorem, the success depends on the choice of the second function. In the above example we chose x^2 as our second function because the terms after the third are all zero.

Example 39 :

If $y = \dfrac{2x+1}{(x+2)(x-3)}$ find y_n.

Solution :

Breaking into partial fractions, we have

$$y = \frac{2x+1}{(x+2)(x-3)} = \frac{3}{5}\,\frac{1}{x+2} + \frac{7}{5}\,\frac{1}{x-3}$$

$$\Rightarrow \quad y_n = \frac{3}{5}\,\frac{(-1)^n\, n!}{(x+2)^{n+1}} + \frac{7}{5}\,\frac{(-1)^n\, n!}{(x-3)^{n+1}} \qquad \text{by (4)}$$

$$= \frac{(-1)^n\, n!}{5}\left[\frac{3}{(x+2)^{n+1}} + \frac{7}{(x-3)^{n+1}}\right].$$

Example 40 :

If $y = \cos^4 x$ find y_n.

Solution :

$$\cos^4 x = (\cos^2 x)^2 = \left(\frac{1+\cos 2x}{2}\right)^2$$

$$= \frac{1}{4}\,(1 + 2\cos 2x + \cos^2 2x)$$

$$= \frac{1}{4}\left\{1 + 2\cos 2x + \frac{1}{2}\,(1+\cos 4x)\right\}$$

$$= \frac{3}{8} + \frac{1}{2}\cos 2x + \frac{1}{8}\cos 4x$$

Hence using (8)

$$y_n = 0 + \frac{1}{2}\,2^n \cos\left(n\,\frac{\pi}{2} + 2x\right) + \frac{1}{8}\,.\,4^n \cos\left(n\,\frac{\pi}{2} + 4x\right)$$

$$= 2^{n-1} \cos\left(n\,\frac{\pi}{2} + 2x\right) + 2^{2n-3} \cos\left(n\,\frac{\pi}{2} + 4x\right).$$

Example 41 :

Find y_n if $y = e^{ax} \sin bx$.

Solution :

$$y = e^{ax} \sin bx$$

$$\Rightarrow \quad y_1 = ae^{ax} \sin bx + e^{ax}\, b \cos bx$$

$$= e^{ax} (a \sin bx + b \cos bx)$$

Let $a = r \cos\phi$, $b = r \sin\phi$.

Then $a \sin bx + b \cos bx = r \sin (bx + \phi)$ .... (10)

$$\therefore \quad y_1 = r\, e^{ax} \sin (bx + \phi)$$

$$\Rightarrow \quad y_2 = r \{ ae^{ax} \sin (bx + \phi) + e^{ax}\, b \cos (bx + \phi) \}$$

$$= r\, e^{ax} \{ a \sin (bx + \phi) + b \cos (bx + \phi) \}$$

$$= r^2\, e^{ax} \sin (bx + 2\phi), \qquad \text{using (10)}$$

Assume that

$$y_m = r^m\, e^{ax} \sin (bx + m\phi)$$

Then $y_{m+1} = r^m \{ ae^{ax} \sin (bx + m\phi) + e^{ax}\, b \cos (bx + m\phi) \}$

$$= r^m\, e^{ax} [\, a \sin (bx + m\phi) + b \cos (bx + m\phi) \,]$$

$$= r^{m+1}\, e^{ax} \sin (bx + \overline{m+1}\,\phi), \qquad \text{using (10)}$$

Hence by induction,

$$y_n = r^n\, e^{ax} \sin (bx + n\phi)$$

where $r = (a^2 + b^2)^{\frac{1}{2}}$ and $\phi = \tan^{-1} \dfrac{b}{a}$.

7.14 Some basic theorems (Mean Value Theorems)

Theorem 8 (a) Rolle's Theorem :

If a function f is

(i) continuous on the closed interval $[\, a, b \,]$

(ii) differentiable on the open interval (a, b) and

(iii) $f(a) = f(b)$,

then there exists at least a point $c \in (a, b)$ such that $f'(c) = 0$.

The proof of this theorem is beyond the scope of the book. We satisfy ourselves with the following geometrical explanation.

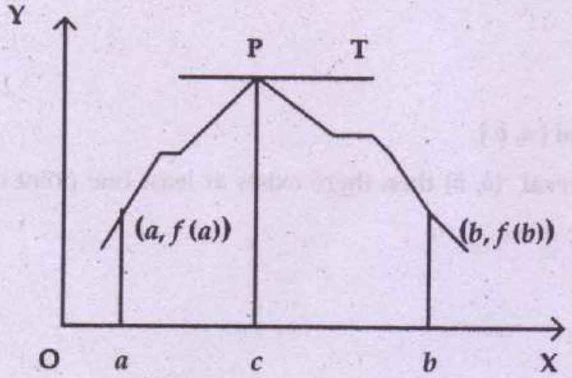


Fig. 12.10

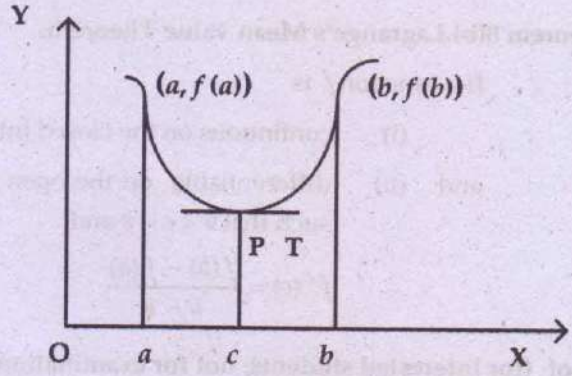


Fig 12.11

If f is a constant in $[a, b]$, $f'(x) = 0$ for all $x \in (a, b)$ and the theorem is trivially true. Now at $x = a$, $f(x)$ has the value $f(a)$. Since $f(b) = f(a)$, if $f(x)$ is not a constant in $[a, b]$, $f(x)$ either increases or decreases as x increases and therefore must either decrease or increase respectively to assume the value $f(b)$.. So there must be a point $c \in (a, b)$ such that for some $\delta > 0$,

either (i) $f'(x) > 0$ for $x \in (c - \delta, c)$ and $f'(x) < 0$ for $x \in (c, c + \delta)$ as in fig. 12.10.

or (ii) $f'(x) < 0$ for $x \in (c - \delta, c)$ and $f'(x) > 0$ for $x \in (c, c + \delta)$ as in fig 12.11.

implying that $f'(c) = 0$ since f is differentiable at every point in (a, b).

Geometrical interpretation of Rolle's theorem :

$f'(c)$ is the slope of the tangent to the curve represented by $y = f(x)$ at the point $(c, f(c))$, where $a < c < b$.

$f'(c) = 0$ implies that the tangent to the curve $y = f(x)$ at $(c, f(c))$ is parallel to the x-axis.

Thus if the function f satisfies all conditions of Rolle's theorem on the interval $[a, b]$, then there exists at least a point $(c, f(c))$ on the curve represented by $y = f(x)$ where the tangent is parallel to the x-axis.

Example 42 :

Verify Rolle's theorem for the function f defined by $f(x) = x^2 - 3x + 2$ on the interval $[1, 2]$.

Solution :

Clearly the given function satisfies the first two conditions of Rolle's theorem on $[1, 2]$.

We observe that $f(1) = f(2) = 0$.

So the third condition of Rolle's theorem is also satisfied by the function f on $[1, 2]$.

We have $f'(c) = 2c - 3$.

$$f'(c) = 0 \quad \Rightarrow 2c - 3 = 0$$

$$\Rightarrow c = \frac{3}{2}$$

$$c = \frac{3}{2} \in (1, 2).$$

Thus there exists a point $c = \frac{3}{2}$ with $1 < c < 2$ such that $f'(c) = 0$.

This is the verification of Rolle's theorem for the given function f on the interval $[1, 2]$.

Theorem 8(b) Lagrange's Mean value Theorem.

If a function f is

(i) continuous on the closed interval $[a, b]$

and (ii) differentiable on the open interval (a, b) then there exists at least one point c such that $a < c < b$ and

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Proof (for interested students; not for examination)

Here f satisfies all conditions of Rolle's theorem except that $f(a) = f(b)$. So assume that $g(x) = f(x) - Ax$ where A is a constant to be chosen suitably so that g satisfies all conditions of Rolle's theorem, i.e. $g(a) = g(b)$ is satisfied. This gives

$$A = \frac{f(b) - f(a)}{b - a}.$$

Since g satisfies all conditions of Rolle's theorem there exists a point $c \in (a, b)$ such that $g'(c) = 0$

i.e. $f'(c) = A$ $\quad (\because g'(x) = f'(x) - A)$

$\Rightarrow \quad f'(c) = \dfrac{f(b) - f(a)}{b - a}$ substituting the value of A.

Geometric Interpretation :

Geometrically the Mean value Theorem says that under the hypotheses we can find a point P $(c, f(c))$ on the curve where $a < c < b$ such that the tangent line at P is parallel to the chord joining $(a, f(a))$ and $(b, f(b))$.

Example 43 :

Verify Lagrange's Mean-value theorem for the function f defined lay $f(x) = 2x^2 - 3x + 7$ on the interval $[1, 2]$.

Solution :

Clearly the given function satisfies all conditions of Lagrange's Mean-value theorem on $[1, 2]$.

We have $f(x) = 2x^2 - 3x + 7$.

So $f(1) = 6$ and $f(2) = 9$.

$$f'(c) = 4c - 3.$$

$$\frac{f(2) - f(1)}{2 - 1} = \frac{9 - 6}{2 - 1} = 3.$$

$$f'(c) = \frac{f(2) - f(1)}{2 - 1}$$

$$\Rightarrow 4c - 3 = 3$$

$$\Rightarrow c = \frac{3}{2}.$$

$$c = \frac{3}{2} \in (1, 2).$$

Thus there exists a point $c = \frac{3}{2}$ with $1 < c < 2$ such that $f'(c) = \frac{f(2) - f(1)}{2 - 1}$.

This is the verification of Lagrange's Mean-value theorem for the given function f on the interval [1, 2].

Theorem 8(c) Cauchy's Mean value Theorem. (For interested students; not for examination)

(For interested students; not for examination)

If f and g are two functions such that

(i) both are continuous on $[a, b]$

(ii) both are differentiable on (a, b)

and (iii) $g'(x) \neq 0$ for any $x \in (a, b)$.

then there exists at least a point c such that $a < c < b$ and

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}.$$

Proof :

We define a new function F (x) as follows.

$$F(x) = f(x) - f(a) - \frac{f(b)-f(a)}{g(b)-g(a)}(g(x) - g(a)). \quad \text{......... (1)}$$

Observe that $g(b) - g(a) \neq 0$. Otherwise we could apply Rolle's theorem to g to obtain $g'(y) = 0$ for some $y \in (a, b)$ contradicting (iii).

Now $F(a) = 0 = F(b)$ and F satisfies other conditions of Rolle's theorem. Hence there exists a point c such that $a < c < b$ and $F'(c) = 0$

i.e. $$f'(c) - \frac{f(b)-f(a)}{g(b)-g(a)} g'(c) = 0$$

$$\Rightarrow \quad \frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}.$$

As an illustration of the utility of these theorems in mathematical analysis consider the following example.

Example 44 :

Prove that if the derivative of a function is zero in an interval then the function is a constant in that interval.

Solution :

Let $f'(x) = 0$ for all $x \in (a, b)$. Let x_1, x_2 be any two numbers in (a, b) such that $a < x_1 < x_2 < b$. Under the hypothesis, Lagrange's Mean value Theorem is applicable in the interval $[x_1, x_2]$ and so we have a $c \in (x_1, x_2)$ such that

$$f'(c) = \frac{f(x_2)-f(x_1)}{x_2 - x_1}.$$

But $f'(c) = 0$ by hypothesis. This implies that $f(x_1) = f(x_2)$ with $x_1 \neq x_2$. Thus $f(x)$ has the same value at any two distinct points of the interval (a, b) and hence is a constant there in.

EXERCISE- 7(m)

1. Verify Rolle's theorem for the function $f(x)=x(x-2)^2$, $0 \le x \le 2$.
2. Examine if role's theorem is applicable to the following functions :

   (i) $f(x) = |x|$ on $[-1, 1]$ (ii) $f(x) = [x]$ on $[-1, 1]$

   (iii) $f(x) = \sin x$ on $[0, \pi]$ (iv) $f(x) = \cot x$ on $[0, \pi]$
3. Verify Lagrange's Mean-Value theorem for

   $F(x) = x^3-2x^2-x+3$ on $[1,2]$
4. Test if Lagrange's mean value theorem holds for the functions given in question no. 2.

5*. (Not for examination)

Verify Cauchy's mean value theorem for the functions x^2 and x^3 in $[1, 2]$.