FINITE MATHEMATICS

Abdulqader Othman

Department of Mathematics
Koya University

2025

List of Tables

List of Figures

Dedication

To my beloved wife and children

Preface

This book is based on lectures developed by the author for B.SC and M.SC. Students at Koya University, Department of Mathematics, and on the lectures taught by the author in postgraduate studies to master's and doctoral students at Salahaddin-Erbil and Sulaymaniyah Universities, Department of Mathematics, as well the book is also the product of the accumulated notes from teaching under the author's postgraduate studies.

In mathematical structures and mathematics education, finite Mathematics is a syllabus and method that depends on foundations of mathematics, while it is independent of calculus. Thereby, since we devoted ourselves to writing the book Foundations of Mathematics (Hamadameen, 2022), it was necessary and inevitable for us to write this book in order to help the student and reader apply mathematical problem solving and logical thinking to real world phenomena, making it an important field of knowledge for students who are pursuing careers in the field of business and its branches (Applied mathematics, social sciences, computer sciences, and applied sciences in the fields of statistics, probabilities, medicine, physics, chemistry, biology, most engineering branches, and other practical professional specializations).

The book takes into consideration the necessity of the contents of this book for students to study mathematics as well as the other applied sciences mentioned above. The contents of this book and its academic goals aim at:

(i) Helping students to be fully familiar with the mathematical induction and its principals. When can mathematical induction be used, what are the conditions for it, what is its application, and where can it be used in real life?

(ii) Motivating students to learn about complex numbers, their properties, their representation in the Cartesian plane, and algebraic operations on them. Polar form of complex numbers, and applications of De Moivre's theorem in the field of complex numbers. What conjugate numbers, their properties, and what are the relations between complex numbers and their conjugates? In addition to absolute value inequalities of complex numbers, their square roots, and roots of unity.

(iii) What are polynomials, and their properties? Quotient of polynomials, long division algorithm of polynomials, their roots, and duplicate roots. Greatest common factor of polynomials, Cardian's method to solve cubic equations, quartic equations, and method to solve them.

(iv) Helping students to be familiar with numerical solutions to nonlinear equations and when to applies to such solutions. Also, helping the students to find approximate values for the roots of nonlinear equations using some practical methods, including Descartes' sign rule and Horner's method. In addition, presenting numerical methods to the student to find the approximate values of the roots of equations, including the bisection method, Newton- Raphson's method, secant method, Birge-Vita method, and Graeffe's root-squaring method. And making the student understand that all of these methods are practical, applicable, and have a solid algorithm for application on the computer.

(v) Considering matrices to students and expressing the system of linear equations as matrices. Types of matrices and their properties, operations on matrices in addition to matrices and linear vectors as matrices, as well as identifying dependent and independent linear vectors and how to find solutions to the linear systems in the form of matrices via operations on rows and columns of the matrices. Moreover, considering determinants as an inevitable result of the matrices, and the related concepts to them. Types of determinants, how to find them, and the general formula for finding determinants. Permutations and determinants, and the relationship between them, and the main property of determinants. Furthermore, Inverse of matrix, elementary transformations of the matrix, and inverse transformations. Equivalence, norm, form, and rank of matrices. Besides, inserting matrix inverse methods. And, how getting inverse of complex matrix, and what are the methods to find the inverse of complex matrix?

(vi) Encouraging students to turn to the system of linear equations and their role in solving real life problems and how to formulate them. Provide some methods like; Cramer's method, Gauss's method, and some other methods. In addition to those methods and their operations on the computer.

(vii) Explaining the concepts of eigenvalues and eigenvectors to the student and explaining their advantages. What are eigenvalues and eigenvectors through changing direction transformations and eigenvalues? Applications of eigenvalues and eigenvectors, and their properties of matrices. Moreover, introducing two different methods for finding eigenvalues.

(viii) Considering each of permutations and combinations and how to formulate them. What is the basic principle in arithmetic? The basic properties of permutations and combinations and the difference between them. Embarking on the binomial theorem and the polynomial theorem.

It is noteworthy that, most of the theorems, corollaries, and exercises in this book are adapted from the references (Balfour and Beveridge, 1972; Britton and Snively, 1954; Brown and Churchill, 2009;

Conte and De Boor, 2017; Fraser, 1958; Froberg, 1965; Goult, 1974; Hoffman and Kunze, 1967; Hohn, 1972; Knopp, 1952; MacDuffee, 1954; Parsonson, 1970; Ralston and Rabinowitz, 2001; Ralston, 1965; Strang, 2006; Uspensky, 1948; Wilkinson, 1971).

The contents of this book are organized as follows: chapter 1 is dedicated to discussing to the mathematical induction, the basic concepts, and the principal of it. Chapter 2, deals with the complex numbers, properties, Cartesian representations, polar form of a complex numbers. In addition to De Moivre's theorem, complex numbers and their conjugates and roots of complex numbers. Chapter 3 is devoted to polynomials in which it defines the concept of polynomials, the properties of polynomials, and the long division algorithm for polynomials. It also shows the relationship between roots and polynomial equations, repeated roots, the greatest common factor of polynomials, solving cubic equations using Cardan's method and solving quadratic equations. Numerical solution of nonlinear equations, finding the differential via Horner's method, and Numerical methods for finding approximate values of them took their place in chapter 4. Chapter 5 deals with the matrices, types of matrices, operations on matrices, matrices partition, vector expression, linearly dependent vectors, and linearly independent vectors. Chapter 6 deals with the determinants, types of determinants, algorithm for finding determinant of a matrix of third order or higher, General methods for finding determinants, permutations and the determinant, and properties of determinants. Chapter 7 is about inverse of a matrix, matrix inverse methods like; the method of adjoint matrix, and the method of elementary transformations. In addition to the method of transformations on rows including; Jacobian method, the method of Triangularization, and the method of Escalator. Moreover, considered inverse of a complex matrix, and method to find the inverse of it. Chapter 8, deals with numerical solution of a system of linear equations, mathematical formulation of linear system, solutions for systems with equal equations and variables including; Cramer's method, Gauss's method, Gauss's method and row echelon form, coefficient matrix partition method, and matrix inverse method. Furthermore, the methods and their operations on the computer. eigenvalues and eigenvectors, and their features. Chapter 9 is about eigenvalue and eigenvector through changes direction, transformations and eigenvalues, polynomial equation of degree n in eigenvalue, eigenvalues and eigenvectors of matrices. Conclusions from eigenvalues, eigenvectors and traces. Methods for finding eigenvalues like; LU method, and Jacobi method. Finally, chapter 10 considered permutations and combinations, and their formulations. Basic principle of the arithmetic of of permutations, and combinations. In addition to difference between combination and permutations, binomial theorem, multinomial theorem, and harmonic series with its properties. And, summation by fragmentation.

It is worth to be mentioned that theorems and their corollaries are printed in *italics*, while, the end of the proofs to theorems and corollaries are indicated by the symbol ⧫.

Abdulqader Othman
Department of Mathematics, Faculty of Science & Health
Koya University
2025

1

Mathematical Induction

1.1 Introduction

There are two kinds of mathematical proofs; mathematical deduction and mathematical induction. For illustration, assume that $S_1, S_2, ..., S_n$ are axioms, and theorems have been proved in the past. So to prove $P \Rightarrow Q$ it is enough to prove that $S_1, S_2, ..., S_n, P \vdash Q$ its true argument. And this process is called deduction. But, mathematical induction is a type of mathematical proof usually used to prove that an equation or difference is true for an infinite set of numbers, such as natural or positive integer numbers (Hamadameen, 2022). This proof involves three stages. First, demonstrate that the initial number in the set meets the requirement. Second, assume the desired condition is satisfied by an arbitrary element in the set. Lastly, algebraically prove that this condition holds true for the next number in the sequence. Let us consider on the Fig. 1.1 below of the similar shapes and Fig. 1.2 on the image of the domino game, and what happens when the first piece falls and its effect on the remaining pieces of the set of pieces.

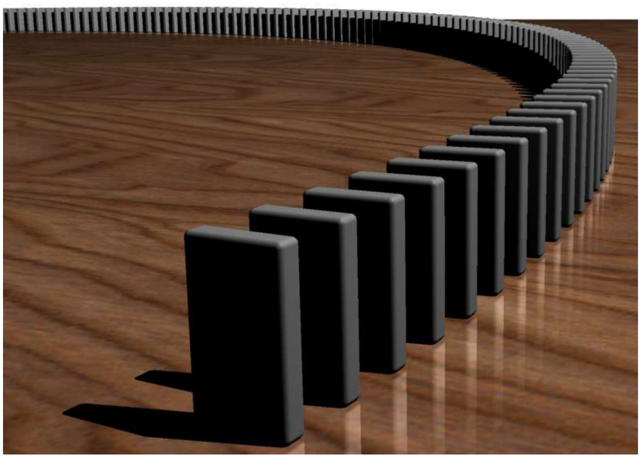


Figure 1.1: Mathematical induction as the effects of dominoes falling

Figure 1.2 The above dominoes are standing, bottom dominoes are in motion

1.2 Mathematical induction

The principle of summation is the basis adopted in mathematical induction. But in the case of a finite algebraic sum, the need arises to write the sum of the formulation instead of writing all the terms, because when the sums are infinite, it is impossible to write all the terms, so two or three terms may be written, followed by three points. For example:

(i) $1 + 2 + 3 + ...$, or

(ii) the nth term, for example; $1 + 2 + 3 + ... + n + ...$, or

(iii) $1 \times 2 + 2 \times 3 + 3 \times 4 + ... + n \times (n + 1) + ...$

However, the writing of the summation terms remains long, tedious, and not well-defined mathematically. For example, the summation of $1+2+3+...+128$, may be the aim is $1+2+3+...+n+...+128$, or summation of 128-term, or may be the aim is $1+2+4+...2^n+...+128$, percissely the summation of 8-term only.

Thereby, the writing of the sign $\sum$ with addition of its lower bond and upper bound is necessary to the summation operation. For example;

(i) $\sum_{n=1}^{\infty} n = 1 + 2 + 3 + ...,$

(ii) $\sum_{n=10}^{17} n = 10 + 11 + 12 + 13 + 14 + 15 + 16 + 17,$

(iii) $\sum_{t=1}^{n} a_{tm} = a_{1m} + a_{2m} + a_{3m} + ... + a_{nm},$

(iv) $\sum_{t=1}^{n} a_{tmk}b_{tmk}c_{tmk} = a_{1mk}b_{1mk}c_{1mk} + a_{2mk}b_{2mk}c_{2mk} + a_{3mk}b_{3mk}c_{3mk} + ... + a_{nmk}b_{nmk}c_{nmk}.$

Example 1.1 $\sum_{t=1}^{7}(t+t^2)^2 = (1+1^2)^2+(2+2^2)^2+(3+3^2)^2+(4+4^2)^2+(5+5^2)^2+(6+6^2)^2+(7+7^2)^2 = 4+36+144+400+900+1764+3136 = 6384.$

Definition 1.1 Mathematical induction is a method for proving that a statement $I(n)$ is true for for every natural number n, that is, that the infinitely many cases $I(0), I(1), I(2), ..., I(n)$ all hold (Anderson, 1979; Bather, 1994).

1.3 Exercises

Solve the following problems:

Q1: Evaluate each of;

(i) $\sum_{k=-5}^{8} k$.

(ii) $\sum_{k=-5}^{0} k^2$.

(iii) $\sum_{n=1}^{5} \frac{1}{n}$.

(iv) $\sum_{n=3}^{7} \frac{1}{(n+1)(n+2)}$.

(v) $\sum_{n=1}^{6} \frac{1}{6}$.

(vi) $\sum_{n=3}^{11}(a_n + b_n + c_n)$.

(vii) $\sum_{n=-3}^{5} \frac{(n+1)(n+2)}{n-4}$.

(viii) $\sum_{n=0}^{\infty}(3-n)^{3n}$.

(ix) $\sum_{n=0}^{\infty}(-1)$.

Q2: State whether the following statements are true or false. Give reasons for your answers.

(i) $\sum_{n=0}^{\infty} cp_n = c\sum_{n=0}^{\infty} p_n$.

(ii) $\sum_{n=1}^{m} f(n) = \sum_{n=1}^{m+1} f(n)$.

(iii) $(\sum_{k=1}^{m} a_k)^2 = \sum_{k=1}^{m} a_k^2 + \sum_{k=1}^{m} 2a_k + \sum_{k=1}^{m} 1$.

(iv) $\sum_{k=1}^{m}(a_k + b_k)^2 = \sum_{k=1}^{m} a_k^2 - \sum_{k=1}^{m} 2a_k b_k + \sum_{k=1}^{m} b_k^2$.

(v) $\sum_{k=0}^{n} a_k = \sum_{k=0}^{n} a_(n-k)$.

(vi) $\sum_{k=1}^{5}(3k + b_k) = c + \sum_{k=1}^{5} b_k$.

(vii) $\sum_{k=1}^{3}(a_k - b_j) = \sum_{k=1}^{3} a_k - \sum_{k=1}^{3} b_k$.

(viii) $\sum_{k=1}^{n}(a_k + c) = \sum_{k=1}^{n} a_k + nc$.

(ix) $\sum_{i=1}^{n}(a_(i+1-a_i) = a_(n+1)+a_1$.

(x) $\sum_{k=2}^{4}[(k+1)^3 - k^3] = 4^3 - 2^3$.

(xi) $\sum_{n=1}^{30}(-1)^n = 0$.

(xii) $1^2 + 2^2 + 3^2 + ... + n^2 = \frac{n(n+1)(2n+1)}{6}, \forall n \geq 1$.

(xiii) $3 + 33 + 333 + ... + 33...3 = \frac{10^{n+1}-9n-10}{27}$.

1.4 The principle of mathematical induction

There are two types of methods for proving mathematical results: deduction and induction. Induction cannot be a method of proof, because it is used in almost all sciences, and in addition to its use in mathematics, it is based on the Intuition principle, not on mathematical logic (Rabinovitch, 1970; Francesco, 1575; Henkin, 1960; Gunderson, 2014). As for mathematical induction, it is used in mathematics only. It is one of the methods of proof that depends on sequential mathematical logic that covers the gaps in induction.

Consider a statement $I(n), n \in \mathbb{N}$. Then to determine the validity of $I(n), \forall n$, use the following principle::

(i) Check whether the given statement is true for $n = 1$.

(ii) Assume that given statement $I(n)$ is also true for $n = k$, where k is any natural number.

(iii) Prove that the result is true for $I(k+1)$ for any $k \in \mathbb{N}$.

If the above-mentioned conditions are satisfied, then it can be concluded that $I(n)$ is true for all n natural numbers.

Example 1.2 Using mathematical method, prove that;
$1 + 2 + 3 + ... + n = \frac{n(n+1)}{2}, n \in \mathbb{N}$.
Solution: Let the given statement $I(n)$ be defined as;
$1 + 2 + 3 + ... + n = \frac{n(n+1)}{2}, n \in \mathbb{N}$.

(i) Put $n = 1$

R.H.S. $I(1) = 1$.

L.H.S. $I(1) = \frac{1(1+1)}{2} = 1$.

$\because$ R.H.S.= L.H.S. for $n = 1$,

$\therefore I(1)$ is true.

(ii) Let us assume that the statement is true for $n = k$, $i.e.I(k)$ is true. Or,

$1 + 2 + 3 + ... + k = \frac{k(k+1)}{2}, k \in \mathbb{N}$.

(iii) To prove that $I(k+1)$ is true;

$I(k+1) = 1+2+3+...+k+(k+1) = I(k)+k+1 = \frac{k(k+1)}{2}+k+1 = \frac{k(k+1)+2(k+1)}{2}$.

$\therefore I(k+1)$ is true.

Thus if $I(k)$ is true, then $I(k+1)$ is also true.

$\therefore I(n), \forall n \in \mathbb{N}$.

Thus, $1 + 2 + 3 + ... + n = \frac{n(n+1)}{2}, n \in \mathbb{N}$.

Example 1.3 Using mathematical method, prove that;
$3^n > 3n - 2, \forall n \in \mathbb{N}$.
Solution: Let the given statement $I(n)$ be defined as;
$3^n > 3n - 2, \forall n \in \mathbb{N}$.

(i) Put $n = 1$

R.H.S. $I(1) = 3^1 = 3$.

L.H.S. $I(1) = 3(1) - 2 = 1$.

$\because$ R.H.S.$>$ L.H.S. for $n = 1$,

$\therefore I(1)$ is true.

(ii) Let us assume that the statement is true for $n = k$, $i.e.I(k)$ is true. Or,

$3^k > 3k - 2, \forall k \in \mathbb{N}$.

(iii) To prove that $I(k+1)$ is true;

$\because 2 \cdot 3^k > 3,$

$\therefore 3^k + 2 \cdot 3^k > 3k - 2 + 3.$

$3^k(1+2) > (3k+3) - 2.$

$\therefore 3^k \cdot 3 > 3(k+1) - 2.$

$\therefore 3^{k+1} > 3(k+1) - 2.$ This is the same inequality, where n has been changed by $k+1$.

$\therefore I(k+1)$ is true.

Thus if $I(k)$ is true, then $I(k+1)$ is also true.

$\therefore I(n), \forall n \in \mathbb{N}$ is true.

Example 1.4 Using mathematical induction, prove that $I(n) = 3^n - 1$ is a multiple of 2.

Solution:

(i) Put $n = 1$

R.H.S. $I(1) = 3^1 - 1 = 3 - 1 = 2.$

2 is a multiple of 2.

$\therefore I(1)$ is true.

(ii) Let us assume that the statement is true for $n = k$, *i.e.*$I(k)$ is true. Or, $3^k - 1, \forall k \in \mathbb{N}$ is a multiple of 2.

(iii) Now, we have to prove that $I(k+1) = 3^{k+1} - 1$ is true.

$\because 3^{k+1} - 1 = 3 \cdot 3^k - 1 = (2+1) \cdot 3^k - 1 = 2 \cot 3^k + 3^k - 1.$

$\because 2 \cot 3^k$ and $3^k + 3^k - 1$ are multiples of 2,

$\therefore 2 \cot 3^k + 3^k - 1$ is a multiple of 2.

$\therefore I(k+1), \forall k \in \mathbb{N}$ is true.

Thus, $I(n) = 3^n - 1, \forall n \in \mathbb{N}$ is a multiple of 2.

1.5 Exercises

Answer the following questions:

Q1: Use mathematical induction to prove each of claims:

(i) $\sum_{l=1}^{n}(l+2)(3l+1) = n(n+2)(n+3)$.

(ii) $\sum_{t=1}^{n} \frac{1}{t(t+1)} = \frac{n}{n+1}$.

(iii) $\sum_{t=1}^{n} ar^{t-1} = \frac{a(r^n-1)}{r-1}$.

(iv) $\sum_{t=2}^{n+1} t2^{t-2} = n2^n$.

(v) $\sum_{r=1}^{n} r^3 = \frac{[n(n+1)]^2}{4}$.

Q2: Express the following mathematical statements by using ($\sum$) symbol.

(i) $a + (a+b) + ... + [a + (n-1)b]$.

(ii) $a - (a-b) + ... + [a - (n-1)b]$.

(iii) $cot\alpha + cot(\alpha+\beta) + ... + cot[(\alpha + (n-1)\beta])$.

(iv) $cot\alpha - cot(\alpha-\beta) + ... + cot[(\alpha - (n-1)\beta])$.

Q3: Prove the following claims:

(i) $a^n + b^n$ is divisible by $a+b, \forall n \in \mathbb{Z}_o^+$.

(ii) $a^n - b^n$ is divisible by $a-b, \forall n \in \mathbb{Z}_e^+$.

(iii) $n^2 - n + 2, \forall n \in \mathbb{Z}_e^+$ is an even number.

(iv) $2^{n+2} + 3^{2n+1} \forall n \in \mathbb{N}$ is a multile of 7.

(v) $a^n \geq n+1, \forall n \in \mathbb{N}, a \geq 2$.

Q4: For all positive real number, prove that $a^{n-1} = 1, \forall a \in \mathbb{N}^+$.
Q5: If

$$U_n - 4U_{n-1} + 4U_{n-2} = 0,$$
$$U_0 = 1,$$
$$U_1 = 4,$$

then prove that

$$U_n = (n+1)^{2^n}.$$

2
Complex Numbers

2.1 Introduction

When we talk about complex numbers and their creation and the imaginary number, the need for complex numbers and imaginary numbers appears, since mathematicians can not figure out the square root of a negative number.

The roots of negative numbers were dealt with for the first time in the sixteenth century, when the scientist Girolamo Cardano in 1545 (Merino, 2006) found that some problems have a solution in terms of square roots of negative numbers, and although he believed that such a solution was unrealistic, he was sure that there was no other solution! People agreed with him that it was useless for nearly a century.

After that the scientist René Descartes in 1637 (Blank, 1999; Jakob and Waerden, 1985) developed the standard formula for imaginary numbers, which later led to the algebraic formula for the complex number $a + ib$. This formula means that the complex number is composed of two parts: the first (real) part a is just a real number, and the second (imaginary) part ib is a real number b multiplied by the imaginary number i, and the sum of these two parts is the complex number. However, Descartes did not believe in complex numbers much, and assumed that if they were used to solve a problem, you would not reach a conclusion! In fact, he called them "imaginary numbers".

For more than a century after that, the opinions of scholars varied between believing in the possibility of the existence of the number i and trying to prove its existence, and refusing to add a new number!

Now complex numbers have become an important part and field of mathematics and have an important and vital role in theoretical applications as well as in practical, productive, and life applications such as; topics that use complex numbers include research on electric current, wavelength, fluid flow and its relationship to obstructions, stress analysis of beams, movement of shock absorbers in automobiles, the study of resonance for structures, design of generators and electric motors and in large matrices used in modeling, and so on.

2.2 Negative discriminant of a polynomial

Consider the following polynomial in a unique variable and in the second degree;

$$ax^2 + bx + c = 0, 0 \neq a, b, c \in \mathbb{R} \tag{2.1}$$

The set solution for (2.1) is $x = \frac{-b \mp \sqrt{b^2-4ac}}{2a}$, *i.e.* the equation has two different roots in $\mathbb{R}$ provided that the distinctive amount $b^2 - 4ac \geq 0$, otherwise the value of the square root will be negative. Thus, the need arises to define another group of numbers called the group of complex numbers, through which the square root of negative numbers can be defined.

2.3 Complex numbers

A complex number is a number that can be expressed in the form $a + ib$ where a and b are real numbers, and i is a symbol called the imaginary unit and satisfying the equation $i^2 = -1$. The set of complex numbers is denoted by $\mathbb{C}$ (Bourbaki, 1994; Andreescu and Andrica, 2006), and can be defined as follows:

Definition 2.1 The formula $\mathbb{C} = \mathbb{R} \times \mathbb{R}\{(x, y)|x, y \in \mathbb{R}\} = \{a + ib, a, b \in \mathbb{R}, i = \sqrt{-1}\}$ is called the set of the complex numbers (Hamadameen, 2022), as shown in the Figure 2.1.

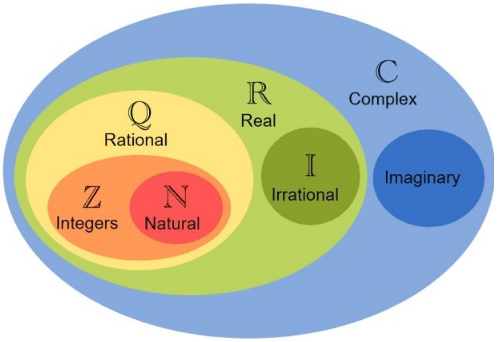


Figure 2.1: Number sets

Notation: The set $\mathbb{C} = \left\{x + iy | x, y \in \mathbb{R}, i = \sqrt{-1}\right\}$, where the imaginary part has the properties:

(i) The term iy is the real number y multiplied by i.

(ii) $iy = yi$.

(iii) $(y_1 + y_2)i = iy_1 + iy_2$.

(iv) If $iy = 0$ then $y = 0$.

2.4 Properties of complex numbers

There are basic properties of the set of complex numbers that distinguish it from other set numbers. We will list them as follows:

(i) The complex number is called equal to zero if both its real and imaginary parts are zero. Or, $z = x + iy = 0 \Leftrightarrow (x = 0) \wedge (y = 0), \forall z \in \mathbb{C}$.

Proof:

$$\text{assume that} \quad z = (x, y) = 0$$
$$\therefore z = x + iy = 0$$
$$\therefore x = -iy$$
$$\therefore x^2 = -y^2$$
$$\therefore x^2 + y^2 = 0$$
$$\therefore (x = 0) \wedge (y = 0).$$

(ii) Addition operation on complex numbers: $If z_1 = x_1 + iy_1, z_2 = x_2 = iy_2$ then $z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2), \forall z_1, z_2 \in \mathbb{C}$.

Proof:

$$\text{assume that} \quad z_1 = (x_1, y_1), z_2 = (x_2, y_2)$$
$$\therefore z_1 + z_2 = (x_1, y_1) + (x_2, y_2)$$
$$= (x_1 + iy_1) + (x_2 + iy_2)$$
$$= (x_1 + x_2) + i(y_1 + y_2).$$

(iii) Multiplication of complex numbers: $If z_1 = x_1 + iy_1, z_2 = x_2 = iy_2$ then $z_1 z_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 - y_1 x_2 y_2), \forall z_1, z_2 \in \mathbb{C}$.

Proof:

$$\text{assume that} \quad z_1 = (x_1, y_1), z_2 = (x_2, y_2)$$
$$\therefore z_1 z_2 = (x_1, y_1)(x_2, y_2)$$
$$= x_1(x_2 + iy_2) + iy_1(x_2 + iy_2)$$
$$= x_1 x_2 + i^2 y_1 y_2 + i(x_1 y_2 + x_2 y_1)$$
$$= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 - y_1 x_2 y_2) \in \mathbb{C}.$$

(iv) Two Complex numbers are equal if and only if the real and imaginary parts of the first number are equal to the corresponding real and imaginary parts of the second number. Or, $z_1 = z_2 \Leftrightarrow (x_1 = x_2) \wedge (y_1 = y_2); \forall z_1 = x_1 + iy_1, z_2 = x_2 + iy_2$.

Proof:

$$
\begin{aligned}
\text{assume that } & z_1 = (x_1, y_1), z_2 = (x_2, y_2) \\
& \because z_1 = z_2 \\
& \therefore x_1 + iy_1 = x_2 + iy_2 \\
& \therefore x_1 + iy_1 - x_2 - iy_2 = 0 \\
& \therefore (x_1 - x_2) + i(y_1 - y_2) = 0 + i0 \\
& \therefore (x_1 - x_2) = 0 \wedge (y_1 - y_2) = 0 \\
& \therefore (x_1 = x_2) \wedge (y_1 = y_2).
\end{aligned}
$$

(v) The product of multiplying a complex number by a real number is a complex number. Or, $az \in , \forall a \in \mathbb{R}, z \in \mathbb{C}$.

Proof:

$$
\begin{aligned}
\text{assume that } & a \in \mathbb{R}, z = (x + iy) \in \mathbb{C} \\
& \because a = (a + i0) \\
& \therefore az \\
& = (a + i0)(x + iy) \\
& = (ax - 0y_1) + i(ay - 0x) \\
& = (ax, ay) \\
& = ax + iay \in \mathbb{C}.
\end{aligned}
$$

(vi) As in real numbers, complex numbers have additive, associative, and distributive properties, and their proofs are left as exercises for the reader.

(vii) Complex numbers have the property of division. Or, $\frac{z_1}{z_2} \in \mathbb{C}, \forall z_1, z_2 \neq 0 \in \mathbb{C}$.

Proof:

$$\text{assume that } z_1 = (x_1, y_1), z_2 = (x_2, y_2)$$
$$\because \frac{z_1}{z_2}$$
$$= \frac{x_1x_2 + y_1y_2}{x^2 + y_2^2} + i\frac{x_2y_1 - x_1y_2}{x^2 + y_2^2} \in \mathbb{C}, x^2 + y_2^2 \neq 0$$
$$.$$

There are a number of distinct properties of complex numbers like adding and multiplying complex numbers. In addition, Euler's and De Moiver's formula on complex numbers, and their theorems on such types of numbers can be found in the foundations of mathematics. For more details, the reader can read Hamadameen's book (Hamadameen, 2022).

Example 2.1 (i) Write the folloing expressions in their simplest form.

(a) $-5(2+3i)+2(3-2i)+(7+5i)$.

(b) $(2,3)(2,-5)(3,2)$.

Solution: (a) $-5(2+3i)+2(3-2i)+(7+5i) = (3-14i)$.

(b) $(2,3)(2,-5)(3,2) = (2,3)[(2,-5)(3,2)] = (2,3)(6+10, 4-15) = (2,3)(16,-11) = (32+33, -22+42) = (65,26) = 65+26i$.

(ii) Find the values of the following.

(a) i^{13}.

(b) i^{28}.

(c) i^{2023}

Solution: (a) $i^{13} = i^{12}i = (i^2)^6 i = (-1)^6 i = (1)i = i$.

(b) $i^{28} = (i^2)^1 4 = (-1)^1 4 = 1$.

(c) $i^{2023} = i^{2022}i = (i^2)^{1011} i(-1)^{1011} i = (-1)i = -1$.

(iii) Solve $(3,4)^2 - 2(x,-y) = (x,y)$.

Solution:

$$(3,4)^2 - 2(x,-y) = (x,y)$$
$$\therefore (9-16, 12+12) + (-2x, 2y) = (x,y)$$
$$\therefore (-7,24) + (-2x,2y) = (x,y)$$
$$\therefore (-7-2x, 24+2y) = (x,y)$$
$$\therefore (x = -7-2x) \wedge (y = 24+2y)$$
$$\therefore x = -\frac{7}{3}, y = -24.$$

2.5 Exercises

Solve the following questions:

Q1: Reduce the following expressions to their simplest form:

(i) $(7-5i)(i-4)$.

(ii) $(a,5)(3,b)\forall a,b \in \mathbb{R}$.

(iii) $(2,3)(3,-3)(3,4) + (-3,7)(7,-6)$.

(iv) $\frac{(3,5)}{(0,2)}$.

(v) $\frac{(2,3)(3,-3)}{(3,0)}$.

(vi) $(8,-5)^2$.

Q2: Find the value of each x, and y in the following equations:

(i) $(7,-2)(x,y) = 3(x,-5y) + (-3,5)$.

(ii) $(2+5i)^2 - 3(x-iy) = (x,y)$.

(iii) $(2+5i)^4 - 3(x-iy) = (x,y)$.

(iv) $(\frac{2+i}{2-i})^2 + \frac{1}{2x+3iy} = 1+i$.

Q3: Find the value of x in $(1,3)+(3,-1)x=(1,2)(0,3)$.
Q4: If $z=(x,y)$ then prove that $\frac{1}{z}=(\frac{x}{x^2+y^2},\frac{-y}{x^2+y^2})$.
Q5: For all $z_1,z_2,z_3\in\mathbb{C}$ prove that:

(i) $z_1+(z_2+z_3)=(z_1+z_2)+z_3$.

(ii) $z_1\cdot(z_2\cdot z_3)=(z_1\cdot z_2)\cdot z_3$.

(iii) $z_1\cdot(z_2+z_3)=(z_1\cdot z_2)+(z_1\cdot z_3)$.

Q6: If $k\in\mathbb{Z}$ then prove that:

(i) $i^{4k+3}=-i$.

(ii) $i^{4k+2}=-1$.

(iii) $i^{4k+1}=i$.

(iv) $i^{4k}=1$.

Q7: Prove that $(\frac{1}{2}+\frac{\sqrt{3}}{2}i)^3=-1$
Q8: If $z,z_1,z_2\in\mathbb{C}$ then prove that:

(i) $R(iz)=-I(z)$.

(ii) $I(iz)=R(z)$.

(iii) $R(z_1+z_2)=R(z_1)+R(z_2)$.

(iv) $I(z_1+z_2)=I(z_1)+I(z_2)$.

Q9: If $\frac{x-iy}{x+iy}=a+ib$ then $a^2+b^2=1$.

2.6 Cartesian representation of a complex number

Real numbers can be represented by points on a straight line called the real number line, and complex numbers can be represented by points in the plane, where any number $z=(x,y)$ can be linked to the coordinates (x,y) of the Cartesian plane.

There is a one-to-one symmetric correspondence between the ordered pairs of real numbers and the points of the Cartesian coordinate plane. Thus, we can find unique symmetry between the complex numbers and the points in the coordinate plane. The real numbers are the ones that can be written in the form $(a, 0)$, the symmetry is a point $(a, 0)$ on the horizontal axis, which is called the real axis. Likewise, purely imaginary numbers, which can be written on the form $(0, b)$ corresponding to the point $(0, b)$ on the vertical axis which it is called the imaginary axis.

These two axes form together what is called the complex plane. It is noted here that while the unit on the real axis is equal to one, the unit on the imaginary axis is equal to $i = \sqrt{-1}$. Thus, the number (a, b) represents the point in the complex plane at a distance a from the real axis and at the distance $-b$ from the imaginary axis, and the following figure shows the Cartesian representation of some complex numbers, where the symbol of the real axis is R, while the imaginary axis with the symbol I. For example, points: $A(0, 0), B(3, 0), C(2, 3), D(-3, 0), E(0, -3), F(-2, -3), g(2, -3),$ and $H(0, 3)$ can be plotted in the complex cartesian as shown in Figure 2.2.

2.7 Adding and subtracting complex numbers

Since the sum of two complex numbers is a complex number whose real part is the sum of the two real parts and its imaginary part is the sum of its two imaginary parts, therefore, the sum represents the resulting new point which is the end of the diagonal of the parallelogram as shown in Figure 2.3, and in which $0z_1, 0z_2$ are adjacent sides and in the same method it is possible to subtract two complex numbers.

This method of addition indicates that the complex number can be considered as a vector at the Cartesian plane according to the rules of directional addition. Therefore, any number z can be represented either in the form $(x + iy)$ or in the form of an ordered pair (x, y) that corresponds to a point in the form of $\overrightarrow{0z}$, and the two components of the vector represent the real and imaginary part of the number. Furthermore, this representation is known as the Argand form (Wells, 2008; Jones, 2011; Flanigan, 1983).

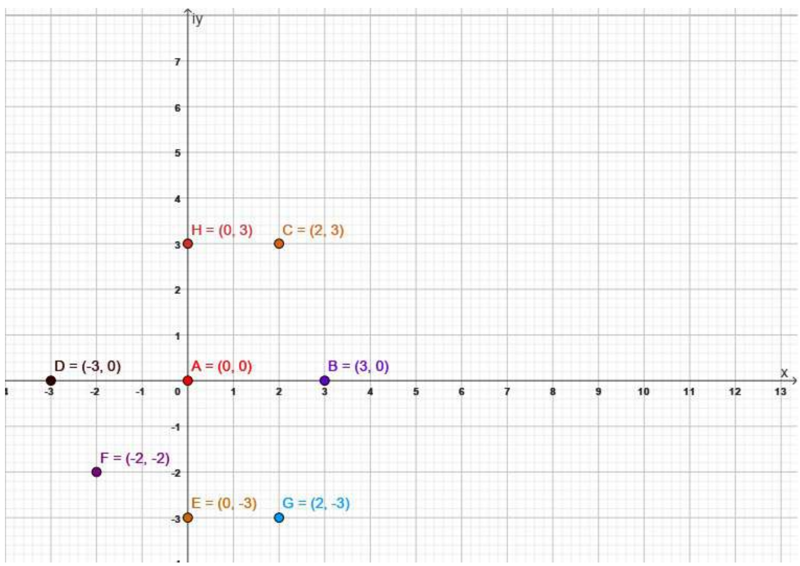


Figure 2.2: Cartesian representation of a complex number

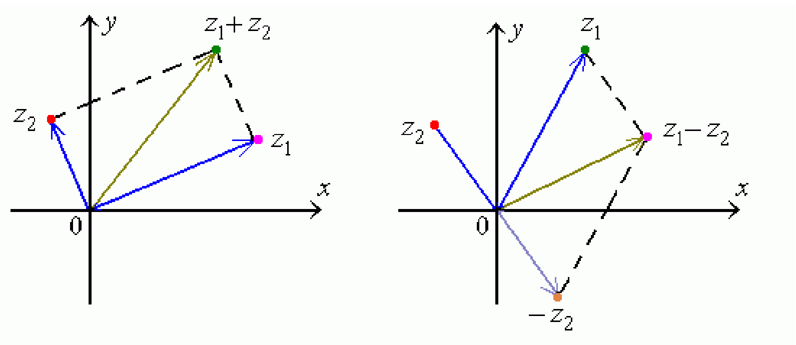


Figure 2.3: Adding and subtracting complex numbers

Example 2.2 if $z_1 = (3+i), z_2 = (-1+2i)$, then:

(1) $z_1 + z_2 = (3+i) + (-1+2i) = (3+i)$. (2) $z_1 - z_2 = (3+i) - (-1+2i) = (4-i)$.

2.8 Polar form of a complex number

The polar form of a complex number is another way to represent a complex number. The form$z = x + yi$ is called the rectangular coordinate form of a complex number.

The horizontal axis is the real axis and the vertical axis is the imaginary axis. We find the real and complex components in terms of r and ϕ where r is the length of the vector and ϕ is the angle made with the real axis, as shown in Figure 2.3.

From Pythagorean Theorem (Sparks, 2008), we have:

$r^2 = a^2 + b^2$.

On the other hand, by using the basic trigonometric ratios (Khan, 2020; Szetela, 1979):

$cos\phi = \frac{x}{r}, sin\phi = \frac{y}{r}$. Or, $rcos\phi = x, rsin\phi = y$, the rectangular form of a complex number is given by $z = x + iy$. Substitute the values of x, y, we get:

$z = x + iy = rcos\phi + i(rsin\phi) = r(cos\phi + isin\phi)$.

In the case of a complex number, r represents the absolute value or modulus and the angle ϕ is called the argument of the complex number. This can be summarized as follows:

The polar form of a complex number:

$z = x + iy = r(cos\phi + isin\phi)$, where $r = |z| = \sqrt{x^2 + y^2}$, $x = rcos\phi, y = rsin\phi$, and $\phi = tan^{-1}(\frac{y}{x}), \forall x > 0$, $\phi = tan^{-1}(\frac{y}{x}) + \pi, \forall x < 0$.

Note: (1) $sin(\phi + 2k\pi) = sin\phi, \forall k \in \mathbb{Z}$. (2) $cos(\phi + 2k\pi) = cos\phi, \forall k \in \mathbb{Z}$.

Example 2.3 Express each of $z_1 = (3, -3), z_2 = (-5, 0)$ by polar form.

Solution: $r = \sqrt{3^2 + (-3)^2} = 3\sqrt{2}$.

$tan\phi = \frac{-3}{3} = -1$.

Therefore, $\phi = 135°$, or $\phi = 315°$.

Thus, $z_1 = (3, -3) = 3\sqrt{2}(cos135°, sin135°)$, or $z_1 = (3, -3) = 3\sqrt{2}(cos315°, sin315°)$.

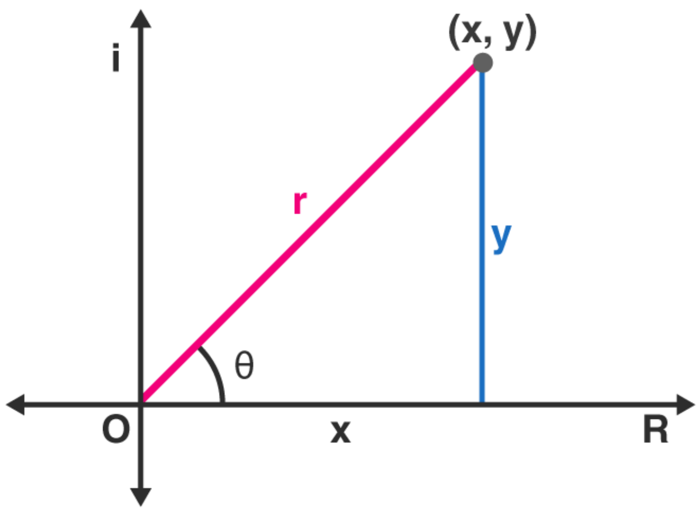


Figure 2.4: The polar form of a complex number

By the same way, $z_2 = (-5, 0) = 5(cos180^\circ + sin180^\circ)$.

2.9 Exercises

Find the polar form of the following complex numbers:

Q1: (1) $-7i$. (2) i. (3) -8.
Q2: (1) $(-\frac{1}{2} + \frac{\sqrt{3}}{2}i)$. (2) $(\frac{1}{2} + \frac{\sqrt{3}}{2}i)$. (3) $(-1 + i)$.
Q3: (1) $(-5 + 3i)$. (2) $(1 - \sqrt{3} + (1 + \sqrt{3})i)$. (3) $(\sqrt{5} - i)$.
Q4: (1) $\sqrt{2} + i$. (2) $\sqrt{2} - i$. (3) $()3 + cos\alpha + isin\alpha)$.
Q5: $(cos\alpha + sin\gamma + i(cos\alpha + sin\gamma))$.
Q6: Find the Cartesian form for the following complex numbers:
(1) $\frac{3}{2}(sin\frac{pi}{4}, cos\frac{pi}{4})$.
(2) $4\sqrt{3}(cos30^\circ, sin30^\circ)$.
(3) $6(cos270^\circ, sin270^\circ)$.
(4) $-2(cos180^\circ, sin180^\circ)$.
(5) $(-cos180^\circ, sin180^\circ)$.
(6) $(-cos180^\circ, sin180^\circ)$.
(7) $\frac{2+\sqrt{3}}{2}(cos30^\circ, sin90^\circ)$.

2.10 Products and quotients of complex numbers in polar form

The complex numbers in the polar system can be multiplied and divided based on the following two theorems with the help of the following lemma to prove those theorems (Phillips, 2020; Ledermann, 2013a; George, 1967; Spiegel et al., 2009; Priestley, 2003).

Lemma:

(i) $cos(\alpha \pm \beta) = cos\alpha cos\beta \pm sin\alpha sin\beta$.

(ii) $sin(\alpha \pm \beta) = sin\alpha cos\beta \pm cos\alpha sin\beta$.

(iii) $cos^2\alpha + sin^2\alpha = 1 = (1, 0) = (cos^2\alpha + sin^2\alpha, 0)$.

Theorem 2.1 *The absolute value of a product of two complex numbers is equal to the product of the absolute values, and the argument of the* *product of two complex numbers is equal to the sum of their arguments. Or,*

$$|z_1 z_2| = r_1 r_2$$
$$arg(z_1 z_2) = arg(z_1) + arg(z_2)$$

Proof

$$\text{Suppose that } z_1 = r_1(cos\phi_1, sin\phi_1), z_2 = r_2(cos\phi_2, sin\phi_2).$$
$$\therefore z_1 z_2 = r_1 r_2(cos\phi_1 cos\phi_2 - sin\phi_1 sin\phi_2, sin\phi_2 cos\phi_1 + cos\phi_2 sin\phi_1)$$
$$= r_1 r_2(cos(\phi_1 + \phi_2), sin(\phi_1 + \phi_2))$$
$$\therefore |z_1 z_2| = r_1 r_2 \left|(cos(\phi_1 + \phi_2), sin(\phi_1 + \phi_2))\right|$$
$$\therefore r_1 r_2 \sqrt{cos^2(\phi_1 + \phi_2) + sin^2(\phi_1 + \phi_2)}$$
$$= r_1 r_2$$
$$\therefore arg(z_1, z_2) = \phi_1 + \phi_2$$
$$= arg(z_1) + arg(z_1). \ \blacklozenge$$

Example 2.4 Find $z_1 z_2$ of $z_1 = 3(cos120°, sin120°), z_2 = 5(cos150°, sin150°)$.

Solution: $z_1 z_2 = 15(cos(120° + 150°), sin(120° + 150°) = 15(0, -1) = (0, -15)$.

$= 15(0, -1) = (0, -15)$.

Theorem 2.2 *The absolute value of a quotient of two complex numbers is equal to the quotient of the absolute values, and the argument of the quotient of two complex numbers is equal to difference between the arguments . Or,*

$$\left|\frac{z_1}{z_2}\right| = \frac{r_1}{r_2}$$
$$arg(\frac{z_1}{z_2}) = arg(z_1) - arg(z_2)$$

Proof

Suppose that $z_1 = r_1(cos\phi_1 + isin\phi_1), z_2 = r_2(cos\phi_2 + isin\phi_2)$.

$$\therefore \frac{z_1}{z_2} = \frac{r_1(cos\phi_1 + isin\phi_1)}{r_2(cos\phi_2 + isin\phi_2)},$$

$$= \frac{r_1}{r_2} \cdot \frac{(cos\phi_1, sin\phi_1)}{(cos\phi_2, sin\phi_2)} \cdot \frac{(cos\phi_2, -sin\phi_2)}{(cos\phi_2, -sin\phi_2)},$$

$$= \frac{r_1}{r_2} \cdot \frac{(cos\phi_1, sin\phi_1)(cos\phi_2, -sin\phi_2)}{(cos^2\phi_2 + sin^2\phi_2, 0)},$$

$$= \frac{r_1}{r_2}(cos\phi_1 cos\phi_2 + sin\phi_1 sin\phi_2 - (cos\phi_1 sin\phi_2 + sin\phi_1 cos\phi_2))$$

$$= \frac{r_1}{r_2}(cos(\phi_1 - \phi_2), sin(\phi_1 - \phi_2)).$$

$$\therefore \left|\frac{Z_1}{z_2}\right| = \frac{r_1}{r_2}\left|(cos(\phi_1 - \phi_2), sin(\phi_1 - \phi_2))\right|.$$

$$= \frac{r_1}{r_2}.$$

$$\therefore arg(\frac{z_1}{z_2}) = \phi_1 - \phi_2,$$

$$= arg(z_1) - arg(z_2). \blacklozenge$$

Example 2.5 If $z_1 = 7(cos95°, sin95°), z_2 = 9(cos65°, sin65°)$ then find $\frac{z_1}{z_2}$.

Solution: $\frac{z_1}{z_2} = \frac{7}{9}(cos30°, sin30°) = \frac{7}{9}(\frac{\sqrt{3}}{2}, \frac{1}{2}) = (\frac{7\sqrt{3}}{18}, \frac{7}{18})$.

2.11 De Moivre's theorem

De Moiver's theorem is one of the most important mathematical theories in the development of analytical geometry. The correct formula for the correct image pattern, and how to visualize it, in addition to using the well-known theory, can be found in the studies (Lial, 2016; Mukhopadhyay, 2006; Brand, 1942).

As mathematics is not just a part of our lives, it is our entire life, it is an integral part of our day, every detail in your life is a set of data, and everything that happens to us is a result of these data, so if we stop for a moment and look around us, we will find that the world is going according to the rules of mathematics and for this studying this theory and getting to know it is very important (Schaumberger, 1968).

Theorem 2.3 *Let* $z = r(cos\phi, isin\phi)$ *be a complex number, and* n *any integer, then* $z^n = r^n(cos(n\phi), isin(n\phi))$.

Proof We will prove the formula by mathematical induction as follows:

$$\text{(1) If } \ n = 1, then$$
$$z^1 = r^1(cos(1\phi), isin(1\phi))$$
$$= r(cos\phi, isin\phi)$$
$$\therefore \text{the formula is true where} n = 1.$$
$$\text{(2) Suppose that the formula is true where } n = k,$$
$$\therefore z^k = r^k(cos(k\phi), isin(k\phi))...(2).$$
$$\text{(3) Now, we have to prove where } n = k + 1.$$
$$\text{Now, multiplying}(2) by (1), \text{we get :}$$
$$z^{k+1} = r^k(cos(k\phi), isin(k\phi))r(cos\phi, isin\phi)$$
$$= (r(cos\phi, isin\phi))^{k+1}$$
$$= r^{k+1}(cos(k + 1)\phi, sin(k + 1)\phi)$$
$$\therefore \text{the theorem is true for all } n \in \mathbb{N}. \ \blacklozenge$$

Corollary *If r= 1, then the De Moivre's theorem became:*
$(cos\phi, isin\phi)^n = (cos(n\phi), (n\phi))$.

Proof The proof has been left as an exercise for a reader. ♦

Example 2.6 Use the corollary of De Moivre's theorem to evaluate $(-\sqrt{3}, 1)^9$.

Solution: The polar form of

$$(-\sqrt{3},1) = 2(cos150°, sin150°).$$
$$\therefore (-\sqrt{3},1)^9 = (2(cos150°, sin150°))^9$$
$$= 2^9(cos9(150°), sin9(150°))$$
$$= 512(cos1350°, sin1350°)$$
$$= 512(0,-1)$$
$$= (0,-512).$$

2.12 Exercises

Solve the following questions:

Q1: Write the result in the algebraic form $x + iy$ for the following:

(i) $2(cos30°, sin30°) \cdot 4(cos45°, sin45°)$.

(ii) $4(cos50°, sin50°) \cdot 5(cos250°, sin250°)$.

(iii) $12(cos245°, sin245°)/3(cos20°, sin20°)$.

(iv) $(-\sqrt{3},1)/(cos150°, sin150°)$.

Q2: Convert the following expressions to the polar form, then find the product and verify the validity of the result by finding the product algebraically:

(i) $(2,-2) \cdot (\sqrt{2},1)$.

(ii) $(1,-\sqrt{3}) \cdot (-2\sqrt{3},-2)$.

(iii) $(4,4) \cdot (-2,-2)$.

Q3: Convert the following expressions to the algebraic form, then find the product and verify the validity of the result by finding the product in the polar form:

(i) $5(cos330°, sin330°) \cdot 4(cos210°, sin210°)$.

(ii) $2(cos30°, sin30°) \cdot 6(cos240°, sin240°)$.

(iii) $4(cos245°, sin245°) \cdot 3(cos20°, sin20°)$.

(iv) $-2(cos180°, sin180°) \cdot 3(cos150°, sin150°)$.

Q4: Use De Moivre's theorem and its corollary to find the result of the following expressions. And, express the results algebraically.

(i) $9(cos30°, sin30°)^3$.

(ii) $\sqrt{3}(cos45°, sin45°)^7$.

(iii) $\sqrt{2}(cos50°, sin50°)^6$.

Q5: If $z = (1, 1)$ what is the result of z^2, z^3?

Q6: By using the polar form prove that if the product of two complex numbers is zero then at least one of these numbers is a zero.

Q7: Prove that, if $z = r(cos\phi, sin\phi)$ then $\frac{1}{z} = \frac{1}{r}(cos\phi, -sin\phi)$.

Q8: Simplify each of:

(i) $(\frac{(1+sin\phi,cos\phi)}{(1+sin\phi,-cos\phi)})^n$.

(ii) $(sin\alpha - sin\beta, cos\alpha - cos\beta)^n$.

Q9: Prove that $(\sqrt{3} + i)^n + (\sqrt{3} - i)^n = 2^{n+1}cos\frac{n\pi}{6}$; $\forall n \in \mathbb{Z}^+$.

Q10: If $sin\alpha + sin\beta + sin\gamma = cos\alpha + cos\beta + cos\gamma$ then prove that $sin3\alpha + sin3\beta + sin3\gamma = 3sin(\alpha + \beta + \gamma)$.

2.13 Conjugate number

Each complex number has the conjugate number. Or, a conjugate of a complex number is another complex number which has the same real part as the original complex number and the imaginary part has the same magnitude but opposite sign. If we multiply a complex number with its conjugate, we get a real number, defined as follows;

Definition 2.2 Defined a conjugate of the complex number $z = (x, y) = x + iy$ is another complex number $\bar{z} = (x, -y) = x - iy$ (Andreescu and Andrica, 2006; Hahn, 1994; Ledermann, 2013b; Schwerdtfeger, 2020; Sikka, 2017), as shown in the Figure 2.2.

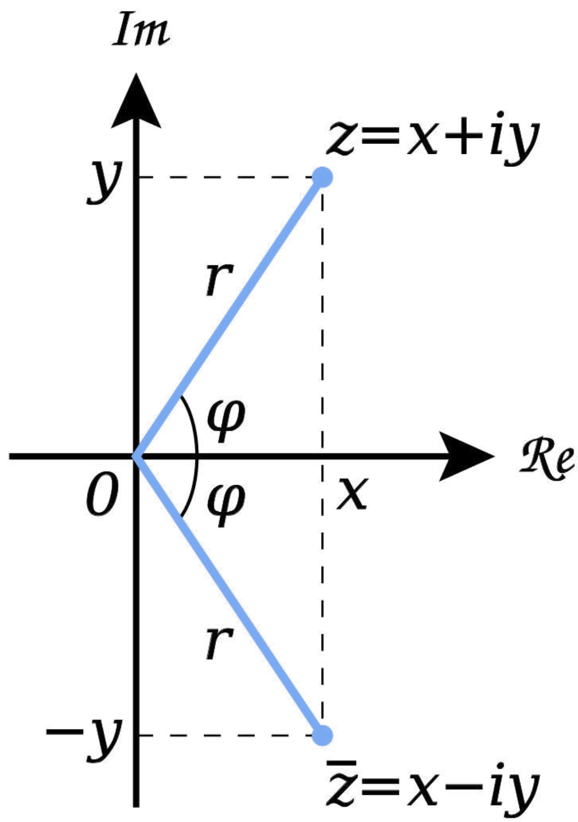


Figure 2.5: Conjugate of complex number

2.14 Properties of complex numbers and their conjugates

In this section, we will study some of the features and characteristics of complex numbers and their conjugates based on the results obtained from some researchers' work in the field of complex numbers (Bourbaki, 1994; Hamadameen, 2022; Axler, 2010; Spiegel et al., 2009; Brown and Churchill, 2009; Apostol, 1974; Apostol and Ablow, 1958; Apostol, 1981).

(i) Conjugate is a factor of the factors of differences of squares of the complex numbers. Mathematically, $(x+iy)(x-iy) = x^2 - i^2y^2 = x^2+y^2$. For example, $5 = 2+3 = 2-3i^2 = (\sqrt{2}-\sqrt{3}i)(\sqrt{2}+\sqrt{3}i)$.

(ii) The sum of any complex number and its conjugate is the real number and is equal to twice the real part of the complex number. Mathematically, $z+\bar{z} = (x,y)+(x,-y) = x+iy+x-iy = 2x = (2x,0), \forall z, \bar{z} \in \mathbb{C}$. For example, $(3+5i)+(3-5i) = 6 = 2(3)$.

(iii) The result of subtracting of the conjugate from a complex number is an imaginary number, and is equal to twice the imaginary part. Mathematically, $z-\bar{z} = (x,y)-(x,-y) = x+iy-x+iy = 2iy = (0,2iy), \forall z, \bar{z} \in \mathbb{C}$. For example, $(\sqrt{5}-\sqrt{7}i)-(\sqrt{5}+\sqrt{7}i) = \sqrt{5}-\sqrt{7}i-\sqrt{5}-\sqrt{7}i = -2\sqrt{7}i$.

(iv) The product of multiplying any complex number by its conjugate is a real number and is called the square of the absolute value of a complex number and is equal to the sum of the square of the real part and the square of the imaginary part. Or, The product is called the square of the length of the complex number or the square of its absolute value.

Proof:

$$\text{assume that } z = (x, y)$$
$$\therefore \bar{z} = (x, -y)$$
$$\text{Also the product} = z\bar{z} = (x^2 + y^2, xy - xy)$$
$$= (x^2 + y^2, 0)$$
$$= x^2 + y^2 + 0i$$
$$= x^2 + y^2$$
$$= r^2$$
$$\therefore z\bar{z} = \bar{z}z = |z|^2$$
$$\therefore r = |z| = \sqrt{z\bar{z}}.$$

See the example in (i).

(v) (a) The conjugate of an algebraic sum is the algebraic sum of conjugates. Or, $\overline{z_1 + z_z} = \bar{z_1} + \bar{z}_2$.

Proof:

$$\text{Assume that } z_1 = (x_1, y_1), z_2 = (x_2, y_2)$$
$$\therefore \bar{z}_1 = (x_1, -y_1), \bar{z}_2 = (x_2, -y_2)$$
$$z_1 + z_2 = (x_1 + x_2, y_1 + y_2)$$
$$\text{L. H. S. is } \overline{z_1 + z_2} = (x_1 + x_2, -y_1 - y_2)$$
$$= (x_1, -y_1) + (x_1, -y_1)$$
$$= \bar{z_1} + \bar{z}_2 ...(1).$$
$$\text{R. H. S. is } \bar{z_1} + \bar{z}_2 = (x1, -y1) + (x_2, -y_2)$$
$$= (x_1 + x_2, -y_1 - y_2)$$
$$= (x_1 + x_2, -y_1 - y_2)$$
$$= \overline{z_1 + z_2} ...(2)$$
$$\because (1) = (2)$$
$$\therefore \overline{z_1 + z_2} = \bar{z_1} + \bar{z}_2.$$

(b) The conjugate of an algebraic subtraction is the algebraic subtraction of conjugates. Or, $\overline{z_1 - z_2} = \bar{z_1} - \bar{z}_2$.

Proof:

The same method of (a).

(vi) Conjugate of the conjugate of a complex number is the complex number itself. Or, $\bar{\bar{z}} = z$.

Proof:

$$\begin{aligned} \text{Assume that } \; z &= (x, y) \\ \therefore \bar{z} &= (x, -y) \\ \therefore \bar{\bar{z}} &= (x, -(-y)) \\ &= (x, y) \\ &= z. \end{aligned}$$

(vii) (a) The conjugate of multiplication of complex numbers is the multiplication of conjugates of them. Or, $\overline{z_1 \cdot z_2} = \bar{z_1} \cdot \bar{z_2}$.
Proof:

$$\begin{aligned} \text{Assume that } \; z_1 &= (x_1, y_1), z_2 = (x_2, y_2) \\ \therefore \bar{z_1} &= (x_1, -y_1), \bar{z_2} = (x_2, -y_2) \\ \therefore z_1 \cdot z_2 &= (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2) \\ \therefore \overline{z_1 \cdot z_2} &= (x_1x_2 - y_1y_2, -x_1y_2 - y_1x_2)...(1). \\ \because \bar{z_1} \cdot \bar{z_2} &= (x_1, -y_1) \cdot (x_2, -y_2) \\ &= (x_1x_2 - y_1y_2, -x_1y_2 - y_1x_2)...(2). \\ \because (1) &= (2) \\ \therefore \overline{z_1 \cdot z_2} &= \bar{z_1} \cdot \bar{z_2}. \end{aligned}$$

(b) The conjugate of division of complex numbers is the division of conjugates of them. Or, $\overline{(\frac{z_1}{z_2})} = \frac{\bar{z_1}}{\bar{z_2}}$.
Proof:

By using (a), we have $\overline{(\frac{z_1}{z_2})}, \frac{\bar{z_1}}{\bar{z_2}}$

$$\therefore \bar{z_2} \cdot \overline{(\frac{z_1}{z_2})}; \bar{z_2} \neq 0$$
$$= \overline{(\frac{z_2 \cdot z_1}{z_2})}$$
$$= \bar{z_1} ...(1).$$
$$\because \bar{z_2} \cdot \frac{\bar{z_1}}{\bar{z_2}}$$
$$= \bar{z_1} ...(2).$$
$$\because (1) = (2)$$
$$\therefore \overline{(\frac{z_1}{z_2})} = \frac{\bar{z_1}}{\bar{z_2}}.$$

(viii) (a) The product of the absolute value of two complex numbers is equal to the product of their absolute values. Or, $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$.

Proof:

By using (iv), we can prove (a), as follows:

$$|z_1 \cdot z_2|^2 = (z_1 \cdot z_2)\overline{(z_1 \cdot z_2)}$$
$$= (z_1 \cdot z_2)(\bar{z_1} \cdot \bar{z_2})$$
$$= (z_1 \cdot \bar{z_1})(z_2 \cdot \bar{z_2})$$
$$= |z_1|^2 \cdot |z_2|^2$$
$$\therefore |z_1 \cdot z_2| = |z_1| \cdot |z_2| .$$

(b) The division of the absolute value of two complex numbers is equal to the division of their absolute values. Or, $\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}$.

Proof:

We prove this claim by using (a) as follows :

$$|z_2|\left|\frac{z_1}{z_2}\right|$$
$$= \left|z_2 \cdot \frac{z_1}{z_2}\right|$$
$$= |z_1| \,...(1).$$
$$\because |z_2|\,\frac{|z_1|}{|z_2|}$$
$$= |z_1| \,...(2).$$
$$\because (1) = (2)$$
$$\therefore \left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}.$$

(ix) (a) $\forall z \in \mathbb{C}, |z| = |\bar{z}|$.

Proof:

Suppose that :

$$z = (x, y)$$
$$\therefore |z| = |(x, y)|$$
$$= \sqrt{x^2 + y^2}$$
$$= \sqrt{x^2 + (-y)^2}$$
$$= |(x, -y)|$$
$$= |\bar{z}|\,.$$

(b) $-|\bar{z}| \leq R(z) \leq |z|$.

Proof:

$$\because y^2 \geq 0$$
$$\therefore -\sqrt{x^2 + y^2} \leq x \leq \sqrt{x^2 + y^2}$$
$$\therefore -|\bar{z}| \leq R(z) \leq |z|\,.$$

(c) $-|\bar{z}| \leq I(z) \leq |z|$.

Proof:

$$\because x^2 \geq 0$$
$$\therefore -\sqrt{x^2+y^2} \leq y \leq \sqrt{x^2+y^2}$$
$$\therefore -|\bar{z}| \leq I(z) \leq |z|\,.$$

2.15 Exercises

Solve the following questions:

Q1: Write the following algebraic expressions in simplest form:

(i) $\frac{(5,5)}{(3,4)} + \frac{(5,-5)}{(3,-4)}$.

(ii) $\frac{2i}{1+i} - \frac{2+i}{2-i}$.

(iii) $\frac{(2,36)}{6,8} + \frac{(7,26)}{(3,-2)}$.

(iv) $\frac{sin\alpha + icos\alpha}{cos\alpha - isin\alpha}$.

(v) $\frac{(a,b)^2}{(a,-b)^2} - \frac{(a,-b)^2}{(a,b)^2}, \forall a,b \in \mathbb{R}$.

(vi) $\frac{(\sqrt{2}+i)^3(2-i)^4}{(2+3i)^4}$.

Q2: If $z \in \mathbb{C}$ then prove that $\overline{\bar{z}+5i} = z - 5i$.

Q3: Find the absolute values of the following complex numbers:

(i) $cos2\beta + isin2\beta$.

(ii) $\frac{(2,-3)(3,-2)}{(3,4)(4,3)}$.

(iii) $\frac{1+\sqrt{3}+i}{1-i}$.

(iv) $\frac{1-\sqrt{3}+2i}{1+i}$.

Q4: If $\frac{(2,2)^2}{(2,-2)^2} + \frac{4}{(x,y)} = (4,4)$ then find the values of x, y.

Q5: Find the number of complex numbers which are conjugate of their own cube?

Q6: If $(x+iy)^4 = a+ib$ then prove that $a^2+b^2 = (x^2+y^2)^4$.

Q7: Find each of x, y of the equation: $(1+i)x + 2(1-2i)y = 3$.

Q8: For each complex number z prove that:

(i) $R(z) = 1/2(z + \overline{z})$.

(ii) $I(z) = 1/2i(z - \overline{z})$.

(iii) $arg(\overline{z}) = -arg(z)$.

Q9: $\overline{(\frac{z_1}{z_2})} = arg(\overline{z_1}) - arg(\overline{z_2})(mod 2\pi)$.

2.16 Absolute value inequalities

We already knew in the previous section that the absolute value of the product of two complex numbers is equal to the product of their absolute values. But the matter is different with regard to the absolute value of the sum of two complex numbers because it depends on the following theorem (Khamsi and Kirk, 2011; Jacobs and Geometry, 1974; Abramowitz et al., 1988; Abramowitz and Stegun, 1948; Apostol, 1967; Krantz et al., 1999).

Theorem 2.4 *the absolute value of the addition of two complex numbers is less than or equal to the addition of their absolute values. Or,* $|z_1 + z_2| \leq |z_1| + |z_2|, \forall z_1, z_2 \in \mathbb{C}$.

Proof

$$\text{Let us consider } z_1 = (x_1, y_1), z_2 = (x_2, y_2)$$
$$\therefore |z_1 + z_2|^2 = (z_1 + z_2)\overline{(z_1 + z_2)}$$
$$= (z_1 + z_2) + (\overline{z_1} + \overline{z_1})$$
$$= z_1\overline{z_1} + z_2\overline{z_2} + z_1\overline{z_2} + z_2\overline{z_1}$$
$$= |z_1|^2 + |z_2|^2 + (z_1\overline{z_2} + z_2\overline{z_1})...(1)$$
$$\because z_2\overline{z_1} \text{ is the conjugate of } z_1\overline{z_2},$$
$$\therefore z_1\overline{z_2} + z_2\overline{z_1}$$
$$= 2(x_1x_2 + y_2y_1)$$
$$= 2R(z_1\overline{z_2}) \leq 2|z_1\overline{z_2}|$$
$$\leq 2|z_1||\overline{z_2}|$$
$$= 2|z_1||z_2|, (|\overline{z_2}| = |z_2|)$$
By substituting in (1), we get;
$$|z_1 + z_2|^2 \leq |z_1|^2 + |z_2|^2 + 2|z_1||z_2|$$
$$= |z_1 + z_2|^2 \leq (|z_1| + |z_2|)^2$$
$$\therefore |z_1 + z_2| \leq |z_1| + |z_2|. \; \blacklozenge$$

Theorem 2.5 *the absolute value of the addition of any numbers of complex numbers is less than or equal to the addition of their absolute values. Or,* $|z_1 + z_2 + ...z_n| \leq |z_1|+|z_2|+|z_3|+...+|z_n|, \forall z_1, z_2, ...z_n \in \mathbb{C}$.

Proof

We prove this theorm by mathematical induction.

(1)if $n = 1$,

then $|z_1| \leq |z_1|$

$\therefore$ the theorem is true.

(2)Suppose that the theorem is true when$n = k$.

Or, $|z_1 + z_2 + ...z_k| \leq |z_1| + |z_2| + |z_3| + ... + |z_k|, \forall z_1, z_2, ...z_k \in \mathbb{C}$.

(3)Now, we have to prove that the theorem is true if $n = k + 1$.

$$|z_1 + z_2 + ...z_k + z_{k+1}| \leq |(z_1 + z_2 + ...z_k) + z_{k+1}|$$
$$\leq |(z_1 + z_2 + ...z_k)| + |z_{k+1}|$$
$$\leq |z_1| + |z_2| + ... + |z_{k+1}|.$$

$\therefore$ the theorem is true where $n = k + 1$.

Thus, the theorem is true for all $n \in \mathbb{N}$. ♦

Example 2.7 If $|z| \leq 2$ then prove that the maximum value of the term $|z^2| + 2$ is 6.

Solution: $|z^2 + 2| \leq |z^2| + |2^2| = 2^2 + 2 = 6$.

2.17 Exercises

Solve the following questions:

Q1: Prove that $|z_1 - z_2| \geq |z_1| - |z_2|$; (Hint: $z_1 = z_2 + (z_1 - z_2)$).

Q2: If $|z| \leq 2$ what is the maximum value of $|z^3 + z^2 + z + 1|$?

Q3: Find the locus of the equation $|z - 3| + |z + 3| = 10$.

Q4: Prove that $|z_1 + z_2|^2 + |z_1 - z_2|^2 = 2(|z_1|^2) + |z_2|^2$. 4

Q5: If $|z_2 - z_1|^2 = |z_2 - z_0|^2 + |z_1 - z_0|^2$ then $z_2 - z_0 = i\lambda(z_1 - z_0), \lambda \in \mathbb{R}$.

Q6: If $z_i, w_i \in \mathbb{C}, i = 1, 2, ..., n$ then prove that:

$|\sum_i^n z_i w_i| \leq \sqrt{\sum_i^n |z_i|^2}\sqrt{\sum_i^n |w_i|^2}$.

2.18 Square root of complex number

Finding the square root of the complex number z is equivalent to finding the solution of the equation:

$$w^2 = z \tag{2.2}$$

To solve (2.2), we suppose that:

$$w = x + iy$$
$$z = a + ib$$

Or,

$$(x + iy)2 = a + iy$$
$$x^2 - y^2 - 2ixy = a + ib$$

By equating the real and imaginary partials in the two equations, we find that:

$$\begin{aligned} x^2 - y^2 &= a \\ 2xy &= b \end{aligned} \tag{2.3}$$

Now, the aim is to find the values of x, y in terms of a, b. By squaring both sides of the equation (2.3), adding them, and simplifying the results, we get:

$$\begin{aligned} x^2 - y^2 &= \sqrt{a^2 + b^2} \\ 2xy &= b \end{aligned} \tag{2.4}$$

The equation (2.4) can be written as:

$$\begin{aligned} x^2 &= \frac{a + \sqrt{a^2 + b^2}}{2} \\ y^2 &= \frac{-a + \sqrt{a^2 + b^2}}{2} \end{aligned} \tag{2.5}$$

Or,

$$\begin{aligned} x &= \mp\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}} \\ y &= \mp\sqrt{\frac{-a+\sqrt{a^2+b^2}}{2}} \end{aligned} \tag{2.6}$$

Thus, (2.3) has the solution $\mp(x+iy)$, where the values of x, y can be requested in (2.6).

Example 2.8 Find the square root of complex number $9+40i$. (Hint: To find the square root of a complex number, we will assume the root to be $a+ib$. Then we can compare it with the original number to find the values of a and b, which will give us the square root).

Solution:

$$\text{Let } 9+40i=(a+ib)^2$$

We will simplify this equation by proceeding as

$$9+40i=a^2+(ib)^2+2(ab)i$$

$$9+40i=a^2-b^2+2(ab)i$$

By comparing the real and imaginary parts :

$$a^2-b^2=9...(1)$$

$$2ab=40...(2)$$

Using equation (1) we can write that :

$$a^2=9+b^2$$

$$\Rightarrow a=\mp\sqrt{9+b^2}$$

Substituting this value in (2) :

$$\mp 2(\sqrt{9+b^2}b=40$$

$$\Rightarrow \mp b\sqrt{9+b^2}=20$$

Squaring both sides of the equation,

$$b^2(9+b^2)=400$$

$$\Rightarrow b^4+9b^2-400=0$$

$$\Rightarrow b^4+25bb^2-16b^2-400=0$$

$$\Rightarrow (b^2-16)(b^2+25)=0$$

$$\Rightarrow b^2=16 \vee b^2=-25(\text{rejected because square cannot be negative}).$$

$$\Rightarrow b=4,-4$$

Substituting the value of b in(1) :

$$\therefore a^2=9$$

$$\Rightarrow a^2=25$$

$$\Rightarrow a=\mp 5$$

$$\therefore \sqrt{9+40i}=\begin{cases} 5+4i \\ -5-4i \end{cases}$$

Example 2.9 Find the square root of $-4-3i$.

Solution: The solution steps have been left as an exercise to the reader , and the result is $\mp\frac{1}{\sqrt{2}}(3+i)$.

2.19 Exercises

Solve the following questions:

Q1: Find the square roots of each of:

(i) i.

(ii) $-3i$.

(iii) $-\frac{1}{2} + \frac{\sqrt{3}}{2}i$

(iv) $a^2 - 1 + 2ai, a \in \mathbb{R}$.

(v) $-21 - 20i$.

(vi) $1 + 4\sqrt{3}i$.

(vii) $-8 - 6i$.

(viii) $3 + 4i$.

Q2: Solve the following equations:

(i) $x^2 + x + 1$.

(ii) $x^2 - \frac{3}{2}x + 1$.

(iii) $x^4 - 1$.

(iv) $x^5 = x$.

(v) $x^4 - 1 - \frac{1}{\sqrt{3}}i$.

(vi) $(1 - i)x^2 + (11 + 9i)x - 20 + 8i = 0$.

(vii) $x^5 + x^3 = 0$.

2.20 Roots of complex numbers

An nth root of a number x is a number r which, when raised to the power n, yields x:

$r^n = x$, where n is a positive integer, sometimes called the degree of the root. A root of degree 2 is called a square root and a root of degree 3, a cube root. Roots of higher degree are referred by using ordinal numbers, as in fourth root, twentieth root,...etc. The computation of an nth root is a root extraction.

Any non-zero number considered as a complex number has n different complex nth roots, including the real ones (at most two). The nth root of 0 is zero for all positive integers n, since $0^n = 0$.

Definition 2.3 If z, w are complex numbers, and $w^n = z$ then w is called the nth root of z (Andreescu and Andrica, 2006; Bak et al., 2010).

Theorem 2.6 *For all nonzero complex number $z = r(cos\phi + isin\phi)$ has nth root in the form:*

$$\sqrt[n]{z} = \sqrt[n]{r}(cos(\frac{\phi + 2\pi k}{n}) + isin(\frac{\phi + 2\pi k}{n})), k = 0, 1, ..., n - 1 \quad (2.7)$$

Proof

By using De Moivre's theorem, we get:

$$\sqrt[n]{z} = (\sqrt[n]{r}(cos(\frac{\phi + 2\pi k}{n}) + isin(\frac{\phi + 2\pi k}{n})))^n$$
$$= r(cos(\phi + 2k\pi) + isin(\phi + 2k\pi))$$
$$= r(cos\phi + isin\phi)$$

$\therefore \forall n \in \mathbb{Z}^+$(2.7) is the (n)th roof for Z.

Now, we are going to prove that for all different values of,
k there are different roots.

Since two complex numbers are equal if their absolute values,
and arguments are equal.

Or if the difference between the two arguments is a multiple of 2.

Suppose that k_1, k_2 two different values of K,
such that $(0 \leq k \leq n-1) \wedge (k_1 < k_2)$.

Obviously, $(\frac{\phi + 2\pi k_1}{n}) \wedge (\frac{\phi + 2\pi k_2}{n})$, are not equal.

Because the difference between them is $(\frac{2\pi(k_2 - k_1}{n})$.

Morever, the difference is less than 2π.

Thus, if w is the nth roor for z, then$w$ is the solution for (2.7)

$\therefore w^n - z = 0,$ and this equation has at most n roots.

$\therefore z = r(cos\phi + isin\phi)$ has exactly n roots. ♦

Example 2.10 Find $\sqrt[3]{8i}$.

Solution:

$$\because 8i = (0,8) = 8(cos90°, sin90°),$$
$$\therefore \sqrt[3]{8i} = \sqrt[3]{(0,8)} = \sqrt[3]{8}(cos\frac{90+2k\pi}{3}, sin\frac{90+2k\pi}{3}).$$
$$= 2(cos30° + k(30°), sin30° + k(30°)), k = 0,1,2.$$
$$k = 0 \Rightarrow r_1 = 2(cos30°, cos30°)$$
$$= (\sqrt{3}, 1)$$
$$= \sqrt{3} + i.$$
$$k = 1 \Rightarrow r_2 = 2(cos150°, cos150°)$$
$$= (-\sqrt{3}, 1)$$
$$= -\sqrt{3} + i.$$
$$k = 2 \Rightarrow r_2 = 2(cos270°, cos270°)$$
$$= (0, -2)$$
$$= -2i.$$

Note: If $k = 3$ then $r_3 = r_1$.

2.21 Exercises

Solve the following problems:

Q1: Express the following roots in polar forms:

(i) $x^4 = -16i$.

(ii) $x^3 = 1 - i$.

(iii) $x^4 = (-\frac{1}{2}, \frac{\sqrt{3}}{2})$.

(iv) $x^6 = ((1+\sqrt{3}) + (1-\sqrt{3})i)$.

(v) $\sum_{i=0}^{4} x^{4-i} = 0$.

Q2: By Solving the equation $x * 5 = i$ algebraically and polarly, prove that:

$cos 18° = \frac{\sqrt{\sqrt{5}+2\sqrt{5}}}{\sqrt[5]{176+8\sqrt{5}}} = \frac{\sqrt{10+2\sqrt{5}}}{4}$.

$sin 18° = \frac{1}{\sqrt[5]{176+8\sqrt{5}}} = \frac{\sqrt{5}-1}{4}$.

2.22 Roots of unity

This section deals with a special case of Theorem 2.6, in which we try to study the equation:

$$x^n = 1 \quad (2.8)$$

whose solutions are known by the nth root of unity;

$1 = (1, 0) = (cos 0°, sin 0°)$.

Thus, the nth root given in the form:

$x = (cos\frac{2k\pi}{n}, sin\frac{2k\pi}{n}), k = 0, 1, 2, ..., n-1$.

If $k = 0$ then $x = 1$. The reminder $n-1$ roots is w^k, where $w = (cos\frac{2\pi}{n}, sin\frac{2\pi}{n})$.

Since, $x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + ..., x + 1)$.

Thus, $w, w^2, w^3, ..., w^{n-1}$ are roots of the equation:

$(x^{n-1} + x^{n-2} + ..., x + 1)$.

For the certain values of n can be solved algebraically, and by comparssion between the algebraic and polar solutions, the algebraic expressions $cos(\frac{2\pi}{n}), sin(\frac{2\pi}{n})$ can be obtained. There are some important interesting properties of the roots of the unity in the literature (Lang, 1984; Bini and Pan, 2012).

Example 2.11 The cube root of the unity is $w = (cos\frac{2\pi}{3}, sin\frac{2\pi}{3})$.

$$w \text{ fulfills the equation } x^2 + x + 1 = 0.$$

$$\because (-\frac{1}{2}, \frac{\sqrt{3}}{2}), (-\frac{1}{2}, -\frac{\sqrt{3}}{2}) \text{ are roots for the equation,}$$

$$\text{besides } cos\frac{2\pi}{3} \text{ is positive and } sin\frac{2\pi}{3} \text{ is negative,}$$

$$\therefore (cos\frac{2\pi}{3}, sin\frac{2\pi}{3}) = (-\frac{1}{2}, \frac{\sqrt{3}}{2}).$$

$$\therefore cos\frac{2\pi}{3} = -\frac{1}{2}, sin\frac{2\pi}{3} = \frac{\sqrt{3}}{2}.$$

Example 2.12 The 5th root of the unity is $w = (cos\frac{2\pi}{5}, sin\frac{2\pi}{5})$.

w fulfills the equation $x^4 + x^3 + x^2 + x + 1 = 0$.

The typical solution of the equation is as follows:

Dividing the equation by x^2, it becomes:

$$x^2 + \frac{1}{x^2} + x + \frac{1}{x} + 1 = 0.$$

$$\text{Let } x + \frac{1}{x} = y,$$

$$\therefore x^2 + \frac{1}{x^2} = y^2 - 2.$$

Thus, we have the equation $y^2 + y + 1 = 0$.

$$\therefore y_1 = \frac{-1 + \sqrt{5}}{2}, y_2 = \frac{-1 - \sqrt{5}}{2}.$$

Now, by solving the equations:

$$x + \frac{1}{x} = y_1,$$

$$x + \frac{1}{x} = y_2.$$

Or,

$$x^2 - y_1x + 1 = 0,$$

$$x^2 - y_2x + 1 = 0.$$

The set solutions of these two equations are:

$$(\frac{-1 + \sqrt{5}}{4}, \frac{\sqrt{10 + 2\sqrt{5}}}{4}), (\frac{-1 + \sqrt{5}}{4}, \frac{-\sqrt{10 + 2\sqrt{5}}}{4}),$$

$$(\frac{-1 - \sqrt{5}}{4}, \frac{\sqrt{10 - 2\sqrt{5}}}{4}), (\frac{-1 - \sqrt{5}}{4}, \frac{-\sqrt{10 - 2\sqrt{5}}}{4}).$$

But each of $cos\frac{2\pi}{5}, sin\frac{2\pi}{5}$ are positive,

$$\therefore (cos\frac{2\pi}{5}, sin\frac{2\pi}{5}) = (\frac{-1 + \sqrt{5}}{4}, \frac{\sqrt{10 + 2\sqrt{5}}}{4}).$$

$$\therefore cos\frac{2\pi}{5} = \frac{-1 + \sqrt{5}}{4}, sin\frac{2\pi}{5} = \frac{\sqrt{10 + 2\sqrt{5}}}{4}.$$

2.23 Exercises

Solve the following questions:

Q1: Prove that the 24th root of the unity $(cos15°, sin15°)$ is fulfills the equation $x^8 - x^4 + 1 = 0$. Furthermore, find an express algebraically and polarity to the equation. (Hint: $x^{24} - 1 = (x^{12} - 1)(x^4 + 1)(x^8 - x^4 + 1)$.)

Q2: If the 7th root of the unity fulfills the equation $x^6+x^5+x^4+x^3+x^2+x+1 = 0$, dividing by x^3, and assuming $y = x + \frac{1}{x}$ then prove that the roots of the equation $y^3+y^2-2y-1 = 0$ are $2cos\frac{2\pi}{7}, 2cos\frac{4\pi}{7}, 2cos\frac{6\pi}{7}$.

Q3: If you know that the 9th root unity is fulfills the equation $x^6+x^3+1 = 0$, then prove that the roots of the equation $y^3-3y+1 = 0$ are $2cos\frac{8\pi}{9}, 2cos\frac{4\pi}{9}, 2cos\frac{2\pi}{9}$.

Q4: Find the 4th root of $-3 - 3i$.

Q5: Find the 4th root of the following complex numbers:

(i) $\frac{-1}{2} + \frac{2}{3}i$.

(ii) $-\frac{3}{4}i$.

Q5: Find the 7th root of the following complex numbers:

(i) $\frac{-3}{2} + \frac{\sqrt{2}}{3}i$.

(ii) $\frac{\sqrt{3}}{5}i$.

(iii) $\sqrt{1 - \frac{2}{3}i}$.

(iv) $(1 + i)^{\frac{1}{3}}$.

(v) $(1 - 2i)^{\frac{1}{5}}$.

(vi) $(2 + 3i)^{\frac{1}{3}}$.

(vii) $(3 - 3i)^{\frac{1}{3}}$.

(viii) $(3 - 3i)^4$.

Q6: Prove that the roots of the equation $x^3 + x^2 - 2x - 1 = 0$ are $2cos\frac{2}{7}\pi, 2cos\frac{4}{7}\pi, 2cos\frac{6}{7}\pi$.

Q7: Prove that the roots of the equation $x^3 - 3x + 1 = 0$ are $2cos\frac{8}{9}\pi, 2cos\frac{4}{9}\pi, 2cos\frac{2}{9}\pi$. Noting that the ninth root of the integer unit satisfies the equation $x^6 + x^3 + 1 = 0$.

3

Polynomials

3.1 Introduction

There are some sets in mathematics with certain conditions and characteristics that have a pivotal role in structural the infrastructure of mathematics, such as; mathematical systems, numerical systems, groups, rings, and fields. In what follows, we begin by defining and studying each of the rings and fields due to their basic role related to polynomials.

3.1.1 Field of real numbers

In mathematics, a real closed field is a field F that has the same first-order properties as the field of real numbers. Some examples are the field of real numbers, the field of real algebraic numbers, and the field of hyper-real numbers. The field of reals is the set of real numbers, which form a field. This field is commonly denoted $\mathbb{R}$.

3.1.2 Real number ring properties

The field F of real numbers has these basic properties (Beachy and Blair, 2006; Fraleigh, 2003; McCoy, 1968):

(i) The addition operation is closed on F. Or, $(a+b) \in F, \forall a, b \in F$.

(ii) The multiplication operation is closed on F. Or, $(a.b) \in F, \forall a, b \in F$.

(iii) Associativity of addition. Or, $a+(b+c) = (a+b)+c, \forall a, b, c \in F$.

(iv) Associativity of multiplication. Or, $a.(b.c) = (a.b).c, \forall a, b, c \in F$.

(v) Commutativity of addition. Or, $a + b = b + a, \forall a, b \in F$.

(vi) Commutativity of multiplication. Or, $a.b = b.a, \forall a, b \in F$.

(vii) Additive identity. Or, there exist a distinct element $0 \in F$ such that $a + 0 = a$.

(viii) Multiplicative identity. Or, there exist a distinct element $1 \in F$ such that $a.1 = a$.

(ix) Additive inverses. Or, $\forall a \in F, \exists(-a) \in F$ called the additive inverse of a, such that $a + (-a) = 0$.

(x) Multiplicative inverses. Or, $0 \neq \forall a \in F, \exists a^{-1} \in F$, called the multiplicative inverse of a, such that $a.a^{-1} = 1$.

(xi) Distributivity of multiplication over addition. Or, $a.(b + c) = (a.b) + (a.c), \forall a, b, c \in F$.

Note:

(1) Since the commutative property is available in the ring of real numbers, it is always a commutative ring.

(2) If the ring has multiplicative inverses element, it called a ring with a multiplicative inverses element.

3.1.3 Fields

Based on Hamadameen (2022), and some others (Beachy and Blair, 2006; Fraleigh, 2003; McCoy, 1968; Sharpe, 1987), the field of real numbers can be defined as follows:

Definition 3.1 Let $\phi \neq A$, and $*, \#$ be binary operations on A. The mathematical system $(A, *, \#)$ is called a field if and only if

(1) $(A, *)$ is a commutative group.

(2) $(A', \#')$ is a commutative group where $A' = A \backslash \{0\}$, 0 is a unit element with respect to $*$, and $\#'$ is a restriction operation on A'.

(3) Distribution laws are fulfilled. Or, if $\forall x, y, z \in A$, then:

(a) $x \cdot (y + z) = x \cdot y + x \cdot z$.

(b) $(y + z) \cdot x = (y \cdot x) + (z \cdot x)$.

3.2 Polynomials

In mathematics, a polynomial is an expression consisting of variables and coefficients, that involves only the operations of addition, subtraction, multiplication, and positive integer powers of variables. An example of a polynomial of a single variable x is $x^4 - 4x + 7$. An example with two variables is $x^3 + 2xy - 1$, and so on...

Polynomials appear in many fields of mathematics and science. For example, they are used to form polynomial equations, which encode a wide range of problems, from elementary word problems to complicated scientific problems; they are used to define polynomial functions, which appear in settings ranging from basic chemistry and physics to economics and social science; they are used in calculus and numerical analysis to approximate other functions. In advanced mathematics, polynomials are used to construct polynomial rings and algebraic varieties, which are central concepts in algebra and algebraic geometry (Prasolov, 2004; Marden, 1949; Bell, 1934).

3.2.1 Polynomial concept

Definition 3.2 The polynomial P is a function $P(x) : \mathbb{R} \rightarrow \mathbb{R} \ni P(x) = a_0 + a_1 x + a_2 x^2 + ... + a_{n-1} x^{x-1} + a_n x^n = \sum_{k=0}^{n} a_k x^k$, where $a_k \in \mathbb{R}$ is a coefficient of $x^k, \forall k = 0, 1, ..., n-1, n$ (Barbeau, 2003; Borwein and Erdélyi, 1995).

3.2.2 Properties of polynomials

Polynomials have some important properties, including the following:

(i) Polynomial equalities. Two polynomials are equal if the corresponding coefficients are equal in each of them, for example, polynomials;

$a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 = b_3x^3 + b_2x^2 + b_1x + b_0 \Leftrightarrow a_4x^4 + (a_3 - b_3)x^3 + (a_2 - b_2)x^2 + (a_1 - b_1)x + (a_0 - b_0) = 0 \Leftrightarrow a_4 = 0, a_3 - b_3 = 0, a_2 - b_2 = 0, a_1 - b_1 = 0, a_0 - b_0 = 0.$

(ii) Addition and subtraction. It is possible to add polynomials $f(x), g(x)$ in the degrees m, and n respectively, such that $n \leq m$, and the degree of the polynomial $(f(x) + g(x)) \leq m$. Morever, polynomials can be added using the associative law of addition, possibly followed by reordering, and combining of like terms (Edward, 1995; Salomon, 2006), for example, polynomials;

If $P = 3x^2 - 2x + 5xy - 2, Q = -3x^2 + 3x + 4y^2 + 8$ then: $PQ = x + 5xy + 4y^2 + 6$.

Note: (1) Subtraction of polynomials is similar. (2) When polynomials are added together, the result is another polynomial (Barbeau, 2003).

(iii) Multiplication. Polynomials can also be multiplied. To expand the product of two polynomials into a sum of terms, the distributive law is repeatedly applied, which results in each term of one polynomial being multiplied by every term of the other (Edward, 1995; Salomon, 2006), for example, polynomials;

If $P = 2x + 3y + 5.Q = 2x + 5y + xy + 1$ then $PQ = 4x^2 + 21xy + 2x^2y + 12x + 15y^2 + 3xy^2 + 28y + 5$.

Note: The product of polynomials is always a polynomial (Barbeau, 2003).

(iv) Composition. Given a polynomial f of a single variable and another polynomial g of any number of variables, the composition $f \circ g$ s obtained by substituting each copy of the variable of the first polynomial by the second polynomial (Barbeau, 2003). For example, if $f(x) = x^2 + 2x, g(x) = 3x + 2$ then $(f \circ g)(x) = f(g(x)) = (3x + 2)(3x + 4)$.

Note: The composition of two polynomials is another polynomial (Kriete, 1998).

(v) Division. The division of one polynomial by another is not typically a polynomial. Instead, such ratios are a more general family of objects, called rational fractions, rational expressions, or rational functions, depending on context (Kriete, 1998; Marecek, 2017; Haylock and Cockburn, 2008). For example, $\frac{1}{x^2+1}$ is not a polynomial, and it cannot be written as a finite sum of powers of the variable x.

Note: For polynomials in one variable, there is a notion of Euclidean division of polynomials, generalizing the Euclidean division of integers. This notion of the division $\frac{a(x)}{b(x)}$ results in two polynomials, a quotient $q(x)$ and a remainder $r(x)$, such that $a = bq + r$ and degree $(r) <$ degree (b). The quotient and remainder may be computed by any of several algorithms, including polynomial long division and synthetic division (Selby and Slavin, 1991; Marecek, 2017; Lipschutz and Lipson, 2018).

(vi) Factoring. All polynomials with coefficients in a unique factorization domain also have a factored form in which the polynomial is written as a product of irreducible polynomials and a constant. This factored form is unique up to the order of the factors and their multiplication by an invertible constant. In the case of the field of complex numbers, the irreducible factors are linear. Over the real numbers, they have the degree either one or two. Over the integers and the rational numbers the irreducible factors may have any degree (Barbeau, 2003). For example, the factored form of $5x^5 - 5 = 5(x-1)(x^2+x+1) = 5(x-1)(x+\frac{1+i\sqrt{3}}{2})(x+\frac{1-i\sqrt{3}}{2})$ over the integers and the reals and the complex numbers respectively.

3.2.3 The quotient of polynomials

Consider the polynomials $f(x), g(x) \neq 0$ over the field $\mathbb{F}$, there exists the polynomials $q(x), r(x)$ over $\mathbb{F}$ such that $r(x) = 0$, or its degree is less than the degree of $g(x)$, in which $f(x) = q(x)q(x) + r(x)$.

This property enables us to recall the long division we learned earlier. For example, if $f(x) = 4x^3 - 2x^2 = 5x + 3, g(x) = 2x^2 + x - 1$. Since the first term of $f(x)$ is $4x^3$ in which can obtained from multiplied of the first term of $g(x)$ by $2x$. Or, $(2x^2)(2x)$, and the difference $f(x) - 2xg(x) = -4x^2 + 7x + 3$ is in the second degree, and by repeating the algorithm, we will get;

$f(x) - 2xg(x) + 2g(x) = qx + 1$.

Since the degree of the reminder of the right hand side is less than the degree $g(x)$ that why we can not go on on the algorithm, and thereby, we have;

$f(x) = (2x - 2)g(x) + q(x) + 1$, where $q(x) = 2x - 2, r(x) = qx + 1$.

This is the general method of the long division. In the case if we assume that $f(x)$ has the degree less than the degree of $g(x)$ then we take;

$q(x) = 0$, and $r(x) = f(x)$.

Now, we are going to prove that the uniqueness of the $q(x), r(x)$.

Suppose that $f(x) = q_1(x)g(x) + r_1(x)$, where $r_1(x) = 0$ or has the degree m less than the degree of $g(x)$.

Thus we have $q(x) - q_1(x)g(x) = r_1(x) - r(x)$.

If the term $q(x) - q_1(x)$ does not vanishing, that means has the degree greater than or equal to zero. Thereby, the left hand side of the previous equation has degree greater than or equal to m, and in that case;

$q(x) = q_1(x)$ and $r(x) = r_1(x)$.

The following example illustrates the algorithm of the long division.

Example 3.1 If $f(x) = x^5 - 4x^3 + 2x^2 - 5x - 8, g(x) = x - 3$ then evaluate $\frac{f(x)}{g(x)}$.

Solution: By long division as shown in Table 3.1, we get the value;
$f(x) = (x^4 + 3x^3 + 5x^2 + 17x + 46)(x - 3) + 130$.

3.2.4 Long division algorithm of polynomials

The following are the steps for the long division of polynomials:

Step1. Arrange the terms in the decreasing order of their indices (if required). Write the missing terms with zero as their coefficient.

Table 3.1: Long devision of polynomials

1	0	-4	2	-5	-8	
	3	9	15	51	38	3
1	3	5	17	46	130	

Step2. For the first term of the quotient, divide the first term of the dividend by the first term of the divisor.

Step3. Multiply this term of the quotient by the divisor to get the product.

Step4. Subtract this product from the dividend, and bring down the next term (if any). The difference and the brought down term will form the new dividend.

Step5. Follow this process until you get a remainder, which can be zero or of a lower index than the divisor.

3.3 Exercises

Solve the following questions:

Q1: What is the degree of a polynomial in x and in x^2 of x^5+4x^2+1?

Q2: Find each of $q(x), r(x)$ for; (1) $f(x) = x^4 + 4x^3 + 4x^2 + 7x + 13, g(x) = x+2$. (2) $f(x) = 3x^5-5x^4+6x^3-8x^2+1x-23, g(x) = x-5$.

3.4 Number of roots of a polynomial equation

Roots of polynomials are the solutions for any given polynomial for which we need to find the value of the unknown variable. If we know the roots, we can evaluate the value of polynomial to zero. Let us assume that the equation:

$f(x) = 2x^6 + 15x^5 + 3x^4 - 94x^3 + 84x^2 + 15x - 25 = 0$. It is a polynomial of the sixth degree and can be analyzing to rewire it as follows:

$(x+5)^2(2x+1)(x-1)^3 = 0$.

Or, each of $(x+5), (2x+1), (x-1)$ is a factor of the polynomial factors. In othere words; $-5, -5, -\frac{1}{2}, 1, 1, 1$ are roots of the polynomial.

In addition, -5 is a repeated (duplicate) root, $-\frac{1}{2}$ is a simple (single) root, and 1 is a triple root. Thereby, the repeated roots appear twice, and the triple roots appear tree times in addition to the single root. Thus, the equation has six roots. The following theorem, called the basic theorem in algebraic which pertain to the roots of polynomial equations.

Theorem 3.1 (The Fundamental Theorem of Algebra)

If $P(x)$ is a polynomial of degree $n \geq 1$, then $P(x) = 0$ has exactly n roots, including multiple and complex roots.

Proof The proof of the theorem relies on attempts by some researchers such as Aigner and Ziegler (1999), Basu (2021), Ahlfors (1979), Shipman (2007), and Aliabadi and Darafsheh (2015) (See Appendix A). ⧫

Theorem 3.2 *Every polynomial of degree $n, n \geq 1$, has exactly n roots, provided that a root repeated by m times is a root of m.*

Proof Suppose that $f(x)$ is a polynomial of degree n, where

$f_n(x) = a_0x^n + a_1x^{n-1} + a_{n-1} + a_0$.

Based on Theorem 3.1, we have,

$f_n(x) = 0$,

has at least one root r_1.

In this case $x - r_1$ is a factor of factors $f_n(x)$. Or,

$f_n(x) = (x - r_1)f_{n-1}(x)$,

where $f_{n-1}(x)$ is a polynomial of the degree $n - 1$, in which the highest power of the variable x is a_0x^{n-1}, and the equation;

$f_{n-1}(x) = 0$,

has at least one root r_2 (Based on Theorem 3.1), and in this case $x - r_2$ is a factor of factors of f_{n-1}, it means;

$f_{n-1}(x) = (x - r_2)f_{n-2}(x)$.

Thus $f_n(x) = (x - r_1)(x - r_2)f_{n-2}(x)$,

where f_{n-2} is a polynomial of the degree $n - 2 \geq 1$, has the highest power for the term a_0x^{n-2}.

By repeating the process n times, we obtain;

$f_n(x) = (x - r_1)(x - r_2)...(x - r_n)f_0(x)$,
where $f_0(x)$ if a polynomial of degree zero $a_0x^0 = a_0$, it means;

$$f_n(x) = a_0(x - r_1)(x - r_2)...(x - r_n) \tag{3.1}$$

The roots of the equation $f_n(x) = 0$ are $r_1, r_2, ...r_n$.

To prove that there can not be more than n roots, let us assume that there is $s \neq r_i, i = 1, 2, ..., n$ for (3.1). That means;

$f_n(s) = a_0(s - r_1)(s - r_2)...(s - r_n)$.

Since any factor of factors in the right hand side of the equation is not equal to zero, hence

$f_n(s) \neq 0$.

Thereby, s does not be a root for $f_n(x) = 0$.

Thus the equation has exactly n roots. ♦

3.5 Complex roots

The following theorem emphasizes that the complex roots of a polynomial equation if exist in a real field occur in the form of duplicate (Gehman, 1941; Ballantine, 1959).

Theorem 3.3 *If the complex number (a, b) is a root of the polynomial equation $f(x) = 0$ on the real field, then then its complex conjugate $(a, -b)$ is also a root of it.*

Proof Suppose that,

$\gamma = (a, b)$ is a root of $f(x) = a_0x^n + a_1x^{n-1} + ... + a_n = 0$.

Then, $f(\gamma) = a_0\gamma^n + a_1\gamma^{n-1} + ... + a_n = 0$.

Therefore, $f(\bar{\gamma}) = a_0(\bar{\gamma})^n + a_1(\bar{\gamma})^{n-1} + ... + a_n = \bar{0} = 0$.

Thus, $\bar{\gamma} = (a, -b)$ is a root of $f(x) = 0$. ♦

Example 3.2 (1) Find the equation if it has the roots, $3, 2+5i, 2-5i$.

(2) Find the equation if it has the roots, $-1, -1, \frac{4}{5}, 3 + 2i$.

(3) Solve the equation; $4x^3 - x^2 - 100x + 25 = 0$.

Solution: (1) $(x-3)(x-2-5i)(x-2+5i) = (x-3)(x^2-4x+29)$.

(2) $(x-(3+2i))(x-(3-2i))(x-\frac{4}{5})(x+1)^2 = 5x^5 - 24x^4 + 26xx^3 + 92x^2 - 15x - 52 = 0$.

(3) $4x^3 - x^2 - 100x + 25 = 0 \Rightarrow x^2(4x-1) - 25(4x-1) = 0 \Rightarrow (4x-1)(x^2-25) \Rightarrow x = \frac{1}{4}, x = 5, x = -5$.

3.6 Bound of the roots

This section discusses how to find the limits of the roots between which the real roots of polynomial equations can lie. Let's consider the following equation:

$$f(x) = a_0x^n + a_1x^{n-1} + ... + a_n = 0 \tag{3.2}$$

where $a_0 \neq 0, a_i \in \mathbb{R}, \forall i$.

By utilizing the fifth property of 3.2.2, and deviding $f(x)$ by $(x - k), k \neq 0$, and if all the coefficints in the third line are positive or zero, then k will be the uppuer bound of the positive roots for (3.2).

For finding bounds of the negative roots, we substitute each x by $-y$ of (3.2) to obtain;

$$f(x) = a_0y^n - a_1y^{n-1} + a_2y^{n-2} - ...(-1)^n a_n = 0 \tag{3.3}$$

Equation (3.4) has roots that are opposite to the sign of Equation (3.3). By the same previous method we find the upper bounds for the roots of (3.4). If we assume that b, then $-b$ is the lower bound for roots of equation (3.3) provided that all the coefficients in (3.3) are positive, the roots are all negative (Hirst and Macey, 1997; Lagrange, 1879; Cauchy, 1828; Marden, 1949; Fujiwara, 1916; Kojima, 1917; Akritas et al., 2008; Ştefănescu, 2007).

Example 3.3 Find the upper and lower bounds of the roots of the equation $4x^3 + 2x^2 - 7 = 0$.

Solution: First, putting $k = 1$, we find that; Since, in the third

Table 3.2: Bounds on roots of polynomial-i

4	2	0	-7	
	4	6	6	1
4	6	6	-1	

line all coefficints are not positive, hence 1 does not necessary to be uppur bond for roots, as shown in Table 3.2. Putting $k = 2$, we obtain that all the coefficints in the third line are positive, as shown in Table 3.3. Thus, 2 is the uppuer bound for the positive roots. To obtaining

Table 3.3: Bounds on roots of polynomial-ii

4	2	0	-7	
	4	10	20	2
4	10	20	33	

the lower bound for the negative roots, we substitute for each x by $-y$, So we get $4y^3 - 2y^2 + 7 = 0$. Now by putting $k = 1$, we get that, all the coefficints in the third line are positive, as shown in Table 3.4. Thereby, -1 is the lower bound for the negative root. The sketch diagram of the function $4x^3 + 2x^2 - 7 = 0$ is shown in Figure 3.1.

Table 3.4: Bounds on roots of polynomial-iii

4	-2	0	7	
	4	2	2	1
4	2	2	9	

3.7 Exercises

Solve the following questions:

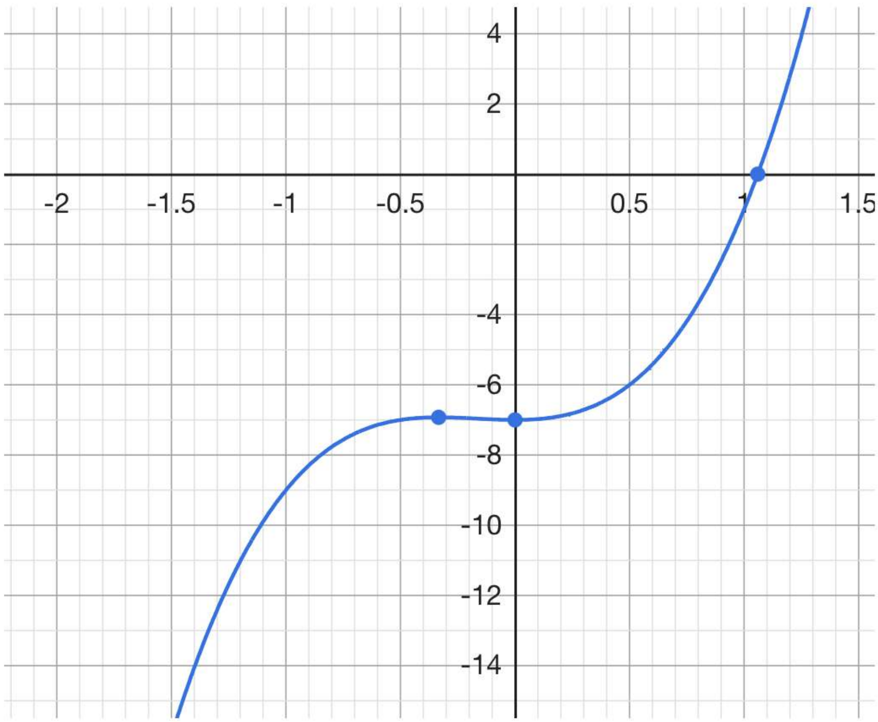
Q1: Solve each of equations:

(1) $(x+5)^2(x-1)(x^2-6x+2) = 0$. (2) $(7x-3)^2(x+4)(x^2+25)-0$. (3) $(2x-1)^3(x+8)(x^2+10x+29) = 0$. (4) $(3x^2-7x)(x^2-12x+3) = 0$.

Q2: Find the equation that has real coefficients and roots shown in each of the following cases:

(1) $0, 3, -3, -2$. (2) $0.0, -1, \frac{1}{3}, \frac{21}{7}$. (3) $-2, \sqrt{11}, -\sqrt{11}$. (4) $1, \frac{1}{2} + 3i, \frac{1}{2} - 3i$. (5) $5, 5, 0, 0, 0, -\frac{1}{5}$.

Q3:



w

Figure 3.1: Bounds on roots for $4x^3 + 2x^2 - 7 = 0$

(i) Find a polynomial equation of the third degree whose coefficients are real and have two roots $6, -1+i$.

(ii) Find a polynomial equation of the forth degree whose coefficients are real and have two roots $i, 3-i$.

Q4: Find the upper and lower bounds of the following equations:
(1) $2x^3 - 9x^2 + 5x + 7 = 0$. (2) $2x^3 + x^2 - 3x + 8 = 0$. (3) $x^5 - 3x^4 - 6x + 10 = 0$. (4) $x^3 - x^2 - 8 = 0$.

Q5: If a, b are integers, prove that for a polynomial equation with arational factors, if $a + \sqrt{b}$ is a root, then $a - \sqrt{b}$ is a root too.

Q6: Use the long division method to show that $x^4 - 4x^3 - x^2 + 14x + 10 = 0$ has a duplicate -1, and find remaining roots.

Q7: If r_1, r_2, r_3 are roots of the equation $x^3 + b_1x^2 + b_2x + b_3 = 0$then prove that $b_1 = -(r_1 + r_2 + r_3), b_2 = r_1r_2 + r_2r_3 + r_3r_1, b_3 = -r_1r_2r_3$.

Q8: If $f(0) \neq 0$, and $r_1, r_2, ..., r_n$ are roots of $f(x) = 0$ then prove that $f(x) = f(0)(1 - \frac{x}{r_1})(1 - \frac{x}{r_2})...(1 - \frac{x}{r_n})$.

3.8 The relationship between roots and polynomial equations

There exists a relationship between a polynomial roots and their coefficients based on the mathematical logic(Dickenstein, 2005; Kudryashov and Demina, 2007). Suppose that $\alpha_1, \alpha_2, ..., \alpha_n$ are roots of the polynomial in the nth degree:

$$\begin{gathered} f_n(x) = a_0x^n + a_1x^{n-1} + ... + a_n \\ = a_0(x - \alpha_1)(x - \alpha_2)...(x - \alpha_n) \\ a_0 \neq 0 \end{gathered}$$

(1) If $n = 2$, the polynomial will be in the second degree, and it can be writtin as follows:

$$\begin{gathered} a_0x^2 + a_1x + a_2 \\ = a_0(x - \alpha_1)(x - \alpha_2) \\ a_0[x^2 - (\alpha_1 + \alpha_2)x + \alpha_1\alpha_2] \end{gathered}$$

Now, by equating the coefficients of powers of x, we find that;

$$\frac{a_1}{a_0} = -(\alpha_1 + \alpha_2)$$
$$\frac{a_2}{a_0} = \alpha_1\alpha_2$$

(2) If $n = 3$, the polynomial will be in the third degree, and it can be writtin as follows:

$$a_0x^3 + a_1x^2 + a_2x + a_3$$
$$= a_0(x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$$
$$= a_0[x^3 - (\alpha_1 + \alpha_2 + \alpha_3 +)x^2 + (\alpha_1\alpha_2 + \alpha_3\alpha_1 + \alpha_3\alpha_2)x - \alpha_1\alpha_2\alpha_3]$$

By equating the coefficients of powers of x, we find that;

$$\frac{a_1}{a_0} = -(\alpha_1 + \alpha_2 + \alpha_3)$$
$$\frac{a_2}{a_0} = \alpha_1\alpha_2 + \alpha_3\alpha_1 + \alpha_3\alpha_2$$
$$\frac{a_3}{a_0} = -\alpha_1\alpha_2\alpha_3$$

By repeating this algorith for $n = 4, 5, ...,$ we conclude that:

$$-\frac{a_1}{a_0} = \sum_{i=1}^{n} \alpha_i$$
$$\frac{a_2}{a_0} = \sum \alpha_1\alpha_2$$
$$-\frac{a_3}{a_0} = \sum \alpha_1\alpha_2\alpha_3$$
$$\vdots$$
$$(-1)^i = \frac{a_i}{a_0} = \sum \alpha_1\alpha_2\alpha_3...\alpha_i$$
$$\vdots$$
$$(-1)^n = \frac{a_n}{a_0} = \alpha_1\alpha_2\alpha_3...\alpha_n$$

Example 3.4 Find the roots of the equation $6x^3-18x^2+24x-12=0$, If the product of two roots is 3.

Solution: Suppose that the roots are r_1, r_2, r_3.

$$\begin{aligned} r_1+r_2+r_3 &= \frac{18}{6} = 3 \\ r_1r_2+r_1r_3+r_3r_3 &= \frac{24}{6} = 4 \\ r_1r_2r_3 &= \frac{12}{6} = 3 \end{aligned} \tag{3.4}$$

Because of the product of two roots is 3. Suppose that those two roots are r_1, r_2.

Thus $r_1r_2 = 3$, implies that $r_3 = 1$.

This implies that $r_1 + r_2 = 2$. Or, $r_1 = 2 - r_2$.

By substituting into the second equation from (3.4), we get;

$(2-r_2)r_2 + (2-r_2) + r_2 = 4$. Or,

$r^2 - 2r_2 + 2 = 0$.

Thereby, $r_2 = \frac{2\pm\sqrt{4-8}}{2} = 1 \pm i$.

Thus, the roots are $1+i, 1-i, 1$.

Example 3.5 Find the cube of the sum of the squares of the roots of the equation $x^4 - 4x^3 + 3x^2 + x - 1 = 0$.

Solution: Suppose that the roots are $\alpha, \beta, \gamma, \delta$.
As, $\alpha + \beta + \gamma + \delta = \frac{4}{1} = 4$.
$\alpha\beta + \beta\gamma + \gamma\alpha + \alpha\delta + \delta\beta + \delta\gamma = \frac{3}{1} = 3$.
But, $(\alpha+\beta+\gamma+\delta)^2 = \alpha^2+\beta^2+\gamma^2+\delta^2+2(\alpha\beta+\beta\gamma+\gamma\alpha+\alpha\delta+\delta\beta+\delta\gamma)$
$16 = \alpha^2 + \beta^2 + \gamma^2 + \delta^2 + 6$.
$\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 10$.
Thus, $(\alpha^2 + \beta^2 + \gamma^2 + \delta^2)^3 = 10^3 = 1000$.

3.9 Exercises

Solve the following questions:

Q1: Solve the equation $2x^3 + 4x^2 + 6x + 4 = 0$, if you know that its roota are α, β, γ, and $\alpha = \beta + \gamma$.

Q2: Solve the equation $4x^3 - 2x^2 - 36x + 20 = 0$, if you know that its roota are α, β, γ, and $\alpha + \beta = 0$.

Q3: Solve the equation $2x^3 - 14x^2 - 42x + 216 = 0$, if you know that its roota are α, β, γ, and $\gamma = \sqrt{\alpha\beta}$.

Q4: Solve the equation $3x^3 + 27x^2 + 18x - 168 = 0$, if you know that its roota are α, β, γ, and $\beta = -2\alpha$.

Q5: Solve the equation $9x^3 - 36x^2 + 44x - 16 = 0$, if you know that its roota are α, β, γ, and if the roots are numeric sequence.

Q6: Solve the equation $3x^3 - 26x^2 + 52x - 29 = 0$, if you know that its roota are α, β, γ, and if the roots are geometric sequence.

Q7: Solve the equation $2x^3 - 12x^2 + 6x + A = 0$, if you know that its roota are α, β, γ, find the value of A, if $\alpha = 2(\beta + \gamma)$.

Q8: Solve the equation $x^3 - 4x^2 + Ax + 46 = 0$, if you know that its roota are α, β, γ, find the value of A, and the roots are numeric sequence.

Q9: Solve the equation $x^3 + Cx^2 + Bx + A = 0$, if you know that its roota are α, β, γ. What is the relationship between A, B, C if the roots are geometric sequence?

Q10: Find the relationship between p, q of the equation $x^3+px+q = 0$, if you know that its roota are α, β, γ, and the equation has the duplicated root.

Q11: If the roots α, β, γ of $x^3 + Bx^2 + Ax + C = 0$ satisfy the relation $\alpha\beta = -\gamma^2$. Prove that $(2A - B^2)^3\gamma = (BA - AC)^3$.

Q12: Solve the equation $4x^4 - 8x^3 + 8x^2 - x - 16 = 0$, if you know that its roota are $\alpha, \beta, \gamma, \delta$, and $\alpha + \beta = 2$.

Q13: Calculate the sum of the square roots of the equation $4x^4 - 8x^3 + 16x^2 - 12x + 8 = 0$, if you know that its roota are $\alpha, \beta, \gamma, \delta$.

3.10 Duplicate roots

As we explained earlier that a is a root of the equation:

$$f(x) = 0 \tag{3.5}$$

if $f(a) = 0$. Let us study the case if a is a root of the equation (3.5), but repeated k times. $f(x)$ can be written as follows:

$$f(x) = (x - a)^k g(x)$$
$$g(a) \neq 0$$

a is called a zero of the polynomial $f(x)$ repeated k times (Barbeau, 2003; Leung et al., 1992; McNamee and Pan, 2013). In the case of $k = 1$ then a is called a simple zero (Krantz et al., 1999). The following theorem proves and explains the fact of the repeated root.

Theorem 3.4 *If a is a zero of the polyomial $f(x)$ and repeated k times, where $k > 1$ then a is a zero repeated $k-1$ times and derived from $f(x)$.*

Proof Let us express $f(x)$ as follows:

$f(x) = (x - a)^k g(x), g(x) \neq 0$.

By derivation, we find that;

$$\begin{aligned} f'(x) &= k(x-a)^{k-1} g(x) + (x-a)^k g'(x) \\ &= (x-a)^{k-1}[kg(x) + (x-a)g'(x)] \end{aligned}$$

Thereby, a is a zero of the $f'(x)$ and repeated $k-1$ times. If the repetition is more than $k-1$ it means that $kg(x)+(x-a)g'(x)$ is divisible by $(x-a)$.

Thus, $g(x)$ should be divisible by $(x-a)$ without remains, or $g(a) = o$. And this is contraduction. ⧫

3.11 Greatest common factor of polynomials

Let p, q be polynomials with coefficients in an integral domain F, typically a field or the integers. A greatest common factor or greatest common divisor (GCF) of p, q is a polynomial d that divides p and q, and such that every common divisor of p and q also divides d. Every pair of polynomials (not both zero) has a $GCF(GCD)$ if and only if F is a unique factorization domain (Basu, 2021; Knuth, 2014; van and Monagan, 2004).

Let us consider the polynomials:

$p = (x-1)^2, q = x^3 - 1$.

$p = (x-1)(x+1), q = (x-1)(x^2+x+1)$.

The $GCF = (x-1)$.

Suppose that we have the polynomials f, f_1. Dividing f by f_1, suppose that the quotient is q_1, and the remainder is f_2. Or,

$f = f_1 q_1 + f_2$.

If we assume that $f_2 \neq 0$, and dividing f_1 by f_2. Let q_2, f_3 be the the quotient and remainder respectively. Or,

$f_1 = f_2 q_2 + f_3$.

Again, we assume that $f_3 \neq 0$, and dividing f_2 by f_3. Let q_3, f_4 be the the quotient and remainder respectively. Or,

$f_2 = f_3 q_3 + f_4$.

By repeating of this process (where the degree of the polynomial is decreasing) as long as the remainder is not equal to zero until we get:

$f_{r-1} = f_r q_r$.

Thus, we obtain the following identities:

$$\begin{aligned} f &= f_1 q_1 + f_2 \\ f_1 &= f_2 q_2 + f_3 \\ &. \\ &. \\ &. \\ f_{r-2} &= f_{r-1} q_{r-1} + f_r \\ f_{r-1} &= f_r q_r \end{aligned}$$

From the identities, f_r is the GCF between f, f_1 because f_{r-1} is divisible by f_r. Also,

$f_{r-2} = f_{r-1}q_{r-1} + f_r$, f_{r-1} is the GCF, because it is divisible by f_r. And whereas f_{r-1}, f_{r-2} are divisible by f_r thus, f is divisible by $f_r, ...$ So we find that f_1, f are divisible by f_r.

Moreover, the identities reflect us that, f_r can be divided by any common factor (d) of polynomials $f, f_1, ..., f_{r-1}$. This can be proven as follows:

Suppose that each of f, f_1 divisible by d such that;

$f_2 = f - f_1 q_1$.

Also, we note that f_2, f_3 are divisible by d such that;

$f_3 = f_1 - f_2 q_2$.

Thus, we find that f_r, f_{r-1} are divisible by d, and since f_r can be divided by any d between f, f_1, So none of these common factor (d)s are of greater degree than f_r. Thus, f_r is the GCF.

If d is the GCF in the same degree, then f_r can be divided by d and the result is fixed amount. So there is an infinite number of common factors for polynomials f, f_1, but they are all of the same degree and in the form of:

cf_r,

where c is a constant, but When dividing by any polynomial, it differs from another polynomial by a constant only, and, in fact is not a new polynomial.

Thus, the GCF is a unique factor regardless of the constant.

Example 3.6 Fin the greatest common factor between the polynomils:

$f = x^6 + 2x^5 + x^3 + 3x^2 + 3x + 2$, $f_1 = x^4 + 4x^3 + 4x^2 - x - 2$

Solution: We divide f by f_1 as follows (we will follow the long division method and write the coefficients of all polynomials) as shown in Table 3.5.

Table 3.5: Greatest common factor (i)

1	2	0	1	3	3	2
1	4	4	−1	−2		
	−2	−4	2	5	3	
	−2	−8	−8	2	4	
		4	10	3	−1	2
		4	16	16	−4	−8
			−6	−13	3	10

1	4	4	−1	−2
1	−2	4		

The quotient is $x^2 - 2x + 4$, and the reminder is $f_2 = -6x^3 - 13x^2 + 3x + 10$.

Now, we divide f_1 by f_2, but the qoutient may give fractional equations, so we will multiply the coefficients of f_1 by 6 (Or, f_3 will produce and multiply by a fixed amount, but this is not an important matter in the processes that we are experimenting with) as shown in Table 3.6.

Again, to avoid the appearance of fractional transactions, we will multiply the last line by the number 6 , and going on with our operations as shown in Table 3.7

Since the last line is a multiple of 19 hence it can be written in the form:

$f_3 = x^2 + 3x + 2.$

Now, we divide f_2 by f_3 as shown in Table 3.8 (iv).

In Table 3.8 (iv), since the reminder is zero hence the GCF is $x^2 + 3x + 2$, and this method is called traditional algorithmic division.

Table 3.6: Greatest common factor (ii)

6	24	24	−6	−12
6	13	−3	−10	
	11	27	4	−12

−6	−13	3	10
−1	−11		

Table 3.7: Greatest common factor (iii)

66	162	24	−72	
66	143	−33	−100	← Turn the sign
	19	57	38	

Table 3.8: Greatest common factor (iv)

−6	−13	3	10
−6	−18	−12	
	5	15	10
	5	15	10
	0	0	0

1	2	3
−6	5	

Example 3.7 Show that the $GCF(x^5 - x^4 - 2x^3 + 2x^2 + x - 1, 5x^4 - 4x^3 - 6x^2 + 4x + 1) = x^3 - x^2 - x + 1$.

Solution: Has been left as an exercise for the reader.

3.12 Exercises

Find the GCF s for the following polynomials:

Q1: $f = 2x^4 + 2x^3 - 3x^2 - 2x + 1, f_1 = x^3 + 2x^2 + 2x + 1$.

Q2: $f = x^4 - 6x^2 - 8x - 3, f_1 = x^3 - 3x - 2$.

Q3: $f = 10x^6 - 9x^5 - 12x^4 + 2x^2 - x - 1, f_1 = 4x^5 + x^4 - 7x^3 - 8x^2 - x + 1$.

3.13 Solving cubic equations using Cardan's method

A cubic equation in one variable is an equation of the form:

$$ax^3 + bx^+cx + d = 0, a \neq 0 \tag{3.6}$$

The solutions of this equation are called roots of the cubic function defined by the left-hand side of the equation. If all of the coefficients a, b, c, and d of the cubic equation are real numbers, then it has at least one real root (this is true for all odd-degree polynomial functions).

Since ancient times, mathematicians have been busy finding the general formula for equations of the third degree and their solutions (Mikami, 1913; Khayyam, 1963; O'Connor and Robertson, 2001; Berggren et al., 1986; Bidyāraṇya and Singh, 1962; Rowe, 1994). Later, the trigonometric solution for the cubic equation with three real roots has been derived and extended (Nickalls, 2006).

Finally, Cardano (1501-1576) came up with a mathematical formulation for the cubic equations as in (3.6), and found a solution to it (Branson, 2013; Cardano and Witmer, 1993; Cardano, 2002). The Cardano's formula, which is similar to the perfect-square method to quadratic equations, is a standard way to find a real root of a cubic equation.

In general, the Algorithm of Cardano's Method was as follows:

Consider the cubic function in (3.6), and for convenient we express it as;

$f(x) = x^3 + ax^2 + bx + c = 0,$

if $c > 0$ there exists a negative root, and if $c < 0$ there exists a positive root.

Assume that;

$x = y + k$, k is arbitrary constant.

From Taylor's expansion (Thomas et al., 2010)(See Appendix B); if the function f is a continuous function and differentiated for any order over the interval $[y, y + k]$ then:

$$f(y+k) = f(k)+f^{'}(k)y+\frac{f^{''}(k)}{2!}y^2+\frac{f^{'''}(k)}{3!}y^2+...+\frac{f^{(n)}(k)}{n!}y^n+... \quad (3.7)$$

From (3.7), we have:

$$f(k) = k^3 + ak^2 + bk + c$$
$$f^{'}(x) = 3k^2 + 2ak + b; x = k$$
$$f^{''}(x) = 2(3k + a); x = k$$
$$f^{'''}(x) = 6; x = k$$

Putting $f^{''}(k) = 0$, we get;

$k = \frac{-a}{3}$.

Thus

$$p^1(\frac{-a}{3}) = b - \frac{-a}{3}$$
$$f(\frac{-a}{3}) = c - \frac{ba}{3} + \frac{2a^3}{27}$$

As well, $x = y - \frac{a}{3}$ so (3.7) became;

$$y^3 + py + q = 0 \quad (3.8)$$

where

$$p = b - \frac{a^3}{3}$$
$$q = c - \frac{ba}{3} + \frac{2a^3}{27}$$

To solve equation (3.8), we use substitution $y = u + v$, it becomes;

$$u^3 + v^3 + (p + 3uv)(u + v) + q = 0 \tag{3.9}$$

Putting $uv = -\frac{p}{3}$, we obtain;

$$u^3 + v^3 = -q$$

So, the solution to equation (3.8) is equivalent to the solution to the two equations in (3.10a) and (3.10b) below:

$$uv = -\frac{p}{3} \tag{3.10a}$$

$$u^3 + v^3 = -q \tag{3.10b}$$

By cubing the equation in (3.10a), we get;

$$u^3v^3 = -\frac{p^3}{27} \tag{3.11}$$

From (3.10b) and (3.111), we notice that we have the product and the sum of two roots of a quadratic equation, and the two roots are u^3, v^3, and the equation is

$$t^2 + qt - \frac{p^3}{27} = 0$$

Assume that $A = u^3, B = v^3$, where

$$A = \frac{-q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}$$

$$B = \frac{-q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}$$

When finding the values of A, B, we get the values of u, v where;

$$u = \sqrt[3]{A}$$

$$u = w\sqrt[3]{A}$$

$$u = w^2\sqrt[3]{A}$$

where $w = (-\frac{1}{2}, \frac{2\sqrt{3}}{2})$ is the cube root of the integer one. Also, we get the value of v in;

$$v = \sqrt[3]{B}$$
$$v = w\sqrt[3]{B}$$
$$v = w^2\sqrt[3]{B}$$

As, we have the condition $uv = \frac{-p}{3}$, not all values are suitable to u. Now, let us check the value $\sqrt[3]{B}$.

$\sqrt[3]{A}\sqrt[3]{B} = \frac{-p}{3}$.

Thus the suitable values of v to $u = \sqrt[3]{A}, u = w\sqrt[3]{A}, u = w^2\sqrt[3]{A}$ are $\sqrt[3]{B}, w^2\sqrt[3]{B}, w\sqrt[3]{B}$ respectivily. Finally, the solution to (3.8) is

$$y_1 = \sqrt[3]{A} + \sqrt[3]{B}$$
$$v = w\sqrt[3]{A} + w^2\sqrt[3]{B}$$
$$v = w^2\sqrt[3]{A} + w\sqrt[3]{B}$$

These values are known as Cardian values (Cardano, 2002) with the aim p, q are ral numbers then the kind of the roots y_1, y_2, y_3 are depend on the function:

$\Delta = 4p^3 + 27q^2, \forall\Delta \in \mathbb{R}$.

(i) $\Delta > 0$. In this case it is

$\sqrt{\frac{q^2}{4} + \frac{p^3}{27}} = \sqrt{\frac{\Delta}{108}}$. The value of the Δ will be a real and positive. Thus the value of A, B are real, and $\sqrt[3]{A}$ is the third real root for A. Since q is a real, and $\sqrt[3]{A}\sqrt[3]{B} = \frac{-p}{3}$.

Also, $\sqrt[3]{B}$ will be a third real root for B.

Thereby, (3.8) has a real root;

$y_1 = \sqrt[3]{A}\sqrt[3]{B}$.

While the other two roots are complex and conjugate.

(ii) $\Delta < 0$. In this case it is

$\sqrt{\frac{q^2}{4} + \frac{p^3}{27}} = i\sqrt{\frac{-\Delta}{108}}$ purely imaginary and both quantities;

$$A = \frac{-q}{2} + i\sqrt{\frac{-\Delta}{108}}$$
$$B = \frac{-q}{2} - i\sqrt{\frac{-\Delta}{108}}$$

they are complex numbers, so the roots of the equation are in (3.8) can be expressed by using the cube roots of complex numbers, so if we assume that;

$\sqrt[3]{A} = a + ib$ is one of the cubic roots of A. Since B is a conjugate for A hence $a - ib$ is one of roots of B So that satisfies the relation: $\sqrt[3]{A}\sqrt[3]{B} = \frac{-p}{3}$.

Or,

$$\sqrt[3]{A} = a + ib$$
$$\sqrt[3]{B} = a - ib$$

From Cardano (2002) it produces the roots:

$$y_1 = 2a$$
$$y_2 = (a + ib)w + (a - ib)w^2 = -a - b\sqrt{3}$$
$$y_3 = (a + ib)w^2 + (a - ib)w = -a + b\sqrt{3}$$

are real and unequal. Obviously if $y_2 \neq y_3$, then $y_1 = y_2$, it implies that

$b = -a\sqrt{3}$.

Or,

$\sqrt[3]{A} = a(1 - i\sqrt{3})$.

Thus, $A = a^3(1 - i\sqrt{3})^3 = -8a^3$ is a real number, and this contradicts the claim that A is a complex number. That why, $y_1 \neq y_2$ so as $y_1 \neq y_3$.

Example 3.8 Solve the equation $x^3 + x^2 - 2 = 0$

Solution: Using transform $x = y - frac13$ the equation transforms into the form;

$y^3 - \frac{1}{3}y - \frac{52}{27} = 0,$
where $\Delta = \frac{52^2}{27} - \frac{4}{27}, p = -\frac{1}{3}, q = -\frac{52}{27}.$
Thus, $\sqrt{\frac{\Delta}{108}} = \frac{5}{\sqrt{27}}.$
We conclude that;

$$A = \frac{26}{27} + \frac{5\sqrt{27}}{27}$$
$$B = \frac{26}{27} - \frac{5\sqrt{27}}{27}$$

Thus,

$$\sqrt[3]{A} = \sqrt[3]{\frac{26}{27} + \frac{5\sqrt{27}}{27}} = \frac{1}{3}\sqrt[3]{26 + 5\sqrt{3}}$$
$$\sqrt[3]{B} = \frac{1}{3}\sqrt[3]{26 - 5\sqrt{3}}$$

We obtain:

$$y_1 = \frac{1}{3}(\sqrt[3]{26 + 15\sqrt{3}} + \sqrt[3]{26 - 15\sqrt{3}})$$
$$y_2 = -\frac{1}{6}(\sqrt[3]{26 + 15\sqrt{3}} + \sqrt[3]{26 - 15\sqrt{3}}) + \frac{i\sqrt{3}}{6}(\sqrt[3]{26 + 15\sqrt{3}} - \sqrt[3]{26 - 15\sqrt{3}})$$
$$y_3 = -\frac{1}{6}(\sqrt[3]{26 + 15\sqrt{3}} + \sqrt[3]{26 - 15\sqrt{3}}) - \frac{i\sqrt{3}}{6}(\sqrt[3]{26 + 15\sqrt{3}} - \sqrt[3]{26 - 15\sqrt{3}})$$

So, the roots of the original equation are:

$$x_1 = \frac{1}{3}(\sqrt[3]{26 + 15\sqrt{3}} + \sqrt[3]{26 - 15\sqrt{3} - 1})$$
$$x_2 = -\frac{1}{6}(\sqrt[3]{26 + 15\sqrt{3}} + \sqrt[3]{26 - 15\sqrt{3} + 5}) +$$
$$\frac{i\sqrt{3}}{6}(\sqrt[3]{26 + 15\sqrt{3}} - \sqrt[3]{26 - 15\sqrt{3}})$$
$$x_3 = -\frac{1}{6}(\sqrt[3]{26 + 15\sqrt{3}} + \sqrt[3]{26 - 15\sqrt{3} + 2})$$
$$-\frac{i\sqrt{3}}{6}(\sqrt[3]{26 + 15\sqrt{3}} - \sqrt[3]{26 - 15\sqrt{3}})$$

But, the equation
$x^3 + x^2 - 2 = 0$ has the root 1, and its two reminder roots are:
$-1 \mp -$.
By comparing these roots with Cardan's forms, we get that:

$$\sqrt[3]{26 + 15\sqrt{3}} + \sqrt[3]{26 - 15\sqrt{3}} = 4$$
$$\sqrt[3]{26 + 15\sqrt{3}} - \sqrt[3]{26 - 15\sqrt{3}} = 2 - \sqrt{3}$$

From these two relations we obtain:

$$\sqrt[3]{26 + 15\sqrt{3}} = 2 + \sqrt{3}$$
$$\sqrt[3]{26 - 15\sqrt{3}} = 2 - \sqrt{3}$$

Example 3.9 Solve the equation $y^3 - 3y + 1 = 0$.
Solution: In this case, we have;

$$p = -3, q = 1, \Delta = -81$$
$$\sqrt{\frac{-\Delta}{108}} = \frac{\sqrt{3}}{2}$$
$$A = \frac{1}{2} + i\frac{\sqrt{3}}{2} = w$$
$$B = -\frac{1}{2} - i\frac{\sqrt{3}}{2} = w^2$$

Thus, the roots are:

$$y_1 = \sqrt[3]{w} + \sqrt[3]{w^2}$$
$$y_2 = w\sqrt[3]{w} + w\sqrt[3]{w^2}$$
$$y_3 = w\sqrt[3]{w} - w\sqrt[3]{w^2}$$

Since the roots $y_{1,2}, y_3$ must be real numbers, although they did not appear here in the form of real numbers, thus we face the real challenge of finding the cube root of the amount w.

3.14 Exercises

Solve the following questions:

Q1: Solve the equations;
(1) $x^3 - 6x + 6 = 0$.
(2) $3x^3 - 6x - 2 = 0$.
(3) $x^3 - 2x + 2 = 0$.
(4) $2x^3 - 3x + 5 = 0$.
(5) $8x^3 + 12x^2 + 30x - 3 = 0$.

Q2: Prove that;
(1) $\sqrt[3]{\sqrt{5}+2} - \sqrt[3]{\sqrt{5}-2} = 1$.
(2) $\sqrt{\sqrt{243}+\sqrt{242}} - \sqrt{\sqrt{243}-\sqrt{242}} = i\sqrt{2}$.

3.15 Solving quartic equations

A quartic equation is one which can be expressed as a quartic function equaling zero. The general form of a quartic equation is

$$ax^4 + bx^3 + cx^2 + dx + e = 0, a \neq 0$$

The quartic is the highest order polynomial equation that can be solved by radicals in the general case (Chávez-Pichardo et al., 2022).

Ferrari (Candido, 1941; Masotti, 1960; Stewart, 2022; Chávez-Pichardo et al., 2022; Chávez-Pichardo et al., 2023) found the general algebraic solution to equations of the fourth degree, and his solution algorithm was as follows:

Consider the equation in (3.12) below:

$$x^4 + ax^3 + bx^2 + cx + d = 0 \tag{3.12}$$

Transform the equation as;

$$x^4 + ax^3 = -bx^2 - cx - d$$

Add $\frac{a^2}{4}x^2$ to both sides of the equation to get;

$$(x^2 + \frac{a}{2}x)^2 = (\frac{a^2}{4} - b)x^2 - cx - d \tag{3.13}$$

It is an equivalent equation for the equation (3.12). If the right side of this equation is a perfect square, then the solution is produced directly, and on the contrary, we add the expression $y(x^2 + \frac{a}{2}x) + \frac{y^2}{4}$ to both sides of the equation (3.13).

Thus, we get a perfect square on the left side of the variable y.

$$(x^2 + \frac{a}{2}x + \frac{y}{2})^2 = (\frac{a^2}{4} - b + y)x^2 + (-c + \frac{1}{2}ay)x + (-d + \frac{1}{4}y^2) \quad (3.14)$$

Now, we have to find the value of y so that the amount;

$$(\frac{a^2}{4} - b + y)x^2 + (-c + \frac{1}{2}ay)x + (-d + \frac{1}{4}y^2)$$

becomes a square for th linear expression,

$$ex + f$$

We know that if

$$Ax^2 + Bx + c = (ex + f)^2 \quad (3.15)$$

then

$$B^2 - 4AC = 0$$

Otherwise, (3.15) is equivalent to the three relations;

$$\begin{aligned} A &= e^2 \\ B &= 2ef \\ c &= f^2 \end{aligned} \quad (3.16)$$

Therefore, the righr hand side of (3.14) is a perfect square for the linear expression $ex + f$ of fulfills the equation;

$$(\frac{1}{2}ay - c)^2 = 4(y + \frac{a^2}{4} - b)(\frac{1}{4}y^2 - d)$$

Or,

$$y^3 - by^2 + (ac - 4d)y + 4bd - a^2d - c^2 = 0 \quad (3.17)$$

Now, we will take the value of y any root of the cubic roots of (3.17), say y_1 which called equation analyzer, and therefore we find that:

$$(x^2 + \frac{a}{2}x + \frac{1}{2}y_1)^2 = (ex + f)^2 \tag{3.18}$$

This can be transformed into two quadratic equations;

$$x^2 + \frac{a}{2}x + \frac{1}{2}y_1) = ex + f$$
$$x^2 + \frac{a}{2}x + \frac{1}{2}y_1) = -ex - f$$

By solving these two last equations, we get the required roots of the equation (3.12).

It is worth noting that Descartes (1954) has come up with an alternative solution to this method, and it has been called the Descartes' method.

Example 3.10 Solve $x^4 + 4x - 1 = 0$.

Solution: $a = b = 0, c = 4, d = -1$.

The cubic equation for solving is

$y^3 + 4y - 16 = 0$.

It has the rationality root 2. Putting $y = 2$, the expression (3.18) becames

$(x^2 + 1)^2 = (\sqrt{2x} - \sqrt{2})^2$.

The roots are

$x^2 + 1 = \sqrt{2x} - \sqrt{2}2$,

$x^2 + 1 = -\sqrt{2x} + \sqrt{2}2$.

Thus, the root values are;

$\frac{1 \mp i\sqrt{\sqrt{8}+1}}{2}, \frac{-1 \mp i\sqrt{\sqrt{8}-1}}{2}$.

3.16 Exercises

Solve the following questions:

Q1 $x^4 - 8x^2 - 4x + 3 = 0$.

Q2 $x^4 + 4x^2 + 4x - 3 = 0$.

Q3 $x^4 - x^2 - 2x - 1 = 0$.

Q4 $x^4 + x^3 + 5x^2 + 5x + 12 = 0.$
Q5 $[(x+2)^2 + x^2]^3 = 8x^4(x+2)^2.$

4

Numerical Solution of Nonlinear Equations

4.1 Introduction

In polynomial equations of the fifth degree or more, as well as equations that include transcendental functions, such as $logx, sinx, cosx...etc$, there is difficulty in finding roots in terms of coefficients and simple algebraic methods. Therefore, it is necessitated searching for numerical methods to find approximate values for the roots of these equations (Lazard, 2009; Billings, 2013; Atkinson, 1991; Mathews, 1992).

Let us consider the following equation;

$$f(x) = 0$$

if α is a root of the equation, then we are looking for approximate values α_n for this root, such that

$$|\alpha - \alpha_n| < \delta$$
$$|f(\alpha_n)| < \epsilon$$

where δ, ϵ are positive amounts.

Before proceeding and delving into the methods of finding approximate solutions to the roots of such equations, we explore some rules, such as: Descartes rule of signs (Anderson, 1979; Wang, 2004), Horner's method for removing roots from polynomials (Horner, 1833; Horner, 1819), and Finding the differential via Horner's method (Cajori, 1911; Volkov, 1990) which would help us to find approximate solutions to nonlinear equations.

4.2 Auxiliary rules for finding approximate roots

A root-finding algorithm is an algorithm for finding zeros, also called roots, of continuous functions. A zero of a function f, from the real numbers to real numbers or from the complex numbers to the complex numbers, is a number x such that $f(x) = 0$. In general, the zeros of a function cannot be computed exactly nor expressed in closed form, root-finding algorithms provide approximations to zeros, expressed either as floating-point numbers or as small isolating intervals, or disks for complex roots. Or, an interval or disk output being equivalent to an approximate output together with an error bound (Vetterling et al., 1992; Press, 1992).

4.2.1 Descartes' rule of signs

This rule is used to find the number of real zeros of polynomials whose coefficients are real numbers. The method relies mainly on counting the number of changes in the signs of non-zero coefficients in the order. Let us assume that the number of signs changing in order is v.

For example, the signs of the coefficients of terms in th polynomial: $p_1(x) = 2x^3 2x - 5$ is in order $(+, +, -)$, so $v = 1$. In $p_2(x) = 3x^4 - 81x^2 - 300x - 445$ is in order $(+, +, -, -, -)$, in this case $v = 1$ also, while in $p_3(x) = 4x^4 - 8x^3 + 81x^2 - 300x - 445$ is in order $(+, -, +, -, -)$, in this case $v = 3$.

4.2.2 Methodology of using Descartes' rule of signs

Supose that k be the number of positive real zeros of the polynomial, so $k \leq v$, and $v - k$ should be an even integer that is positive or equal to zero.

Based on the rule $v = 1$ and in the case $p_1(x)$, so $k = 1$ and the number of the real roots is just one root. By the same way, we conclude that $k = 1$ in the case of the polynomial $p_2(x)$. And, $k = 1$ or $k = 3$ in the case of $p_3(x)$, and Neither of them can be determined in this way.

It is easy to prove that if r is a zero of the polynomial $p(x)$, then $-r$ is a zero of the polynomial $p(-x)$. Thus, we can find the number of negative real zeros by applying $p(-x)$ If we substitute of all x by $-x$ in $p_1(x), p_2(x)$ we get;

$p_1(-x) = -2x^3 - 2x - 5,$

$p_2(-x) = 3x^4 - 8x^3 - 81x^2 + 300x - 445.$

In the case $p_1(-x)$, the boundary signs, in order, are; $(-, -, -)$ so $v = 0$, and $p_1(x)$ It has no negative real zeros. And it can be concluded that $p_1(x)$ has one positive real number, and couple complex root (conjugate).

In what related to the polynomial $p_2(-x)$ the signs, in order are; $(+, -, -, +, -)$, thereby $v = 3$, and;

$v - k = 0$, or

$v - k.$

Thus, $p_2(x)$ either it has one negative root, or three negative real roots.

4.2.3 Horner's method for removing roots from polynomials

Horner's method enables us to remove any root from the roots of a polynomial equation , which can be summarized as follows:

Suppose the polynomial;

$$P(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + ... + a_n = 0$$

has a root α, it is possible to find a polynomial of degree $(n-1)$ such that

$$(x-\alpha)(x^{n-1} + b_1 x^{n-2} + b_2 x^{n-3} + b_{r-1} x^{n-r} + b_r x^{n-r-1} + ... + b_{n-1})$$
$$\equiv x^n + a_1 x^{n-1} + a_2 x^{n-2} + ... + a_r x^{n-r} + ... + a_n$$

Now, by comparing the coefficient of x^{n-r} we get:

$$b_r = a_r + \alpha_{r_1}$$
$$b_0 = 1$$

For all values of r, from these relations we get;

$$b_1 = a_1 + \alpha$$
$$b_2 = a_2 + a_1\alpha + \alpha^2$$
$$.$$
$$.$$
$$.$$
$$b_n = a_n + a_{n-1}\alpha + a_{n-2}\alpha^2 + ... + \alpha^n = P(\alpha) = 0$$

To find these values, we follow the following process as shown in Table 4.1:

Table 4.1: Horner's method for removing roots from polynomials

1	a_1	$a_2...$	a_n	α
	α	$a_1\alpha + \alpha^2...$		
1	$a_1 + \alpha$	$a_2 + a_1\alpha + \alpha^2...$	b_n	

If α is not a root then;

$$P(\alpha) = b_n \neq 0$$

This process can be used to find the numerical value of $P(x), \forall x$. Whenever one of the roots of the equation is removed, we obtain a new equation of degree $(n-1)$.

Example 4.1 The polynomial $2x^4 + 2x^3 - x - 2 = 0$ has two roots $x = -2, x = 1$. it is required to remove them, and find the resulting equation.

Solution: After performing the first operation (dividing by $x-1$), it produces the numbers in the third line, which are the coefficients of a third-degree polynomial equation;

$x^3 + 3x^2 + 3x + 2 = 0.$

Table 4.2: Horner's method for removing roots from polynomials

1	2	0	-1	-2	1
	1	3	3	2	Operation I
1	3	3	2	0	-2
	-2	-2	-2		Operation II
1	1	1	0		

And, after performing the first operation (dividing by $x+2$), it produces the numbers in the third line, which are the coefficients of a second-degree polynomial equation;

$x^2+x+1=0.$

Example 4.2 The polynomial $2x^5-6x^4+8x^3+4x^2-20x-8=0$ has two roots $x=2, x=-1$. it is required to remove them, and find the resulting equation.

Solution: $2x^5-6x^4+8x^3+4x^2-20x-8=x^5-3x^4+4x^3+2x^2-10x-4=0$

Table 4.3: Horner's method for removing roots from polynomials

1	-3	4	2	-10	-4	2
	2	-2	4	12	4	Operation I
1	-1	2	6	2	0	
	-1	2	-4	-2		Operation II
1	-2	4	2	0		

After performing the first operation (dividing by $x-2$), it produces the numbers in the third line, which are the coefficients of a forth-degree polynomial equation;

$x^4-x^3+2x^2+6x+2=0.$

And, after performing the first operation (dividing by $x+1$), it produces the numbers in the third line, which are the coefficients of a third-degree polynomial equation;

$x^3-2x^2+4x+2=0.$

Example 4.3 Find the value of $P(0.5)$ by usung Horner's method, if $P(x) = 7x^4 - 2x^3 + 3x + 8 = 0$.

Solution: We arrange the problem in Table 4.4 as:

Table 4.4: Horner's method for removing roots from polynomials

4	-6	0	3	-5	0.5
	2	-2	-1	1	
2	-2	-1	1	-4	

The quotient is $2x^3 - 2x^2 - x + 1$, and $P(-\frac{1}{2}) = -4$.

4.2.4 Finding the differential via Horner's method

It had been proved that by repeatedly dividing a polynomial of degree n by the linear term $(x - \alpha)$ we obtain the numerical value of the polynomial and its differentiation at $x = \alpha$ (Horner, 1833; Horner, 1819; Cajori, 1911; Pan, 1997; Sitton et al., 2003).

Suppose that

$$P(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + ... + a_n \tag{4.1}$$

Dividing by $(x - \alpha)$, to get

$$P(x) = (x - \alpha)q_1(x) + P(\alpha) \tag{4.2}$$

where $q_1(x)$ is a polynomial in degree $(n - 1)$, and dividing it over $(x - \alpha)$, we get;

$$q_1(x) = (x - \alpha)q_2(x) + q_1(\alpha)$$

where $q_2(x)$ is a polynomial in degree $(n-2)$, and in compensation for $q_1(x)$ in (4.2), the result will be;

$$P(x) = P(\alpha) + (x - \alpha)q_1(\alpha) + (x - \alpha)^2q_2(x) \tag{4.3}$$

Again, dividing $q_2(x)$ by $(x - \alpha)$, and in compensation for (4.3), we obtain;

$$P(x) = P(\alpha) + (x-\alpha)q_1(\alpha) + (x-\alpha)^2 q_2(\alpha) + (x-\alpha)^3 q_3(x)$$

where $q_3(x)$ is a polynomial in degree $(n-3)$, and by repeating this process n times we get;

$$P(x) = P(\alpha)+(x-\alpha)q_1(\alpha)+(x-\alpha)^2 q_2(\alpha)+(x-\alpha)^3 q_3(\alpha)+...+(x-\alpha)^n q_n \tag{4.4}$$

where $q_n = a_0$ is a constant.

Using the Taylor's expansion (Thomas et al., 2010) for the polynomial $P(x)$ near $x = \alpha$, we get:

$$P(x) = p(\alpha)+\frac{(x-\alpha)}{1!}p^{(1)}\alpha+\frac{(x-\alpha)^2}{2!}p^{(2)}\alpha+...+\frac{(x-\alpha)^n}{n!}p^{(n)}\alpha \tag{4.5}$$

where $n! = n(n-1)(n-2)...3.2.1$, $p^{(n)}(x) = (\frac{d^n p}{dx^n})x = \alpha$.

The series (4.5) is finat because $p(x)$ is polynomial in the degree n, thus;

$$\begin{aligned} &P^{(n+1)}\alpha = 0 \\ &P^{(n+2)}\alpha = 0 \\ &. \\ &. \\ &. \\ &P^{(n+r)}\alpha = 0, r \geq 1 \end{aligned}$$

By comparing the terms in the equations (4.4) and (4.5), we find that:

$$q_1(\alpha) = \frac{p^{(1)}\alpha}{1!}$$
$$q_2(\alpha) = \frac{p^{(2)}\alpha}{2!}$$
$$q_3(\alpha) = \frac{p^{(3)}\alpha}{3!}$$
$$.$$
$$.$$
$$.$$
$$q_n(\alpha) = \frac{p^{(n)}\alpha}{n!}$$

Generally;

$$q_j(\alpha) = \frac{p^{(j)}\alpha}{j!}, j = 1, 2, 3, ..., n.$$

Now, it can be used Horner's method, to find $p(z_0), p^1(z_0)$, where $z_0 = x_0 + iy_0$ is a complex number, as follows:

Assume that;

$$P(z) = a_0 z^n + a_1 z^{n-1} + a_2 z^{n-2} + ... + a_{n-1} z + a_n$$

It can be finding the polynomials:

$$q(z) = b_0 z^{n-2} + b_1 z^{n-3} + ... + b_{n-2}$$
$$d(z) = (z - z_0)(z - z_0) = z^2 - 2x_0 z + x^2 0 + y^2 0$$

where

$$p(z) = q(z)d(z) + b_{n-1} z + b_n \tag{4.6}$$

In compensation for $p(z), q(z), d(z)$ in (4.6), and compare cofficients, we find that:

$$b_k = a_k + 2x_0 b_{k-1} + (-x_0^2 - y_0^2) b_{k-2}, k = 1, 2, 3, ..., n - 1$$
$$b_0 = a_0, b_{-1} = 0$$
$$b_n = a_n + (-x_0^2 - y_0^2) b_{n-2}$$

Furthermore,

$$P(z_0) = b_{n-1}z_0 + b_n$$
$$p'(z_0) = 2iy_0(c_{n-3}z_0 + c_{n-2}) + b_{n-1}$$

where

$$c_{-1} = 0, c_0 = b_0$$
$$c_k = b_k + 2x_0c_{k-1} + (-x_0^2 - y_0^2)c_{n-2}, k = 1, 2, 3, ..., n-3$$
$$c_{n-2} = b_{n-2} + (-x_0^2 - y_0^2)c_{n-4}$$

Example 4.4 If $p(x) = x^3 + 2x^2 - 3x + 1$ then find each of $p(2), p^{(3)}(2)$.

Solution: It can be found $p(2), p^{(3)}(2)$ by utilizing Horner's method table as follows:

Table 4.5: Finding the differential via Horner's method

1	2	-3	1	2
	2	8	10	
1	4	5	11	p(2)= 11
	2	12		
1	6	17		$17 = \frac{p^{(1)}(2)}{1!}$
	2			
1		8		$8 = \frac{p^{(2)}(2)}{2!}$
1				$1 = \frac{p^{(3)}(2)}{3!}$
				$p^{(3)}(2) = 6$

Example 4.5 Consider the polynomial $p(x) = x^3 + 3x - 1$, and evaluate each of $p(1+i), p^{(1)}(1+i)$.

Solution:

$$x_0 = y_0 = 1$$
$$a_0 = 1, a_1 = 0, a_2 = 3, a_3 = -1$$
$$P(1+i) = b_2(1+i) + b_3$$
$$b_{-1} = 0, b_0 = a_0 = 1$$
$$b_1 = a_1 + 2b_0 - 2b_{-1} = a_1 + 2 = 0 + 2 = 2$$
$$b_2 = a_2 + 2b_1 - 2b_0 = 3 + 4 - 2 = 5$$
$$b_3 = a_3 - 2b_1 = -1 - 2 = -3$$
$$\therefore p(1+i) = 5(1+i) - 3 = 2 + 5i$$
$$\because P^{(1)}(1+i) = 2i(c_0(1+i) + c_1) + b_2 = 2i(1 + i + c_1) + 5$$
$$c_1 = b_1 - 2c_{-1} = b_1 = 2$$
$$\therefore p^{(1)}(i+1) = 2i(3+i) = -2 + 6i$$

4.3 Numerical methods for finding approximate values of the roots of equations

This section deals with some numerical methods for finding approximate values for the real or complex roots of polynomial equations or equations of transcendental functions. Most of these methods require finding the maximum or minimum, or both of the roots. Therefore, the following lemma is of great importance before embarking on the various methods.

Lemma *Assume that α is a root of the following equation polynomial:*
$P(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + ... + a_n = 0.$
If $\lambda = max\,|a_i|\,, i = 1, 2, ..., n$ *then* $-1 - \lambda \leq \alpha \leq 1 + \lambda.$

4.3.1 Bisection method

The bisection method is a root-finding method that applies to any continuous function for which one knows two values with opposite signs. The method stated based on repeatedly bisecting the defined interval and then selecting the subinterval in which the function changes sign, and therefore must contain a root, as shown in Figure 4.1.

The method is a very simple, robust, and relatively slow. It often used to obtain a rough approximation to a solution, and used as a starting point for more rapidly converging methods, and the method has more than one name (Burden and Faires, 1985). The algorithm of the method can be described step by step as follows:

Let us assume that $f(x)$ It is a continuous function on the closed interval $[b_0, c_0]$, and it defined as:

$$\begin{aligned} &f(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + ... + a_n = 0 \\ &a_0 = -1 - \lambda \\ &b_0 = 1 + \lambda \\ &f(a_0) \cdot f(b_0) \leq 0 \end{aligned}$$

Thus,

- Step$_1$: Put $j = 0$.
- Step$_2$: Put $c_j = \frac{a_j + b_j}{2}$.
- Step$_3$: If $f(a_j) \cdot f(c_j) \begin{cases} \leq 0; a_j = a_{j+1}, c_j = b_{j+1} \\ > 0; c_j = a_{j+1}, b_j = b_{j+1} \end{cases}$
- The function $f(x)$ has a root in the Nested Interval $[a_{j+1}, b_{j+1}] = [a_1, b_1] \subseteq [a_0, b_0]$ (Königsberger, 2004; Fridy, 2000).
- Put $j = 1$, and repeat from Step$_2$ until $|c_{j+1} - c_j| \leq e$, where e is small enough.

Example 4.6 Find the negative root of $x^3 - 21x + 3500 = 0$, by the bisection method, correct to three decimal places.

Solution:

$$\begin{aligned} f(x) &= x^3 - 21x + 3500 = 0 \\ f(-x) &= -x^3 + 21x + 3500 = 0 \end{aligned}$$

The negative root of $f(x) = 0$ is a positive root of $f(-x) = 0$. Therefore, we have to find the positive root of $f(-x) = 0$.

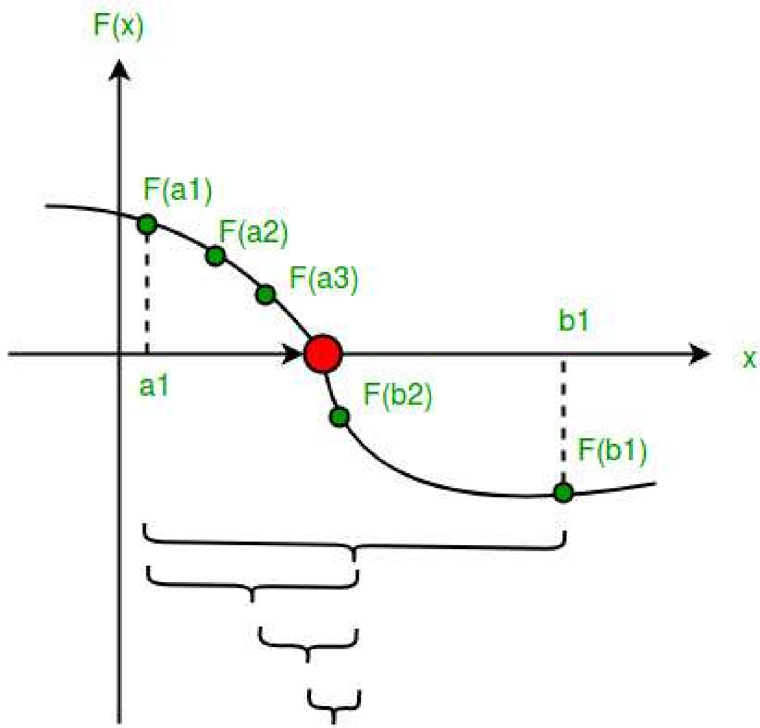


Figure 4.1: Bisection method

Let us take,

$$\begin{aligned}
&f(-x) = \psi(x) = -x^3 + 21x + 3500 = 0\\
&\psi(x) = x^3 - 21x - 3500 = 0\\
&\psi(0) = -ve\\
&\psi(1) = -ve\\
&\psi(2) = -ve\\
&.\\
&.\\
&.\\
&\psi(14) = -ve\\
&\psi(15) = -ve\\
&\psi(16) = +ve
\end{aligned}$$

Thus, a root lies between $[15, 16]$.
By putting;

$$\begin{aligned}
x_0 &= \frac{15+16}{2}\\
&= 15.5\\
f(15.5) &= -ve
\end{aligned}$$

Thus, a root lies between $[15.5, 16]$.

By repeating the algorithm, we will find that the approximate positive root is 15.644. Thereby, the positive root of $\psi = 15.644$, hence the negative of the equation is -15.644, as shown in Table 4.6.

Example 4.7 Use bisection method to find roots of the equation $x^3 - x - 1 = 0$.

Solution: Since the coefficients are $a_1 = 0, a_2 = -1, a_3 = -1$, hence

$\lambda = max\,|a_j| = 1; j = 1, 2, 3.$

Thereby, the roots of the equation are exists in $[-2, 2]$. Or, $a_0 = -2, b_0 = 2$.

Bisides,

Table 4.6: Bisection method iterations

j	x_1	x_2	$f(x_1)$	$f(x_2)$	x	$f(x)$
1	15	16	$-ve$	$+ve$	15.5	$+ve$
2	15.5	16	$-ve$	$+ve$	15.75	$+ve$
3	15.5	15.75	$-ve$	$+ve$	15.6250	$-ve$
4	15.625	15.75	$-ve$	$+ve$	15.6875	$+ve$
5	15.625	15.6875	$-ve$	$+ve$	15.6563	$+ve$
6	15.625	15.6563	$-ve$	$+ve$	15.6407	$-ve$
7	15.6407	15.6563	$-ve$	$+ve$	15.6485	$+ve$
8	15.6407	15.6485	$-ve$	$+ve$	15.6446	$+ve$
9	15.6407	15.6446	$-ve$	$+ve$	15.6427	$-ve$
10	15.6427	15.6460	$-ve$	$+ve$	15.6444	$+ve$
11	15.6427	15.6444	$-ve$	$+ve$	15.64355	

$$f(a_0) \cdot f(b_0) \leq 0$$

Now, we are follow the following steps:

(i) Put $j = 0$

(ii) Utilize $c_j = \frac{a_j + b_j}{2} = \frac{a_0 + b_0}{2} = 0$.

(iii) Since, $f(a_0) \cdot f(c_0) = f(-2) \cdot f(0) > 0$, hence $a_1 = c_0 = 0, c_0 = c_1 = 2$.

Therby, the new interval is $[0, 2]$.

(iv) Put j= 1, and repeat the steps, the roots lie in the $[1, 2]$. By going on we find the root lies at $[1.3246, 1.3250]$. Or, the root is 1.3250. What related to the other two roots, we will obtaining them, by dividing the equation over $x_1 - 1.325$, and finally we solve the second degree equation by the constitution method.

Example 4.8 Find a root of $\frac{x^2}{4} - sinx = 0$ by using the bisection method.

Solution: As in the Example 4.6, we obtain the root as shown Table 4.7.

Table 4.7: Bisection method iterations

j	a_{j-1}	b_{j-1}	c_{j-1}	$f(c_j) \cdot f(a_j)$
1	1.5	2	1.75	$-ve$
2	1.75	2	1.8752	$-ve$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
13	1.93372	1.93384	1.93378	$+ve$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.

4.3.2 Newton-Raphson's method

Newton-Raphson Method is a powerful technique for solving equations numerically. It is most commonly used for approximation of the roots of the real functions. Newton Rapson Method was developed by Isaac Newton and Joseph Raphson (Atkinson, 1991; Süli and Mayers, 2003; Atkinson, 1991; Ypma, 1995). Newton Raphson Method involves iteratively refining an initial guess to converge it toward the desired root. In this section, we will study the method and its steps to calculate the roots, as well as the applications of it.

The Newton-Raphson method is an iterative numerical method used to find the roots of a real-valued function. This formula is named after Sir Isaac Newton and Joseph Raphson, as they independently contributed to its development. Newton Raphson Method is an algorithm to approximate the roots of zeros of the real-valued functions, using guess for the first iteration (x_0) and then approximating the next iteration (x_1) which is close to roots, using the following formula;

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

where, x_0 is the initial value of x, $f(x_0)$ is the value of the equation at initial value, and $f'(x_0) \neq 0$ is the value of the first order derivative of the equation or function at the initial value x_0. As shown in Figure

4.2.

In the general form, the Newton-Raphson method formula is written as follows:

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$$

where, x_{n-1} is the estimated $(n-1)$th root of the function, $f(x_{n-1})$ is the value of the equation at $(n-1)$th estimated root, and $f'(x_{n-1})$ is the value of the first order derivative of the equation or function at x_{n-1}.

Now, let us describe the algorithm step by step:

- Step$_1$: Draw a graph of $f(x)$ for different values of x as shown in Figure 4.2.
- Step$_2$: A tangent is drawn to $f(x)$ at x_0. This is the initial value.
- Step$_3$: This tangent will intersect the $X-$ axis at some fixed point $(x_1, 0)$ if the first derivative of $f(x)$ is not zero *i.e.* $f'(x_0) \neq 0$.
- Step$_4$: As this method assumes iteration of roots, this x_1 is considered to be the next approximation of the root.
- Step$_4$: Now steps $(2-4)$ are repeated until we reach the actual root x^*.

Note:

(i) Conditions required to apply the Newton-Raphson method.

The following conditions must be met when using the Newton-Raphson method to find the approximate roots of the equation $f(x) = 0, \forall x \in [a, b]$.

(a) $f(x) \neq 0$, and exists.

(b) Signal $f'(x), \forall x \in [a, b]$ does not change.

(c) If $f(a)\dot{f}(b) < 0$ then $\left|\frac{f(a)}{f'(a)}\right| < |b-a|$, and $\left|\frac{f(b)}{f'(b)}\right| < |b-a|$.

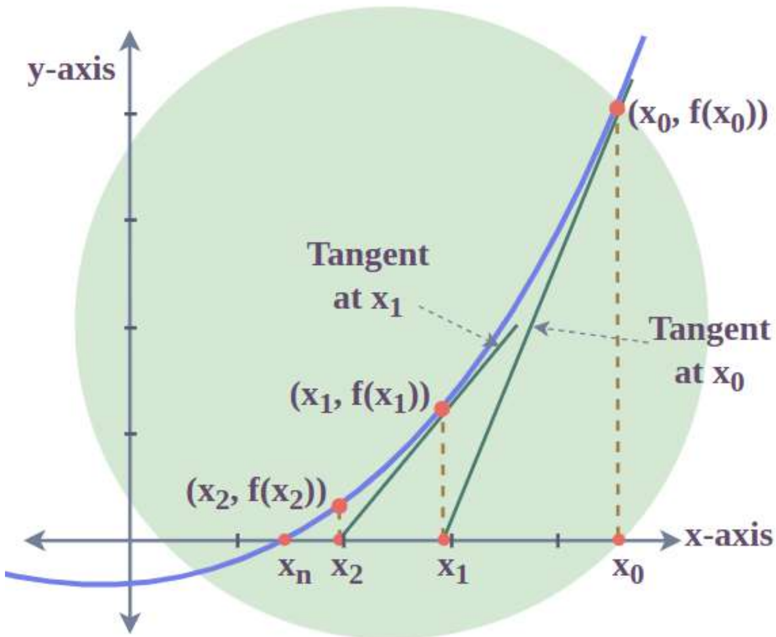


Figure 4.2: Newton-Raphson method

It should be noted that, (a) and (b) mean that $f(x) = 0$ has a unique root at $[a, b]$, while (c) means that there exist at least one root at $[a, b]$.

(ii) Slope-intercept equation of a line.

The slope-intercept equation of any line is represented as $y = mx + c$, Where m is the slope of the line and c is the $x-$intercept of the line. Using the same formula we, get;

$$y = f(x_0) + f'(x_0)(x - x_0)$$

Here $f(x_0)$ represents the c and $f'(x_0)$ represents the slope of the tangent m. As this equation holds true for every value of x, it must hold true for x_1. Thus, substituting x with $x - 1$, and equating the equation to zero as we need to calculate the roots, we get:

$$0 = f(x_0) + f'(x_0)(x_1 - x_0)$$
$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

Thus, Newton Raphson's method was mathematically proved and accepted to be valid.

(iii) Convergence of Newton Raphson method.

The necessary condition for the convergence of the Newton-Raphson method is:

$$|f(x) \cdot f''(x)| < |f'(x)|^2$$

Example 4.9 For the initial value $x_0 = 3$, approximate the root of $f(x) = x^3 - 3x + 1$.

Solution:

$$\begin{aligned}
&x_0 = 3, f(x) = x^3 - 3x + 1\\
&f'(x) = 3x^2 - 3\\
&f'(x_0) = 3(9) - 3 = 24\\
&f(x_0) = f(3) = 27 - 3(3) + 1 = 19\\
&x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}\\
&= 3 - \frac{19}{24}\\
&= 2\frac{5}{24}.
\end{aligned}$$

Example 4.10 For the initial value $x_0 = 2$, approximate the root of $f(x) = x^2 - 2 = 0$.

Solution:

$x_0 = 2, f(x) = x^2 - 2$

$f'(x) = 2x$

$f'(x_0) = 2(2) = 4$

$f(x_0) = f(2) = 2^2 - 2 = 4$

$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$

$x_1 = 2 - \frac{2}{4}$

$x_1 = \frac{3}{2}$

Using Newton Raphson method again :

$x_2 = \frac{17}{12}$

$x_3 = \frac{577}{408}$

Therefore, the root of the equation is approximately$x = \frac{577}{408} \approx 1.414.$

Example 4.11 Use $x_0 = 1.5$ as the initial point to find a root of

$$f(x) = sin(x) - \frac{x^2}{4}$$

Solution:

$$f(x) = sin(x) - \frac{x^2}{4}$$
$$f'(x) = cos(x) - \frac{x}{2}$$
$$x0 = 1.5$$
$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}, \forall n \in \mathbb{N}$$

Accoeding to Table 4.8, the root is $x_4 = 1.93375$.

Table 4.8: Newton Raphson method iterations

n	x_n	$f(x_n)$	$f'(x_n)$	$x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$
0	1.50000	0.434995	−0.67927	−0.64039
1	2.14039	−0.30319	−1.6095	0.1884
2	1.95201	−0.02437	−1.34805	0.0181
3	1.93393	−0.00023	−1.32217	0.00018
4	1.93375	0.000005	−1.3219	$\leq \frac{1}{2}10^{-5}$
.	.	.	.	.

Example 4.12 Find the nth root of number γ by using Newton Raphson.

Solution:

$$x = \sqrt[n]{\gamma}$$
$$f(x) = x^n - \gamma$$
$$f'(x) = nx^{n-1}$$
$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$
$$x_{k+1} = \frac{1}{n}[(n-1)x_k + \frac{\gamma}{x_k^{n-1}}]$$

For a radicand γ, beginning from some initial value x_0 and using Newton Raphson method repeatedly with successive values of k, one obtains after a few steps a sufficiently accurate value of $\sqrt[n]{\gamma}$ if x_0 was not very far from the searched root.

For cub root $\sqrt[3]{\gamma}$;

$$x_{k+1} = \frac{1}{3}[2x_k + \frac{\gamma}{x_k^2}]$$

For example, if one wants to compute $\sqrt[3]{2}$, and uses $x_0 = 1$, at the fifth step gives $x_5 = 1.259921049894873$.

Note: Newton Raphson method advantage and disadvantage.

(i) The method allows us to guess the roots of an equation with a small degree very efficiently and quickly.

(ii)
- The method tends to become very complex when the degree of the polynomial becomes very large.
- The difficulty when using Newton-Raphson's method to find roots is the need to find the numerical value of the derivative $f'(x)$ when $x = x_i; i = 0, 1, 2, ...$

4.3.3 Secant method

The use of an approximation of the derivative $f'(x)$ when $x = x_i, i = 1, 2, 3, ...$ by utilizing transversal straight (Fuzzy, 2009; Fusy, 2005; Vázquez-Ávila, 2021) that connects the two points $(x_{n-1}, f(x_{x-1})), (x_n, f(x_n))$ is to cover the disadvantage of Newton Raphson's method. Thus, the secant method is a root-finding algorithm that uses a succession of roots of secant lines to better approximate a root of a function f, as sketched in Figure 4.3. And, it can be thought of as a finite-difference approximation of Newton's method (Papakonstantinou and Tapia, 2013; Avriel, 1976).

The derivation of the method can be summarized as following:

- Step$_1$: Select two initial approximations x_0, x_1 to the root.
- Step$_2$: Calculate the function's values at these points, $i.e. f(x_0), f(x_1)$.

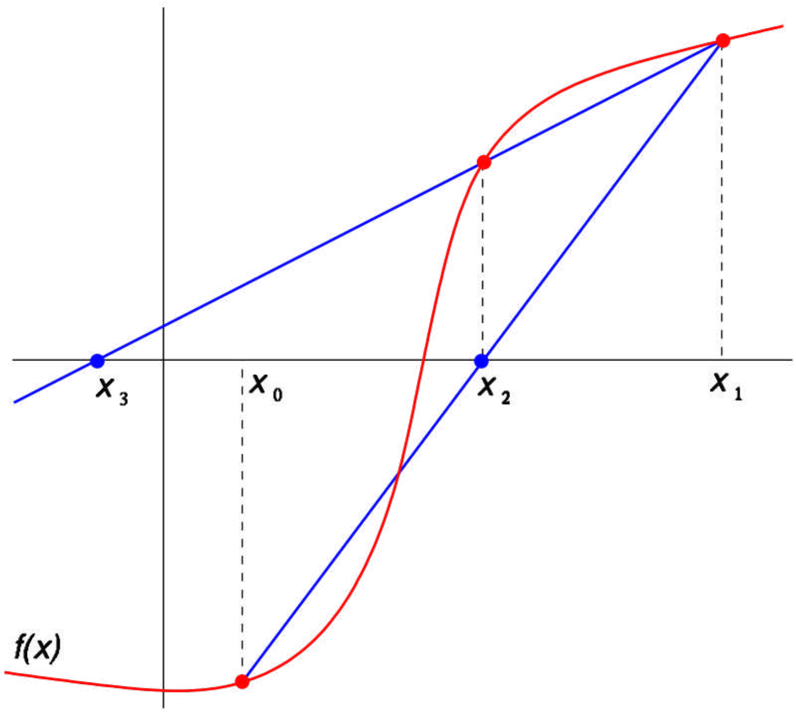


Figure 4.3: Secant method

- Step$_3$: Apply the Secant Method formula to find the next approximation x_2.
- Step$_4$: Repeat the process until an acceptable level of accuracy is reached or a maximum number of iterations is achieved.

Mathematically, assume that the initial values x_0, x_1, we construct a line through the points $(x_0, f(x_0)), (x_1, f(x_1))$, as shown in figure 4.3. In slope–intercept form, the equation of this line is

$$y = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

The root of this linear function, that is the value of x such that $y = 0$ is

$$x = x_1 - f(x_1 \frac{x_1 - x_0}{f(x_1) - f(x_0)}$$

we use this new value of x as x_2 and repeat the process, using x_1, x_2 instead of x_0, x_1 respectively. We continue this process, solving for x_3, x_4, etc., until we reach a sufficiently high level of precision. A sufficiently small difference between x_n, x_{n-1}:

$$x_2 = x_1 - f(x_1 \frac{x_1 - x_0}{f(x_1) - f(x_0)}$$
$$x_3 = x_2 - f(x_2 \frac{x_2 - x_1}{f(x_2) - f(x_1)}$$

.

.

.

$$x_n = x_{n-1} - f(x_{n-1}) \frac{x_{n-1} - x_{n-2}}{f(x_{n-1}) - f(x_{n-2})}$$

Note:

(i) The iterates x_n of the secant method converge to a root of f if the initial values x_0, x_1 are sufficiently close to the root.

(ii) Using the secant method needs two initial approximate values for the root, while the Newton-Raphson method requires just one initial approximation value.

Example 4.13 Find a root of the equation $2x - 2 - sinx$, where the root located between $0.4, 0.6$.

Solution:

$$x_0 = 0.4, x_1 = 0.6$$
$$f(x) = 2x - 2 - sinx$$
$$f(x_0) = 0.3891$$
$$f(x_1) = -0.4352$$
$$x_2 = 0.6 - \frac{(0.2)(-0.4352)}{0.8243} = 0.404$$
$$x_3 = 0.494 - \frac{(-0.106)(0.0042)}{0.4392} = 0.49501$$
$$f(x_3) = 0.00003$$
$$x_4 = 0.4950 - \frac{(0.001)(0.00003)}{-0.00417} = 0.4950$$

Thus, the desirable value is $x_4 = 0.4950$.

Example 4.14 Compute the root of the equation $x^2 e^{\frac{-x}{2}} = 1$ in the interval $[0, 2]$ using the secant method. The root should be correct to three decimal places.

Solution:

$$x0 = 1.42, x1 = 1.43, f(x0) = –0.0086, f(x1) = 0.00034$$
$$x_2 = x_1 – \frac{x_0 – x_1}{f(x_0) – f(x_1)} f(x_1)$$
$$= 1.43 – \frac{1.42 – 1.43}{0.00034 – (–0.0086)}(0.00034)$$
$$= 1.4296$$
$$f(x_2) = –0.000011(–ve)$$
$$x_3 = x_2 – \frac{x1 – x2}{f(x_1) – f(x_2)} f(x_2)$$
$$= 1.4296 – \frac{1.42 – 1.4296}{0.00034 – (–0.000011)}(–0.000011)$$
$$= 1.4292$$

The value of each x_2, x_3 matching up to three decimal places, thereyby, the required root is $x_3 = 1.429$.

4.3.4 Birge-Vieta method

Combining Newton-Raphson's method (Atkinson, 1991; Süli and Mayers, 2003; Atkinson, 1991; Ypma, 1995) to find the roots of the polynomial $f(x) = x^n + a_1x^{n-1} + ... + a_n = 0$ with Horner's method (Horner, 1833; Horner, 1819) to find $f(x), f'(x)$ values for all values of x in the closed interval $[a, b]$ is called Birge-Vieta method (Funkhouser, 1930; Vinberg, 2003; Djukić et al., 2011; Britannica et al., 1993). In other words, Newton-Raphson's method can be used to find a root of a polynomial equation via Horner's method.

Example 4.15 Find all roots of $x^4 - 5x^3 + 5x^2 + 5x - 7 = 0$, if the equation has a roots in the interval $[-1, 3]$.

Solution:

$$f(x) = x^4 - 5x^3 + 5x^2 + 5x - 7$$

By putting $x_0 = 3$, and find $f(3), f'(3)$ as shown in Table 4.9:

Table 4.9: Birge-Vieta method (i)

1	-5	5	5	-7	3
	3	-6	-3	6	
1	-2	-1	2	-1	f(3)= -1
	3	3	6		
1	1	2	8		$f'(3) = 8$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 3 - \frac{-1}{8} = 3.125$$

Now, by putting $x_1 = 3.125$, and repeating the previous process, we find that the result as shown in Table 4.10:

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 3.125 - \frac{0.238}{11.838} = 3.105$$

Table 4.10: Birge-Vieta method (ii)

1	-5	5	5	-7	3.125
	3.125	-5.859	-2.689	7.238	
1	-1.875	-0.859	2.316	0.238	f(3.125)= 0.238
	3.125	3.906	9.522		
1	1.250	3.047	11.838		$f'(3.125) = 11.838$

Again, put $x_3 = 3.1048$ as the absolut approximate value, and to find the other roots, repeat the process in the beginning and by utilizing $x_0 = -1$. And, the better is to reduction the orignal equation in the fourth degree to the thied degree by dividing $f(x)$ over $(x - 3.1048)$.

$$f_1(x) = x^3 - 1.895x^2 - 0.8840x + 2.2552$$

And, utilizing Horner's method, to find $f_1(-1), f_1'(-1)$, as shown in Table 4.11

Table 4.11: Birge-Vieta method (iii)

1	-1.895	-0.8840	2.2552	-1
	-1	2.895	-2.0110	
1	-2.895	2.0110	0.2442	$f_1(-1) = 0.2442$
	-1	3.895		
1	-3.895	5.9060		$f_1'(-1) = 5.9060$

$$x_1 = -1 - \frac{0.2442}{5.9060} = -1.041$$

Morever, as in Table 4.12:

$$x_2 = -1.041 - \frac{-0.0063}{6.3125} = -1.0400$$

Again repeating the process, we get $x_3 = -1.0399$. Thereby, the second required root is

Table 4.12: Birge-Vieta method (iv)

1	-1.895	-0.8840	2.2552	-1.041
	-1.041	3.0564	-2.2615	
1	-2.936	2.1724	0.0063	
	-1.041	4.1401		
1	-3.977	6.3125		

$$x_3 = -1.0399$$

On the other hand, the remaind two roots can be obtaind by solving the equation in the second degree, after dividing $f_1(x)$ by $(x+1.0399)$.

$$x^2 - 2.936x + 2.1724 = 0$$
$$x_4 = \frac{2.936 + \frac{8328}{31589}i}{2}$$
$$x_5 = \frac{2.936 - \frac{8328}{31589}i}{2}$$

Thus, the set of the roots are:

$$x_1 = 3.125$$
$$x_2 = 3.1048$$
$$x_3 = -1.0399$$
$$x_4 = \frac{367}{250} + \frac{4164}{31589}i$$
$$x_5 = \frac{367}{250} - \frac{4164}{31589}i$$

4.3.5 Graeffe's root-squaring method

Graeffe's method or Dandelin-Lobachesky-Graeffe method is an algorithm for finding all of the roots of a polynomial in the same time (Bini and Pan, 2012). The method has been developed independently and Relatively simultaneously by Germinal Pierre Dandelin, Lobachevsky and Karl Heinrich Gräffe (Alston, 1959) .

The method separates the roots of a polynomial by squaring them repeatedly. This squaring is done implicitly, only working on the coefficients. Finally,Viète's formulas are used in order to approximate the roots (Best, 1949).

The method is derived as follows in the following steps:

Suppose that what is required is to find the real roots of the polynomial equation;

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + ... + a_n = 0$$

By putting the even terms on one side of the equation and the odd terms on the other side of the equation and squaring both sides, we get;

$$(x^n + +a_2 x^{n-2} + a_4 x^{n-4} ...+)^2 = (a_1 x^{n-1} + a_3 x^{n-3} + ...)^2$$

Now, by substituting y for x^2, we get;

$$yx^n + +b_1 y^{n-1} + b_2 y^{n-2} ... + b_n = 0$$

where,

$$\begin{aligned} b_1 &= -a_1^2 + 2a_2 \\ b_2 &= a_2^2 - 2a_1 a_3 + 2a_4 \\ b_3 &= -a_3^2 + 2a_2 a_4 - 2a_1 a_5 + 2a_6 \\ &. \\ &. \\ &. \\ b_n &= (-1)^n a_n^2 \end{aligned}$$

Or;

$$(-1)^k b_k = a_k^2 - 2a_{k-1} a_{k+1} + 2a_{k-2} a_{k+2} + ...$$

By repeating this process $r-$times, we get the following equation;

$$x^n + c_1 x^{n-1} + c_2 x^{n-2} + ... + c_n = 0$$

where the roots of the equation are;

$$P_1, p_2, ..., p_n$$

and the roots of the original equation are;

$$z_1, z_2, ..., z_n$$

Obviously,

$$p_j = z_j^{2r}, j = 1, 2, ..., n$$

If we assume;

$$|p_n| < |p_{n-1}| < ... < |p_2| < |p_1|$$
$$|z_n| < |z_{n-1}| < ... < |z_2| < |z_1|$$

then;

$$c_1 = -\sum_k p_k \simeq -p_1$$
$$c_2 = -\sum_{j,k} p_j p_k \simeq -p_1 p_2$$
$$c_3 = -\sum_{j,k,l} p_j p_k p_l \simeq -p_1 p_2 p_3$$
$$\vdots$$

Implies that;

$$p_1 =\simeq -c_1$$
$$p_2 =\simeq -\frac{c_2}{c_1}$$
$$p_3 =\simeq -\frac{c_3}{c_2}$$
$$\vdots$$

Generally;

$$z_j = (p_j)^{\frac{1}{2r}}, \forall j$$

It is worth noting that the sign of the root can be obtained by substituting into the equation.

Example 4.16 Find the roots of the equation:

$$4x^3 - 32x^2 + 68x - 40 = 0$$

Solution: By dividing the equation by 4, putting the even terms on one side and the odd terms on the other side, and also by squaring both sides, we have;

$$(x^3 + 17x)^2 = (8x^2 + 10)^2$$

by substituting y for x^2, we get;

$$y^3 - 30y^2 + 129y - 100 = 0$$
$$(y^3 + 129y)^2 = (30y^2 + 100)^2$$

Again, substituting z for y^2, we get;

$$z^3 - 642z^2 + 10641y - 10000 = 0$$
$$(z^3 + 10641z)^2 = (642z^2 + 10000)^2$$

Next, substituting r for z^2, we get;

$$r^3 - 390882r^2 + 100390881r - 10^8 = 0$$

Thereby, the approximate absolute values of the roots are;

$$|z_1| = (390882)^{\frac{1}{8}} = \frac{12166}{2433}$$
$$|z_2| = (\frac{100390881}{390882})^{\frac{1}{8}} = \frac{2467}{1233}$$
$$|z_3| = (\frac{10^8}{100390881})^{\frac{1}{8}} = \frac{2050}{2051}$$

Thus, the $\{z_1, z_2, z_3\} = \{\frac{12166}{2433}, \frac{2467}{1233}, \frac{2050}{2051}\}$.

4.4 Exercises

Solve the following questions:

Q1: Find the number of real positive and negative roots of the polynomials;

(i) $3x^3 + 3x^2 - 3x + 1$.

(ii) $x^3 - 3x^2 + 3x - 1$.

(iii) $x^4 + 3x^2 - 4x + 1$.

(iv) $x^5 - 3x^3 + 3x^2 - x + 1$.

(v) $x^6 - x^2 + x$.

(vi) $x^4 + 3x^2 + 1$.

(vii) $x^6 - x^5 + x^2 + x$.

(viii) $x^6 - x^5 + 2x^4 - x^3 + x^2 + x$.

Q2: Using Horner's method, find the differential values of the following functions;

(i) $3x^3 - 2x + 1$, at $x = 3$.

(ii) $2x^4 + x^2 - 1$, at $x = 1.5$.

(iii) $2.1x^3 + 3.02x^2 - 1.78$, at $x = -1.41$.

Q3: The equation; $x^8 - 170x^6 + 7.392x^4 - 39.712x^2 + 51.200 = 0$ has eight roots. If four of them are known, they are $\{2, 10, \sqrt{2}, 8\}$, then remove these roots to obtain an equation in the fourth degree.

Q4: Use the bisection method to find the roots of the equations, in which have roots between $[0.5, 1]$;

(i) $p_1(x) = x^3 - 2x - 1$.

(ii) $p_2(x) = x^4 - 3x^3 + x^2 + 1$.

(iii) $p_3(x) = x - 0.2sinx$.

Q5: If the initial interval $[a_0, b_0]$ in the case the bisection method used, and if $M = |b_0 - a_0|$, the values $x_0, x_1, x_2, ...$ have been found by the method. Prove that $|x_{i+1} - x_i| = \frac{M}{2^{i+2}}, i = 0, 1, 2, ...$

Q6: Use Nweton-Raphson method to find the roots of the equations, if they have roots between $x = 1, x = 2$.

(i) $x^3 - 2x - 1 = 0$.

(ii) $x^6 + x^4 + x^3 + 1 = 0$.

Q7: Use Secant method to find the roots of the equations, if they have roots between $x = 1, x = 2$.

(i) $x^3 - 2x - 1 = 0$.

(ii) $x^6 + x^4 + x^3 + 1 = 0$.

Q8: Use Newton-Raphson method to find a root of the equation; $z^5 + (7-2i)z^4 + (20-12i)z^3 + (20-28i)z^2 + (19-12i)z + (13-26i) = 0$

[Hint: $z_0 = 3i$, Ans.: $z_8 = -1 + 2i$]

Q9: What is the positive value of x that makes the function;
$y = \frac{tanx}{x^2}$ has a minimum value.

[Hint: Ans.: $x = 0.94775$]

Q10: Find the intersection of the equations;
$y = ex^{-2}, y = loge(x + 2), x > 2$.

Q11: Use Birge-Vieta method to find the roots of;
$x^3 - 3x + 1 = 0$.

Q12: Use the Root-squaring method to find the roots of;
$x^3 - 5x^2 - 17x + 20 = 0$.

Q13: Apply Newton-Raphson's method to the equations;
$x^3 - N = 0, X^p - M = 0$ to show that; $N^{\frac{1}{3}}, M^{\frac{1}{p}}$, and find $\sqrt{10}, \sqrt[3]{10}$.

Q14: Consider the polynomial;
$p(x) = a_0x^n + a_1x^{n-1} + ... + a_{n-1}x + a_n$, where $b_0 = a_0, b_i = a_i + zb_{i-1}, i = 1, 2, ..., b_n = p(z)$. Prove that;
$\frac{p(x)-p(z)}{x-z} = \sum_{i=0}^{n-1} b_i x^{n-1-i}$.

Q15: In the previous question if b_i is known, and $c_0 = b_0, c_i = b_i + zc_{i-1}, i = 1, 2, ..., n-1$. Prove that $c_{n-1} = p^1(z)$.

5

Matrices

5.1 Introduction

A matrix is a rectangular array arranged in rows and columns, which is used to represent a mathematical object or a property of such an object. The study of matrices takes an important position in terms of its use in the fields of knowledge and science in general and in the field of mathematics in particular. The matrices had been used for the first time by mathematician Arthur Cayley (1822-1995) (Dossey et al., 1987; Cayley, 1858a; Cayley, 1894; Dieudonné, 1978) as a shorthand to express a system of linear equations.

There are numerous applications of matrices whether in mathematics or in science and other fields. Some of them take advantage of the compact representation of a set of numbers in a matrix. Including:

- Algebraic aspects and generalizations (Coburn, 1955; Brown, 1991).

- Game theory and economics (Fudenberg and Tirole, 1991).

- Graph theory (Godsil and Royle, 2001; Punnen and Kabadi, 2002).

- Analysis and geometry (Lang, 1984; Wright, 2006; Lang, 2012a; Trudinger, 1983; Ŝolín, 2005).
- Probability theory and statistics (Latouche and Ramaswami, 1999; Srinivasan and Mehata, 1976; Healy, 2000; Krzanowski, 2000; Conrey, 2007; Brézin et al., 2006).
- Symmetries and transformations in physics (Bessis et al., 1980; Itzykson and Zuber, 1980).
- Linear combinations of quantum states (Schiff, 1968; Böhm, 2013; Weinberg, 1995).
- Normal modes (Wherrett, 1986; Riley et al., 1999).
- Geometrical optics (Guenther, 1990).
- Electronics (William et al., 2007; Nilsson and Riedel, 2001; Dimo, 1975; Baker, 2019; Slurzberg and Osterheld, 1950; Callegaro, 2012; Ushida et al., 2003; Kline, 2019; Dybkaer, 2003; Mohr and Phillips, 2014).
- Chemstry (McNaught et al., 1997a; Fifield and Haines, 2000; Delmonte, 1997; McNaught et al., 1997b; Arruda, 2007).
- Biology (Caswell, 2000; Bruce and Shernock, 2002; Zhang et al., 2010).

In short, we can say that matrices have applications in almost all areas of life. This chapter will be a tool to learn about matrices, the basic concepts of them, their types, how to express them mathematically, and the basic operations on them.

5.2 Linear equations

Definition 5.1 Let F be a field, $a_1, a_2, ...a_n \in F$, and $x_1, x_2, ..., x_n$ be unknown variables. The mathematical form:

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b \tag{5.1}$$

is called a linear equation, and $a_1, a_2, ...a_n$ are its cofficints and b is the absolute term (Barnett et al., 2019; Charles, 1892; David, 1890; Wilson and Tracey, 1925).

Definition 5.2 A set of equations in the form:

$$
\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1\\
&a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2\\
&.\\
&.\\
&.\\
&a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m
\end{aligned}
\tag{5.2}
$$

is called a system of linear equation of m equations, and n varisbles (Barnett et al., 2019; Charles, 1892; David, 1890; Wilson and Tracey, 1925).

Definition 5.3 The $(n - tuple)$n $x_1, x_2, ...x_n \in F$ which satisfies (5.2) is called the solution of the linear equation system. (Cullen, 2012).

Definition 5.4 Let F be a field, $a_1, a_2, ...a_n \in F$, and $x_1, x_2, ..., x_n$ be unknown variables. The mathematical form:

$$a_1x_1 + a_2x_2 + ... + a_nx_n = 0 \tag{5.3}$$

is called a homogeneous linear equation, (Barnett et al., 2019; Charles, 1892; David, 1890; Wilson and Tracey, 1925).

Definition 5.5 A set of equations in the form:

$$
\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = 0\\
&a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = 0\\
&.\\
&.\\
&.\\
&a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = 0
\end{aligned}
\tag{5.4}
$$

is called system of homogeneous linear equations (Strang, 2006; Strang, 2011; Apostol, 1967).

Note: When ($b_1 = b_2 = ... = b_m = 0$), the $n - tiple$ $(0, 0, ..., 0)$ is a solution to the $m-$ homogeneous linear equations.

Definition 5.6 A linear combination is an expression constructed from a set of terms by multiplying each term by a constant and adding the results. Mathematically, a linear combination of x_1 and x_2 would be any expression of the form $ax + by$, where a, b are constants in the field F (Strang, 2022; Lay, 2003; Axler, 2015).

Note: In the system (5.2), if we multiply the equations by $c_1, c_2, ..., c_m$ respectively, and add them, the result will be:

$$\begin{aligned}
&(c_1a_{11} + c_2a_{21} + ... + c_ma_{m1})x_1 \\
&+ (c_1a_{12} + c_2a_{22} + ... + c_ma_{m2})x_2 \\
&. \\
&. \\
&. \\
&+ (c_1a_{1n} + c_2a_{2n} + ... + c_ma_{mm})x_n \\
&= c_1b_1 + c_2b_2 + ... + c_mb_m
\end{aligned}$$

Example 5.1 Solve the following system of linear equations:

$$\begin{aligned}
&2x_1 + 3x_2 + 8x_3 + x_4 = 1...(1) \\
&x_1 + x_2 + 3x_3 - x_4 = 2...(2) \\
&3x_1 - 4x_2 + 8x_3 - x_4 = 5...(3)
\end{aligned}$$

Solution: Multiplying (2) by 2 and subtracting it from (1) we get;

$$x_2 + 2x_3 + 3x_4 = 2...(4)$$

Again Multiplying (2) by 3 and subtracting it from (3) we get;

$$x_2 + x_3 - 2x_4 = 1...(5)$$

Solving (4) , (5) simultaneously, , we get;

$x_2 = 6x_4, x_3 = 1 - 5x_4$.

By putting the values of x_2, x_3 at one of the equations of the system, we obtain;

$x_1 = -1 + 10x_4$.

Thereby, the set solution of the system is $(-1+10x_4, 6x_4, 1-5x_4, x_4)$.

It should be noticed that, we have found x_1, x_2, x_3 in terms of x_4.

Thus, x_1, x_2, x_3 are dependent variables, and x_4 is independent variable. Furthermore, the set solution of the system is infinite, because, the system consisted of four variables with three equations.

Example 5.2 Solve the following system of linear equations:

$$
\begin{aligned}
&6x_1 + 3x_2 + 2x_3 = -6 \\
&12x_1 - 3x_2 + x_3 = 15 \\
&-24x_1 - 3x_2 + 4x_3 = 24
\end{aligned}
$$

[Hint: Ans.: $(0, -4, 3)$]

The solution has been left as an exercise for the reader.

5.3 Exercises

Solve the following system of linear equations:

Q1:

$$
\begin{aligned}
&x_1 + x_2 = 6 \\
&3x_1 - 2x_2 = 2
\end{aligned}
$$

Q2:

$$
\begin{aligned}
&x_1 + x_2 + x_3 = 8 \\
&x_1 - x_2 + 3x_3 = 13 \\
&24x_1 + 3x_2 - x_3 = 7
\end{aligned}
$$

Q3:

$$
\begin{aligned}
&x_1 + x_2 + x_3 = 4 \\
&2x_3 + 3x_2 + 2x_1 = 5 \\
&3x_1 + 4x_2 - 3x_3 = -3
\end{aligned}
$$

Q4:

$$\begin{aligned} &x_1 + x_2 - 2x_3 = 3 \\ &3x_1 + x_2 - 6x_3 = 8 \\ &x_2 + 7x_2 + 3x_3 = 0 \\ &-x_1 + 8x_2 + 7x_3 = -4 \end{aligned}$$

Q5:

$$\begin{aligned} &x_1 + x_2 + 3x_3 + x_4 = 0 \\ &x_1 - x_2 - x_3 - x_4 = 0 \\ &3x_1 + x_2 + 5x_3 + 3x_4 = 0 \\ &x_1 + 5x_2 + 11x_3 + 8x_4 = 0 \end{aligned}$$

5.4 Matrices

Definition 5.7 A matrix is a rectangular array of numbers (objects), called the entries of the matrix. Matrices are subject to standard operations such as addition and multiplication. A matrix over a field F is a rectangular array of elements of F. A real matrix and a complex matrix are matrices whose entries are respectively real numbers or complex numbers (Lang, 1984; Lang, 2002; Fraleigh, 2003; Nering, 1970).

Note:

(i) A matrix can be represented as:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

Or, the matrix A has m rows, and n columns. Mathematicaly, $A = (a_{ij}) \in F, \forall i, j; 1 \leq i \leq m, 1 \leq j \leq n$.

(ii) If a matrix consists of m rows and n columns, it is called a matrix $m \times n$, and $m \times n$ represents its degree.

(iii) Capital letters are used to denote matrices. For example, $A_{5\times3}$ is a mtrix has five rows and three columns.

(iv) A matrix of degree $m \times n$ is defined as a function from the set $(i, j) : i = 1, 2, ..., m \wedge j = 1, 2, ..., n \to F$. It means this function is fixed each location of element (i, j) of the field elements F, and covering all possibilities, the matrix is formed.

5.5 Types of matrices

There are different types of matrices, and the following is a list of these types in some detail, based on some of the studies of scientists (Lang, 1984; Lang, 2002; Fraleigh, 2003; Nering, 1970).

5.5.1 Row matrix

Definition 5.8 A row matrix has only one row but any number of columns. A matrix is said to be a row matrix if it has only one row. Or, $A = [a_{ij}]_{1\times n}$ is a row matrix of order 1×5.

Example 5.3 $A = \begin{bmatrix} \sqrt{-2} & 7 & 3 & 4 & 1 \end{bmatrix}$ is a row matrix of order 1×5.

5.5.2 Column matrix

Definition 5.9 A column matrix has only one column but any number of rows. A matrix is said to be a column matrix if it has only one column. Or, $A = [a_{ij}]_{m\times1}$ is a column matrix of order $m \times 1$.

Example 5.4 $A = \begin{bmatrix} 4 \\ -5 \\ \sqrt{3} \end{bmatrix}$ is a column matrix of order 3×1.

5.5.3 Square matrix

Definition 5.10 A square matrix has the number of columns equal to the number of rows. A matrix in which the number of rows is equal to the number of columns is said to be a square matrix. Thus an $m \times n$ matrix is said to be a square matrix if $m = n$, and is known as a square matrix of order 'm'. Or, $A = [a_{ij}]_{m\times m}$ is a square matrix of order m.

Example 5.5 $A \begin{bmatrix} -1 & 2 & 4 & 2 & -\frac{3}{5} \\ \frac{1}{2} & -7 & 6 & -3 & 3 \\ -\frac{7}{8} & -2 & -3 & 0 & 9 \\ 1 & -2 & -7 & 3 & 5 \\ 7 & -\frac{3}{5} & -5 & 3 & 0 \end{bmatrix}$
is a square matrix of order 5.

5.5.4 Trace of a square matrix

Definition 5.11 The trace of an $n \times n$ square matrix A is defined as (Weisstein, 2002; Lipschutz and Lipson, 2018):

$$T(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + ... + a_{nn}$$

where a_{ii} denotes the entry on the ith row and ith column of A. The entries of A can be real numbers, complex numbers, or more generally elements of a field F. The trace is not defined for non-square matrices.

Example 5.6 $A \begin{bmatrix} -1 & 2 & 4 & 2 \\ \frac{1}{2} & -7 & 6 & -3 \\ -\frac{7}{8} & -2 & -3 & 0 \\ 1 & -2 & -7 & 3 \end{bmatrix}$
$T(A) = \sum_{i=1}^{4} a_{ii} = (-1) + (-7) + (-3) + (+3) = -8.$

Note: The trace is a linear mapping. That is for each matrices A, B, and a constant c, has the following basic properties:

(i) $T(A + B) = T(A) + T(B)$.

(ii) $T(cA) = cT(A)$.

(iii) $T(A) = T(A^T)$.

(iv) $T(AB) = T(BA)$.

5.5.5 Rectangular matrix

Definition 5.12 A matrix is said to be a rectangular matrix if the number of rows is not equal to the number of columns.

Example 5.7 (i) $A \begin{bmatrix} -1 & 2 & 4 & 2 \\ \frac{1}{2} & -7 & 6 & -3 \\ -\frac{7}{8} & -2 & -3 & 0 \\ 1 & -2 & -7 & 3 \\ 7 & -\frac{3}{5} & -5 & 3 \end{bmatrix}$ is a square matrix of order 5×4.

(ii) $B = \begin{bmatrix} -1 & 2 & 4 & 2 & 7 \\ \frac{1}{2} & -7 & 6 & -3 & 2 \\ -\frac{7}{8} & -2 & -3 & 0 & -3 \end{bmatrix}$ is a square matrix of order 3×5.

5.5.6 Diagonal matrix

Definition 5.13 A square matrix $A = [a_{ij}]_{m\times m}$ is said to be a diagonal matrix if all its non-diagonal elements are zero, that is a matrix $A = [a_{ij}]_{m\times m}$ is said to be a diagonal matrix if $a_{ij} = 0$, when $i \neq j$.

Example 5.8 (i) $A = \begin{bmatrix} -3 \end{bmatrix}$ is a matrix of order 1.

(ii) $B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & -3 \end{bmatrix}$ is a matrix of order 3.

5.5.7 Scalar matrix

Definition 5.14 A diagonal matrix is said to be a scalar matrix if all the elements in its principal diagonal are equal to some non-zero constant. A diagonal matrix is said to be a scalar matrix if its diagonal elements are equal, that is, a square matrix $A = [a_{ij}]_m \times m$ is said to be a scalar matrix if: $\begin{cases} a_{ij} = 0, \text{when } i \neq j. \\ a_{ij} = k, \text{when } i = j, \quad \text{for some constant } k. \end{cases}$

Example 5.9 (i) $A = \begin{bmatrix} 2 \end{bmatrix}$ is a matrix of order 1.

(ii)

(iii) $B = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$ is a matrix of order 2.

(iv) $C = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$ is a matrix of order 5.

5.5.8 Zero matrix

Definition 5.15 A matrix is said to be zero matrix or null matrix if all its elements are zero, and denoted by O.

Example 5.10 (i) $A = \begin{bmatrix} 0 \end{bmatrix}$ is a matrix of order 1.

(ii)

(iii) $B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ is a matrix of order 2.

(iv) $C = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ is a matrix of order 4.

5.5.9 Unit matrix

Definition 5.16 If a square matrix has all elements 0 and each diagonal elements are they are units, it is called identity matrix and denoted by I. Or, the square matrix $A = [a_{ij}]_{m\times m}$ is an identity matrix if:

$$\begin{cases} a_{ij} = 1, \text{when } i = j. \\ a_{ij} = 0, \text{when } i \neq j \end{cases}$$

Example 5.11 (i) $A = \begin{bmatrix} 1 \end{bmatrix}$ is a matrix of order 1.

(ii)

(iii) $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is a matrix of order 2.

(iv) $C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is a matrix of order 3.

Note:

(i) We denote the identity matrix of order n by I_n. When the order is clear from the context, and we simply write it as I.

(ii) A scalar matrix is an identity matrix when $k = 1$.

(iii) Every identity matrix is clearly a scalar matrix, but the vice versa is not true.

5.5.10 Upper triangular matrix

Definition 5.17 A square matrix in which all the elements below the diagonal are zero is known as the upper triangular matrix, and denoted by U.

Example 5.12 $A = \begin{bmatrix} 5 & 4 & 3 & -1 \\ 0 & 2 & -7 & 4 \\ 0 & 0 & -1 & -\frac{3}{4} \\ 0 & 0 & 0 & 9 \end{bmatrix}$ is the U in order 4.

5.5.11 Lower triangular matrix

Definition 5.18 A square matrix in which all the elements above the diagonal are zero is known as the upper triangular matrix, and denoted by L.

Example 5.13 $B = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 0 \\ -1 & 2 & 4 & 0 & 0 \\ 5 & 3 & -2 & 9 & 0 \\ 2 & 7 & 3 & 9 & \frac{7}{9} \end{bmatrix}$ is the L in order 5.

5.5.12 Transpose of a matrix

Definition 5.19 The transpose of a matrix A, denoted by A^T (Whitelaw, 2019; Cayley, 1858a), may be constructed by any one of the following methods:

(i) Reflect A over its main diagonal (which runs from top-left to bottom-right) to obtain A^T.

(ii) Write the rows of A as the columns of A^T.

(iii) Write the columns of A as the rows of A^T.

Or, the *ith* row, *jth* column element of A^T is the *jth* row, *ith* column element of A: $[A^T]_{ij} = [A]_{ji}$. If A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix.

Example 5.14 If $A = \begin{bmatrix} 5 & 0 & 6 & -1 \\ 2 & 2 & 3 & 5 \\ -1 & 2 & 4 & 0 \\ 5 & 3 & -2 & 9 \\ 2 & 7 & 3 & \frac{7}{9} \end{bmatrix}$ then,

$$A^T = \begin{bmatrix} 5 & 2 & -1 & 5 & 2 \\ 0 & 2 & 2 & 3 & 7 \\ 6 & 3 & 4 & -2 & 3 \\ -1 & 5 & 0 & 9 & \frac{7}{9} \end{bmatrix}$$

5.5.13 Symmetric matrix

Definition 5.20 A symmetric matrix is a square matrix that is equal to its transpose. Or, A is a symmetric if and only if $A = A^T$ (Bellman, 1997; Zielke, 1985).

Note:

(i) Since equal matrices have equal dimensions, hence, only square matrices can be symmetric.

(ii) The entries of a symmetric matrix are symmetric with respect to the main diagonal. So if a_{ij} denotes the entry in the *ith* row and *jth* column then; A is symmetric $\Leftrightarrow a_{ij} = a_{ji}, \forall i, j$.

Example 5.15 The following 4×4 matrix is symmetric:

$$A = \begin{bmatrix} 1 & 7 & 3 & 6 \\ 7 & 3 & 2 & 1 \\ 3 & 2 & 5 & 0 \\ 6 & 1 & 0 & 9 \end{bmatrix}. \text{ Since } A = A^T.$$

5.5.14 Skew symmetric matrix

Definition 5.21 A skew symmetric (antisymmetric or antimetric) matrix is a square matrix whose transpose equals its negative. Or, A is skew symmetric matrix if and only if $A^T = -A$ (W et al., 1997; Lipschutz and Lipson, 2018; Lipschutz and Lipson, 2001a).

Note: In terms of the entries of the matrix, if a_{ij} denotes the entry in the ith row and jth column then;

A is skew symmetric $\Leftrightarrow a_{ji} = -a_{ji}, \forall i, j$.

Example 5.16 $A = \begin{bmatrix} 0 & 3 & -60 \\ -3 & 0 & -7 \\ 60 & 7 & 0 \end{bmatrix}$ is skew-symmetric because;

$$-A = \begin{bmatrix} 0 & -3 & 60 \\ 3 & 0 & 7 \\ -60 & -7 & 0 \end{bmatrix} = A^T.$$

5.5.15 Matrix conjugate

Definition 5.22 The matrix A is the conjugate matrix of matrix B if the elements of A are the complex conjugate numbers of the elements of matrix B, and it denoted by $\overline{A}$ (Arfken, 1985; Ayres, 1962; Courant and Hilbert, 2008).

Example 5.17 If $A = \begin{bmatrix} 3i & 4+i & 2 \\ 7-i & 3+7i & 4-10i \\ -8i & 11+3i & 0 \\ 6 & 1-i & 0 \end{bmatrix}$, then

$$\overline{A} = \begin{bmatrix} -3i & 4-i & 2 \\ 7+i & 3-7i & 4+10i \\ 8i & 11-3i & 0 \\ 6 & 1+i & 0 \end{bmatrix}$$

is the conjugate of A, since all entries of matrix $\overline{A}$ are conjugated. In other words, the numbers in matrix $\overline{A}$ have the same real part as numbers in matrix A, but their complex part have the opposite sign.

Note: The conjugate transpose of A is denoted by $\overline{A'} = A^{\star}$.

5.5.16 Hermitian matrix

Definition 5.23 A Hermitian matrix (self adjoint matrix) is a complex square matrix that is equal to its own conjugate transpose that is, the element in the ith row and jth column is equal to the complex conjugate of the element in the jth row and ith column.

Or, A is Hermittian matrix, if and only if $a_{ij} = \overline{a_{ji}}, \forall i, j$. In matrix form, A is Hermittian matrix, if and only if $A = \overline{A^T}$. (Liu and Li, 2015; Frankel, 2011; Sylvester et al., 1855).

Example 5.18 The following matrices are Hermitian:

(i) $A = \begin{bmatrix} 3 & 7-2i \\ 7+2i & \frac{1}{3} \end{bmatrix}$.

(ii) $B = \begin{bmatrix} 2 & 1-2i & 2-6i \\ 1+2i & 4 & 1 \\ 2+6i & 1 & -5 \end{bmatrix}$.

Note: Hermitian matrix denoted by A^H.

5.5.17 Skew Hermitian matrix

Definition 5.24 A square matrix with complex entries is said to be skew Hermitian (anti-Hermitian matrix) if its conjugate transpose is the negative of the original matrix.

Or, A is a skew Hermitian matrix if and only if $A^H = -A$, where A^H denotes the conjugate transpose of the matrix A. This means that;

A is a skew Hermitian matrix if and only if $a_{ij} = -\overline{a_{ji}}, \forall i, j$ (Hom and Johnson, 1985; Horn, 1985; Meyer and Stewart, 2023).

Example 5.19 The following matrices are Hermitian:

(i) $A = \begin{bmatrix} 0 & 1+i \\ -1+i & 0 \end{bmatrix}$.

(ii) $B = \begin{bmatrix} 0 & 2i & 2-3i \\ 2i & 0 & -4 \\ -2-3i & 4 & 0 \end{bmatrix}$.

5.6 Exercises

What are the following types of matrices?

Q1:

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & -4 \\ 0 & -4 & 5 \end{bmatrix}.$$

Q2:

$$\begin{bmatrix} 0 & 1-2i & 5i \\ -1-2i & 0 & 3 \\ 5i & 3 & i \end{bmatrix}.$$

Q3:

$$\begin{bmatrix} 0 & -2 & -3 & -4 \\ 2 & 0 & -5 & -6 \\ 3 & 5 & 0 & -7 \\ 4 & 6 & 7 & 0 \end{bmatrix}.$$

Q4:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

Q5:

$$\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{-2}{3} \\ \frac{-2}{3} & \frac{2}{3} & \frac{-1}{3} \end{bmatrix}.$$

Q6:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{5}{13} & \frac{12}{13} \\ 0 & \frac{-12}{13} & \frac{5}{13} \end{bmatrix}.$$

Q7:

$$\begin{bmatrix} i & i \\ i & i \end{bmatrix}.$$

Q8:

$$\begin{bmatrix} \frac{-6}{7} & \frac{2}{7} & \frac{3}{7} \\ \frac{2}{7} & \frac{-3}{7} & \frac{6}{7} \\ \frac{3}{7} & \frac{6}{7} & \frac{2}{7} \end{bmatrix}.$$

Q9:

$$\begin{bmatrix} a^2 - b^2 - c^2 + d^2 & 2(ab - cd) & 2(ac - bd) \\ 2(ab + cd) & -a^2 + b^2 - c^2 + d^2 & 2(bc - ad) \\ 2(ac - bd) & 2(bc + ad) & -a^2 - b^2 + c^2 + d^2 \end{bmatrix}.$$

Q10:

$$\begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & 2 & 1 & i \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

Q11:

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ i & 1 & 0 & 0 \\ 2 & 2 & 3 & 0 \\ -1 & 4 & -i & 1 \end{bmatrix}.$$

Q12:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

5.7 Operations on matrices

In this section, we review some of the basic operations on the set of matrices: addition, subtraction, multiplication, and division are not defined. It is worth noting that these operations are inherited from the regular operations in the field F, with emphasis on the symbols $\boxdot, \boxplus, \boxminus, \boxtimes$ to Scalar multiplication, adding, subtratind, and maltiplying matrices respectively to show that these operations are special on the set of matrices. But for convenience, we will use the usual symbols.

5.7.1 Equality of matrices

Definition 5.25 Two matrices A and B are said to be equal if they are of the same degree and all of their corresponding elements are equal. Or, $A_{m\times n} = B_{m\times n} \Leftrightarrow a_{ij} = b_{ij}, \forall i, j$ (Finkbeiner, 2013; Bellman, 1997; Schwartz, 2001).

5.7.2 Scalar multiplication and scalar division

Definition 5.26 Multiplying by a constant of a matrix is defined by multiplying the constant by each element of the matrix, and the symbol $\boxdot$ is used to indicate it.

Or, if the scalar is k, and the matrix is a_{mn} then:
$k \boxdot (a_{ij})_{mn} = (k \boxdot a_{ij}); \forall i, j$ (Strang, 2006; Strang, 2012; Strang, 2012; Lay, 2003).

Now, we can define a scalar division of a matrix, as follows;

Definition 5.27 Scalar division by a constant of a matrix is defined by dividing each element of the matrix by the constant and the symbol $\frac{1}{k}\boxdot$ is used to indicate it.

Or, if the scalar is k, and the matrix is a_{mn} then:
$\frac{1}{k} \boxdot (a_{ij})_{mn} = (\frac{1}{k} \boxdot a_{ij}); \forall i, j.$

Example 5.20 (i) $3 \boxdot \begin{bmatrix} 2 & 0 & 0 & 0 \\ i & 1 & 0 & 0 \\ 2 & 2 & 3 & 0 \\ -1 & 4 & -i & 1 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 3i & 3 & 0 & 0 \\ 6 & 6 & 9 & 0 \\ -3 & 12 & -3i & 3 \end{bmatrix}.$

(ii) $\frac{1}{3} \boxdot \begin{bmatrix} 2 & 0 & 0 & 0 \\ i & 1 & 0 & 0 \\ 2 & 2 & 3 & 0 \\ -1 & 4 & -i & 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & 0 & 0 & 0 \\ \frac{i}{3} & \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & \frac{2}{3} & 1 & 0 \\ \frac{-1}{3} & \frac{4}{3} & \frac{-i}{3} & \frac{1}{3} \end{bmatrix}.$

5.7.3 Adding and subtracting matrices

Definition 5.28 Matrix addition is the operation of adding two matrices by adding the corresponding entries together.

Or, $(a_{ij})_{mn} \boxplus (b_{ij})_{mn} = (a_{ij} \boxplus b_{ij})_{mn} = c_{ij}$ (Anton and Rorres, 2013; Lipschutz and Lipson, 2018; Lipschutz and Lipson, 2001a; Riley et al., 1999).

Now, we can define subtraction of matrices, as follows;

Definition 5.29 The two matrices whose difference is calculated have the same number of rows and columns. The subtraction of the two matrices can also be defined as addition of A and $-B$.

Or, $(a_{ij})_{mn} \boxminus (b_{ij})_{mn} = (a_{ij} \boxplus (-1 \boxdot b_{ij}))_{mn} = c_{ij}$

Example 5.21 (i) $\begin{bmatrix} 1 & 3 & -1 & 3 \\ 2 & 4 & 0 & 5 \\ -5 & 4 & 3 & -2 \\ 3 & 4 & 0 & 6 \end{bmatrix} \boxplus \begin{bmatrix} 6 & 3 & 7 & 3 \\ 0 & 4 & 9 & 5 \\ 5 & -4 & 3 & 2 \\ 13 & 0 & 10 & -6 \end{bmatrix} = \begin{bmatrix} 7 & 6 & 6 & 6 \\ 2 & 8 & 9 & 10 \\ 0 & 0 & 6 & 0 \\ 16 & 4 & 10 & 0 \end{bmatrix}$.

(ii) $\begin{bmatrix} 1 & 3 & -1 \\ 2 & 4 & 0 \\ -5 & 4 & 3 \end{bmatrix} \boxminus \begin{bmatrix} 6 & 3 & 7 \\ 0 & 4 & 9 \\ 5 & -4 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 4 & 0 \\ -5 & 4 & 3 \end{bmatrix} \boxplus \begin{bmatrix} -6 & -3 & -7 \\ 0 & -4 & -9 \\ 5 & -4 & -7 \end{bmatrix} = \begin{bmatrix} -5 & 0 & -8 \\ 2 & 0 & -9 \\ 0 & 0 & -4 \end{bmatrix}$.

5.7.4 Multiplying matrices

Matrix multiplication was first described by Jacques Philippe Marie Binet in 1812 (Souza and Tatiana, 2018), to represent the composition of linear maps that are represented by matrices. Matrix multiplication is thus a basic tool of linear algebra, and as such has numerous applications in many areas of mathematics, as well as in applied mathematics, statistics, physics, economics, and engineering (Van et al., 1991). Computing matrix products is a central operation in all computational applications of linear algebra.

Definition 5.30 Matrix multiplication is a binary operation that produces a matrix from two matrices. For matrix multiplication, the number of columns in the first matrix must be equal to the number of rows in the second matrix. The resulting matrix, known as the matrix product, has the number of rows of the first and the number of columns of the second matrix. The product of matrices A and B is denoted as AB (Hoffman and Kunze, 1967; Hohn, 1972; Strang, 2006; Strang, 2022).

Or, if A is an $m \times n$ matrix, and B is an $n \times p$ matrix,

$$A = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ . & . & ... & . \\ . & . & ... & . \\ . & . & ... & . \\ a_{m1} & a_{m2} & ... & a_{mn} \end{bmatrix}, B = \begin{bmatrix} b_{11} & b_{12} & ... & b_{1p} \\ b_{21} & b_{22} & ... & b_{2p} \\ . & . & ... & . \\ . & . & ... & . \\ . & . & ... & . \\ b_{n1} & b_{n2} & ... & b_{np} \end{bmatrix},$$

the matrix product $C = A \boxdot B$ is defined to me $m \times p$ matrix

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix},$$

where

$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + ... + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$

$\forall i = 1, 2, .., m; j = 1, 2, ..., p.$

Implies that, the dot product of the ith row of A and the jth column of B (Lipschutz and Lipson, 2018; Lipschutz and Lipson, 2001b; Riley et al., 1999; Adams and Essex, 2018; Horn and Johnson, 2012).

Thus,

$$C = A \boxdot B$$

$$= \begin{bmatrix} a_{11}b_{11} + ... + a_{1n}b_{n1} & a_{11}b_{12} + ... + a_{1n}b_{n2} & ... & a_{11}b_{1p} + ... + a_{1n}b_{np} \\ a_{21}b_{11} + ... + a_{2n}b_{n1} & a_{21}b_{12} + ... + a_{2n}b_{n2} & ... & a_{21}b_{1p} + ... + a_{2n}b_{np} \\ . & . & ... & . \\ . & . & ... & . \\ . & . & ... & . \\ a_{m1}b_{11} + ... + a_{mn}b_{n1} & a_{m1}b_{12} + ... + a_{mn}b_{n2} & ... & a_{m1}b_{1p} + ... + a_{mn}b_{np} \end{bmatrix}$$

Example 5.22 Consider $A = \begin{bmatrix} 1 & 3 & -1 \\ 2 & \frac{1}{2} & 0 \\ -5 & 0 & 3 \end{bmatrix}, B = \begin{bmatrix} 4 & 3 & \frac{1}{3} \\ 2 & 4 & 0 \\ -\frac{1}{5} & 4 & 3 \end{bmatrix}$.

Find;

(i) $A \boxdot B$.

(ii) $B \boxdot A$.

(iii) what do you notice?

Solution:

(i) $A \boxdot B = \begin{bmatrix} \frac{51}{5} & 11 & -\frac{8}{3} \\ 9 & 8 & \frac{2}{3} \\ -\frac{103}{5} & -3 & \frac{22}{3} \end{bmatrix}$.

(ii) $B \boxdot A = \begin{bmatrix} \frac{25}{3} & \frac{27}{2} & -3 \\ 10 & 8 & -2 \\ -\frac{36}{5} & \frac{7}{5} & \frac{46}{5} \end{bmatrix}$.

(iii) We note that $A \boxdot B \neq B \boxdot A$

Note: In general, $A \boxdot B \neq B \boxdot A$, for some matrices A, B.

5.7.5 Matrices partition

The aim of matrices partition with large degrees is for scientific and applied purposes, including ease of operations on them, and control over them through partial matrices. Any matrix may be interpreted as a partitioned matrix in one or more ways, with each interpretation defined by how its rows and columns are partitioned (Macedo and Oliveira, 2013).

Definition 5.31 A partitioned matrix is a matrix that is interpreted as having been broken into sections called submatrices (Eves, 1980). A matrix interpreted as a partitioned matrix can be visualized as the original matrix with a collection of horizontal and vertical lines, which partition it, into a collection of smaller matrices (Anton and Rorres, 2013).

Example 5.23

$$\left[\begin{array}{cc:cc} 1 & 2 & 2 & 7 \\ 1 & 5 & 6 & 7 \\ \hdashline 3 & 3 & 4 & 5 \\ 3 & 3 & 6 & 7 \end{array}\right]$$

can be partitioned into four 2×2 submatrices:

$A_{11} = \begin{bmatrix} 1 & 2 \\ 1 & 5 \end{bmatrix}, A_{12} = \begin{bmatrix} 2 & 7 \\ 6 & 7 \end{bmatrix}, A_{21} = \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix}, A_{22} = \begin{bmatrix} 4 & 5 \\ 6 & 7 \end{bmatrix}.$

Thus, the partitioned matrix can then be written as:

$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$

Note:

(i) The matrix $A_{4\times 4}$ has became $A_{2\times 2}$ after partition.

(ii) Partitioning the rows of the second matrix is the same as partitioning the columns in the first matrix.

(iii) Whenever partitioning of matrices, it is taken into account to obtain the largest number of zero and unary matrices for the purpose of the operations on them.

Example 5.24 Consider

$$A = \left[\begin{array}{cc:ccc:c} 1 & 3 & 0 & 0 & 0 & 3 \\ 3 & 4 & 0 & 0 & 0 & 5 \\ \hdashline 0 & 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hdashline 1 & 3 & 4 & 1 & 3 & 1 \end{array}\right], B = \left[\begin{array}{cc:cc:c} 1 & 0 & 1 & 2 & 2 \\ 0 & 0 & 5 & 4 & 1 \\ \hdashline 1 & 2 & 1 & 3 & 3 \\ 3 & 2 & 2 & 4 & 1 \\ 2 & 2 & 3 & 4 & 2 \\ \hdashline 0 & 0 & 2 & 2 & 1 \end{array}\right]$$

Evaluate $A \boxdot B$.

Solution: $A_{11} = \begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix}, A_{12} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, A_{13} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}.$

$A_{21} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, A_{22} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, A_{23} = \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix}.$

$A_{31} = \begin{bmatrix} 1 & 3 \end{bmatrix}, A_{32} = \begin{bmatrix} 4 & 1 & 3 \end{bmatrix}, A_{33} = \begin{bmatrix} 1 \end{bmatrix}.$

$B_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, B_{12} = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}, B_{13} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$

$B_{21} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \\ 2 & 2 \end{bmatrix}, B_{22} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 3 & 4 \end{bmatrix}, B_{23} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}.$

$B_{31} = \begin{bmatrix} 0 & 0 \end{bmatrix}, B_{32} = \begin{bmatrix} 2 & 2 \end{bmatrix}, B_{33} = \begin{bmatrix} 1 \end{bmatrix}.$

So the product is equal to

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \boxdot \begin{bmatrix} b_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} =$$

$$\begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} & A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32} & A_{11}B_{13} + A_{12}B_{23} + A_{13}B_{33} \\ A_{21}B_{11} + A_{22}B_{21} + A_{23}B_{31} & A_{21}B_{12} + A_{22}B_{22} + A_{23}B_{32} & A_{21}B_{13} + A_{22}B_{23} + A_{23}B_{33} \\ A_{31}B_{11} + A_{32}B_{21} + A_{33}B_{31} & A_{31}B_{12} + A_{32}B_{22} + A_{33}B_{32} & A_{31}B_{13} + A_{23}B_{32} + A_{33}B_{33} \end{bmatrix}$$

To obtain the desired result, $3 \times 9 = 27$ small matrix multiplication operations are required.

Thus, $A \boxdot B =$

$$\begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix} & \begin{bmatrix} 22 & 20 \\ 33 & 32 \end{bmatrix} & \begin{bmatrix} 8 \\ 15 \end{bmatrix} \\ \begin{bmatrix} 1 & 2 \\ 3 & 2 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 9 & 11 \\ 4 & 6 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 7 \\ 2 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 14 & 16 \end{bmatrix} & \begin{bmatrix} 33 & 44 \end{bmatrix} & \begin{bmatrix} 25 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 22 & 20 & 8 \\ 3 & 0 & 33 & 32 & 15 \\ 1 & 2 & 9 & 11 & 7 \\ 3 & 2 & 4 & 6 & 2 \\ 14 & 16 & 33 & 44 & 25 \end{bmatrix}$$

Example 5.25 Find the product of the following:

$$\left[\begin{array}{cc:cc} 1 & 0 & 1 & 0 \\ 0 & 3 & 0 & 2 \\ \hdashline 0 & 0 & 2 & 0 \\ 4 & 3 & 0 & 5 \end{array}\right] \boxdot \left[\begin{array}{cc} 3 & 5 \\ 0 & 4 \\ \hdashline 1 & 3 \\ 6 & 1 \end{array}\right].$$

Solution: After dividing the two matrices into submatrices - and after performing the operations, we find that:

$$\begin{bmatrix} A_{11} & A_{21} \\ A_{21} & A_{22} \end{bmatrix} \boxdot \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} = \begin{bmatrix} A_{11} \boxdot B_{11} \boxplus A_{12} \boxdot B_{21} \\ A_{21} \boxdot B_{11} \boxplus A_{22} \boxdot B_{21} \end{bmatrix}. \text{ where,}$$

$$A_{11} \boxdot B_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \boxdot \begin{bmatrix} 3 & 4 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 0 & 15 \end{bmatrix}.$$

$$A_{12} \boxdot B_{21} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \boxdot \begin{bmatrix} 1 & 3 \\ 6 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 12 & 2 \end{bmatrix}.$$

$$A_{22} \boxdot B_{11} = \begin{bmatrix} 0 & 0 \\ 4 & 3 \end{bmatrix} \boxdot \begin{bmatrix} 3 & 4 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 12 & 31 \end{bmatrix}.$$

$$A_{22} \boxdot B_{21} = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} \boxdot \begin{bmatrix} 1 & 3 \\ 6 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 30 & 5 \end{bmatrix}.$$

Thereby,

$$A_{11} \boxdot B_{11} \boxplus A_{12} \boxdot B_{21} = \begin{bmatrix} 3 & 4 \\ 0 & 15 \end{bmatrix} \boxplus \begin{bmatrix} 1 & 3 \\ 12 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 7 \\ 12 & 17 \end{bmatrix}.$$

$$A_{21} \boxdot B_{11} \boxplus A_{22} \boxdot B_{21} = \begin{bmatrix} 0 & 0 \\ 12 & 31 \end{bmatrix} \boxplus \begin{bmatrix} 2 & 6 \\ 30 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 42 & 36 \end{bmatrix}.$$

Thus, the matrix product is

$$\begin{bmatrix} \begin{bmatrix} 4 & 7 \\ 12 & 17 \end{bmatrix} \\ \begin{bmatrix} 2 & 6 \\ 42 & 36 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 4 & 7 \\ 12 & 17 \\ 2 & 6 \\ 42 & 36 \end{bmatrix}.$$

5.8 Vectors

Vectors were introduced in geometry and physics for quantities that have both a magnitude and a direction, such as displacements, forces and velocity. Such quantities are represented by geometric vectors in the same way as distances, masses and time are represented by real numbers. The term vector is also used, in some contexts, for tuples, which are finite sequences of numbers of a fixed length (Ivanov, 2001; Heinbockel, 2001; Itō, 1993).

Definition 5.32 A vector is an object that has both a magnitude and a direction. Or, a vector is a directed line segment, whose length is the magnitude of the vector and with an arrow indicating the direction. The direction of the vector is from its tail to its head (Hamadameen, 2022; Ivanov, 2001; Hoffman and Kunze, 1967; Hohn, 1972; Strang, 2006).

5.8.1 Vector expression

A vector with $m-$dimension on the field F means an ordered set of $m-$ elements. Or, $\vec{A} = [a_1, a_2, ..., a_m]$, where $a_i \in F, \forall i = 1, 2, ..., m$. The component $a_i, i = 1, 2, ..., m$.

$\vec{A} = \begin{bmatrix} a_1 & a_2 & . & . & . & a_m \end{bmatrix}$

is called a row vector.

While,

$$\vec{A} = \begin{bmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_m \end{bmatrix}$$

is called a column vector.

Note:

(i) A matrix $A_{m\times n}$ can be expressed as a vector in m rows and n columns.

(ii) The laws of the operations on matrices apply to vectors as well.

Example 5.26 If $\vec{x}_1 = [1, -3, 5, 7], \vec{x}_2 = [1, -2, -1, -5]$. Find:

(i) $\vec{x}_1 \boxplus 2\vec{x}_2$.

(ii) $\vec{x}_1 \boxminus 4\vec{x}_2$.

(iii) $-\frac{1}{3}\vec{x}_1 \boxdot 3\vec{x}_2$.

(i) $\vec{x}_1 \boxplus 2\vec{x}_2 = [1, -3, 5, 7] \boxplus (2 \boxdot [1, -2, -1, -5]) = [1, -3, 5, 7] \boxplus [2, -4, -2, -10] = [3, -7, 3, -3]$.

(ii) $\vec{x}_1 \boxminus 4\vec{x}_2 = [1, -3, 5, 7] \boxplus (-4 \boxdot [1, -2, -1, -5]) = [1, -3, 5, 7] \boxplus [-4, 8, 4, 20] = [-3, 5, 9, 27]$.

(iii) $-\frac{1}{3}\vec{x}_1 \boxdot 3\vec{x}_2 = (-\frac{1}{3}[1, -3, 5, 7]) \boxdot (3[1, -2, -1, -5]) = [\frac{-1}{3}, 1, \frac{-5}{3}, \frac{-7}{3}] \boxdot [3, -6, -3, -15] = \frac{-1}{3} \times 3 + 1 \times -6 + \frac{-5}{3} \times -3 + \frac{-7}{3} \times -15 = 33$.

5.8.2 Linearly dependent vectors

Definition 5.33 A sequence of vectors $v_1, v_2, ..., v_m$ defined on the field F are said to be linearly dependent, if there exist m scalars $a_1, a_2, ..., a_m$ in F, not all zero, such that;

$a_1 v_1 + a_2 v_2 + ... + a_m v_m = 0$,

where 0 denotes the zero vector (Shilov, 1977; Shilov, 2012).

5.8.3 Linearly independent vectors

Definition 5.34 In the previous definition, if there is no such scalars $a_1, a_2, ..., a_m$, the vectors $v_1, v_2, ..., v_m$ are said to be linearly independent (Shilov, 1977; Shilov, 2012).

Example 5.27 Consider the vectors $v_1 = [2, 2, -3], v_2 = [-4, -4, 6]$, $v_3 = [3, 1, -4]$. Prove that:

(i) v_1, v_2 are perpendicular vectors.

(ii) v_1, v_3 are not perpendicular vectors.

Solution:

(i) $2v_1 + v_2 = 2[2, 2 - 3] + [-4, -4, 6] = [4, 4, -6] + [-4, -4, 6] = [0, 0, 0]$. So, it enables us to find $a_1 = 2, a_2 = 1$, where $2v_1 + v_2 = 0$.

Thereby, v_1, v_2 are perpendicular vectors.

(ii) $a_1 v_1 + a_2 v_2 = a_1[2, 2, -3] + a_2[3, 1, -4] = [2a_1 + 3a_2, 2a_1 + a_2, -3a_1 - 4a_2] = [0, 0, 0]$.

Or, $2a_1 + 3a_2 = 0, 2a_1 + a_2 = 0, -3a_1 - 4a_2 = 0$.

By solving these equations, implies $a_1 = a_2 = 0$.

Thus, v_1, v_3 are not perpendicular vectors.

5.9 Exercises

Solve the following questions:

Q1: Consider the mtrices $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.
Find each of :

(i) A^2, B^2, C^2.

(ii) $A \boxdot B, B \boxdot A, A \boxdot C, C \boxdot B, B \boxdot C, C \boxdot A$.

Q2: Evaluate the result of:

(i) $\begin{bmatrix} -2 & 3 & -6 \\ -5 & -4 & -5 \\ 0 & 1 & -9 \end{bmatrix} \boxplus \begin{bmatrix} -1 & 3 & -4 \\ 0 & 2 & -5 \\ -1 & 0 & -1 \end{bmatrix}$.

(ii) $\begin{bmatrix} -6 & 1 & 0 \\ 4 & -2 & -1 \end{bmatrix} \boxplus \begin{bmatrix} -5 & 0 & -3 \\ 0 & 1 & -3 \end{bmatrix} \boxplus \begin{bmatrix} -1 & 3 & 0 \\ 0 & -4 & 0 \end{bmatrix}$.

Q3: Consider $A = \begin{bmatrix} 1 & 2 & -3 \\ 3 & 4 & 5 \\ 2 & -1 & 7 \end{bmatrix}, B = \begin{bmatrix} 3 & 1 & 5 \\ 1 & -4 & 7 \\ 2 & 1 & 3 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.
Evaluate each of the following:

(i) $3A$.

(ii) $A \boxplus 2B$.

(iii) $A \boxminus 3C$.

(iv) Find the matrix D, in which $A \boxplus D = B$.

(v) Find the matrix E, in which $\frac{A-2C}{2} + 3E = 5B$.

Q4: Assume that each of A, B are matrices in the order 2×2, 0 is a zero matrix, and $k \in F$. Prove that:

(i) $-(A \boxplus B) = (-A) \boxplus (-B)$.

(ii) $B \boxminus A = -(A \boxminus B)$.

(iii) $0 \boxdot k = 0$.

(iv) $k(A \boxplus B) = k(A) \boxplus k(B)$.

(v) $A \boxplus B = B \boxplus A$.

Q5: Consider $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, B = \begin{bmatrix} 2 & -3 \\ -1 & 5 \end{bmatrix}, C = \begin{bmatrix} 1 & 4 \\ 5 & 1 \end{bmatrix}$.
Evaluate each of the following:

(i) $A \boxdot B$.

(ii) $B \boxdot C$.

(iii) $A^2 \boxminus 5A \boxminus B$.

(iv) $A \boxdot B \boxplus A \boxdot C$.

(v) $(A \boxplus B)^2$.

(vi) $A^2 \boxplus 2A \boxplus B + B^2$.

(vii) $A(B \boxdot C)$.

(viii) A^3.

(ix) $A(B \boxplus C)$.

(x) $(A \boxminus B)(A \boxplus B)$.

(xi) $A^2 - B^2$.

(xii) $(A \boxdot B)C$.

Q6: Assume that
$$A = \begin{bmatrix} -1 & -2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}, B = \begin{bmatrix} -3 & -6 & 2 \\ 2 & 4 & -1 \\ 2 & 3 & 0 \end{bmatrix}, C = \begin{bmatrix} -5 & -8 & 0 \\ 3 & 5 & 0 \\ 1 & 2 & -1 \end{bmatrix}.$$
Prove that:

(i) $A^2 = B^2 = C^2 = 1$.

(ii) $A \boxdot B = B \boxplus A = C$.

(iii) $B \boxdot C = C \boxplus B = A$.

(iv) $A \boxdot C = C \boxplus A = B$.

Q7: Assume that

$$L = \begin{bmatrix} 3 & 2 & 2 \\ -1 & 0 & -1 \\ -3 & -3 & -2 \end{bmatrix}, M = \begin{bmatrix} -1 & -2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}, k \in F.$$

Prove that $[kM \boxdot (1-k)L]^2 = 1$.

Q8: If $Z = \begin{bmatrix} -1 & -2 & 1 \\ 2 & 1 & -3 \\ -5 & 2 & 3 \end{bmatrix}, X = \begin{bmatrix} 2 & 5 & -1 & -7 \\ -2 & 1 & 3 & 4 \\ 3 & 2 & 1 & 2 \end{bmatrix},$

$Y = \begin{bmatrix} 3 & 6 & 0 & -6 \\ -1 & 2 & 4 & 5 \\ 4 & 3 & 2 & 3 \end{bmatrix}$, then $ZX = ZY$, but $X \neq Y$.

Q9: Find

$$\left[\begin{array}{ccc|cc} 0 & 2 & 1 & 0 & 0 \\ -2 & 0 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] \boxdot \left[\begin{array}{cc|ccc} 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & -2 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right].$$

Q10: Find

$$\left[\begin{array}{cc|cc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 1 \end{array}\right]^2.$$

Q11: Find

$$\left[\begin{array}{c|cc|cc} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right]^2.$$

Q12: Prove the following below:

(i) $(A^T)^T = A, \forall A.$

(ii) $(A \boxplus B)^T = A^T \boxplus B^T, \forall A, B.$

(iii) $(A \boxdot B)^T = B^T \boxdot A^T, \forall A, B.$

(iv) Every square matrix can be expressed as the sum of a symmetric matrix and a skew symmetric matrix.

(v) If A is a skew symmetric matrix, then A^2 is a symmetric matrix.

(vi) If A is a squar matrix, then $A \boxdot A^T$ is a symmetric matrix.

(vii) If A is a squar matrix, then $A \boxplus A^T$ is a symmetric matrix.

(viii) If A is a squar matrix, then $A \boxminus A^T$ is a skew symmetric matrix.

(ix) If A, B are skew symmetric matrices, then each of $A \boxplus B, A \boxminus B$ is a skew symetric matrix.

(x) Give an example to show that, if A, B are symmetric, where $A \boxdot B$ is not a symmetric matrix.

Q13: Assumme that $A \boxdot B = C$, where;

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 2 & 4 \\ 2 & 0 & 2 & 0 & 1 & 6 \\ 5 & 0 & 0 & 1 & 3 & 2 \\ 3 & 1 & 1 & 2 & 2 & 1 \\ 2 & 0 & 2 & 3 & 2 & 0 \\ 4 & 1 & 0 & 4 & 2 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 & 1 & 2 & 1 & 7 \\ 2 & 0 & 2 & 1 & 6 & 8 \\ 3 & 0 & 3 & 2 & 1 & 8 \\ 4 & 1 & 1 & 1 & 6 & 2 \\ 5 & 0 & 2 & 2 & 1 & 1 \\ 6 & 0 & 3 & 1 & 6 & 3 \end{bmatrix}.$$

Find each of the following:

(i) C_{43}.

(ii) C_{55}.

(iii) C_{53}.

(iv) C_{46}.

(v) C_{66}.

(vi) C_{53}.

Q14: Consider $v_1 = [1, 2, -3, 4], v_2 = [3, -1, 2, 1], v_3 = [1, -5, 8, -7]$.

Prove that these vectors are perpendicular.

6
Determinants

6.1 Introduction

Determinants are mathematical objects that are very useful in the analysis and solution of systems of linear equations. The determinant is a scalar value that is a function of the entries of a square matrix. The determinant of a matrix A is commonly denoted $det(A)$, or $|A|$. Its value characterizes some properties of the matrix and the linear map represented by the matrix.

Determinant occurs throughout mathematics. It can be used:

(i) To represent the coefficients in a system of linear equations.

(ii) To solve equations (Cramer's rule), although other methods of solution are computationally much more efficient.

(iii) For defining the characteristic polynomial of a matrix, whose roots are the eigenvalues.

(iv) The signed $n-$dimensional volume of a $n-$dimensional parallelepiped is expressed by a determinant.

(v) The determinant of (the matrix of) a linear transformation determines how the orientation and the $n-$dimensional volume are transformed.

(vi) In calculus with exterior differential forms and the Jacobian determinant, in particular for changes of variables in multiple integrals.

6.2 The concept of determinate and its precise definition

Definition 6.1 A determinate is a function whose domain is the set of square matrices and whose codomain of F, and denoted by $det(A)$ or $|A|$. (Hoffman and Kunze, 1967; Hohn, 1972; Strang, 2012; Strang, 2022; Whitelaw, 2019; Lang, 2012b).

From this definition we can find the mathematical definition of the determinate as follows;

Definition 6.2 Let $\mathbb{M} = \{\bigcup_{i=1}^{n} A_i : A_i \text{ is a nonsingular squar matrix}\}$. The $|A| = f : \mathbb{M} \to F$.

6.3 Types of determinants

Generally, there are commonly three types of determinants.

6.3.1 First order determinant

This is used for the calculation of the determinant for a matrix of order 1. For example, if $[a] = A$, then the determinant of A will be equal to a.

Example 6.1 (i) If $A = [a]$, then $|A| = a, \forall a$.

(ii) If $A = [5]$, then $|A| = 5$.

(iii) If $A = [-\frac{3}{8}]$, then $|A| = \left|-\frac{3}{8}\right| = -\frac{3}{8}$.

6.3.2 Second order determinant

This is used for matrices of order 2. The determinant of a matrix of order 2 can be calculated by first multiplying the diagonally opposite elements in the matrix and then finding the difference between these two products. Or, It is the product of the elements of the main diagonal minus the product of the elements of the secondary diagonal.

Example 6.2 (i) If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$

(ii) If If $A = \begin{bmatrix} -3 & -\frac{7}{9} \\ -4 & 11 \end{bmatrix}$, then $\begin{vmatrix} -3 & -\frac{7}{9} \\ -4 & 11 \end{vmatrix} = (-3)(11) - (-4)(-\frac{7}{9}) = -\frac{325}{9} = -35\frac{1}{9}.$

6.3.3 Third order determinant

This is used for matrices of order 3. The determinant of a matrix of order 3 can be calculated by first adding the product of the diagonally opposite elements of the matrix and then subtracting the sum of elements perpendicular to the line. Or, It is defined as the algebraic sum of the following product, as shown in Figurs 6.1 and 6.2, taking into account the sign:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}).$$

Or,

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Example 6.3 $|A| = \begin{vmatrix} 3 & 4 & 5 \\ -4 & 6 & 3 \\ 1 & -4 & 3 \end{vmatrix} = 3\begin{vmatrix} 6 & 3 \\ -4 & 3 \end{vmatrix} - 4\begin{vmatrix} -4 & 3 \\ 1 & 3 \end{vmatrix} + 5\begin{vmatrix} -4 & 6 \\ 1 & -4 \end{vmatrix} = 3(30) - 4(-15) + 5(10) = 90 + 60 + 50 = 200.$

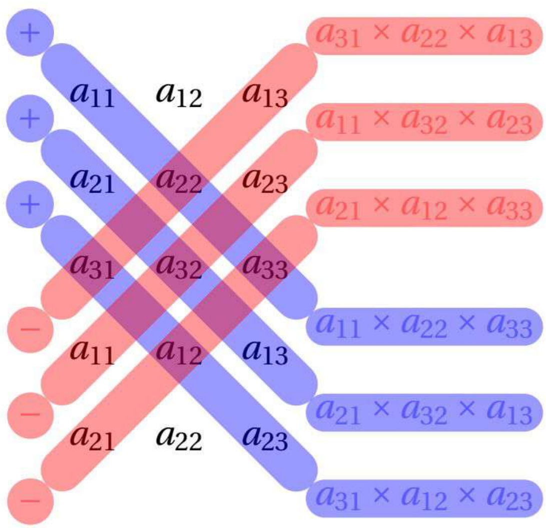


Figure 6.1: Determinant of a 3×3 matrix

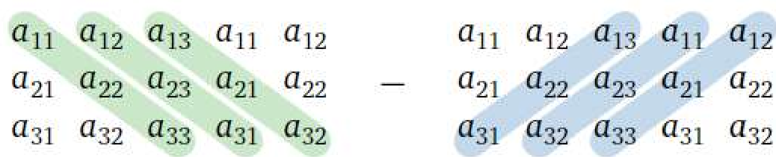


Figure 6.2: Cofactor of a 3×3 matrix

6.3.4 Algorithm for finding determinant of a matrix of third order or higher

In the previous section, we have seen that the determinant of matrix is the sum of products of elements of any row (or any column) and their corresponding cofactors. Thus, here are the steps to find the determinant of matrix (a 3×3 matrix or any other matrix).

- Step1: Choose any row or column. We usually choose the first row to find the determinant.
- Step2: Find the co-factors of each of the elements of the row (column) that we have chosen in Step 1.
- Step3: Multiply the elements of the row (column) from Step 1 with the corresponding cofactors obtained from Step 2.
- Step 4: Add all the products from Step 3 which would give the determinant of the matrix.

6.4 General methods for finding determinants

There are some general methods for finding the determinants of square matrices, such as; cofactors and permutations. In this section, we will discuss them in detail in terms of the algorithm, and illustrative examples to prove the effectiveness of the methods and their practical applicability.

6.4.1 Finding the determinants using cofactors

Definition 6.3 Let A be a square matrix of size n. The (i, j) minor refers to the determinant of the $(n-1) \times (n-1)$ submatrix A_{ij} formed by deleting the *ith* row and *jth* column from A (or sometimes just to the submatrix A_{ij} itself). The corresponding cofactor is the signed minor

$(-1)^{i+j} |A_{ij}|$ (Lang, 1984; Shilov, 2012; Britton and Snively, 1954; Hoffman and Kunze, 1967; Hohn, 1972; Strang, 2006; Norman, 1986).

Example 6.4 Compute the determinant of

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Solution: $A_{11} = (-1)^{1+1} a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}.$

$A_{12} = (-1)^{1+2} a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = -a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}.$

$A_{13} = (-1)^{1+3} a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$

Thus,

$$|A| = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$
$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$
$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

Example 6.5 Find the determinant of $A = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix}$.

Solution: The cofactor expansion of A along the first column is

$$|A| = 1 \begin{vmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} - 1 \begin{vmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} + 1 \begin{vmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{vmatrix} - (-1) \begin{vmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -11 & 1 & 1 \end{vmatrix}.$$

Calculating the 3 by 3 determinant in each term,

$$\begin{vmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = -4, \begin{vmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 4, \begin{vmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = -4, \begin{vmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -11 & 1 & 1 \end{vmatrix} = -4.$$

Thus, we obtain $|A| = 1(-4) - 1(4) + 1(-4) - 1(-1)(-4) = -16$.

6.4.2 Finding the determinants using permutations

This method depends on the permutation of a set whose number of elements is equal to the degree of the square matrix which the goal is to finding its determinant. If the degree of a matrix is n, then this method depends on all permutations $(1, 2, ..., n)$.

Definition 6.4 A permutation of a set is an arrangement of its members into a sequence or linear order, or if the set is already ordered, a rearrangement of its elements. The permutation refers to the process of changing the linear order of an ordered set (Gove, 1963).

Mathematically, a permutation of a set S is defined as a bijection from S to itself. That is, it is a function $f : S \to S$ for which every element occurs exactly once as an image value. Such a function $f : S \to S$ is equivalent to the rearrangement of the elements of S in which each element $i = f(i), \forall i \in S$, described by the function σ (McCoy, 1968; Nering, 1970).

Example 6.6 Consider $S = \{1, 2, 3\}$. The permutations of S are all bijective functions $\sigma_i : S \to S, i = 1, 2, ..., 6$ as shown below:

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$
$$\sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \sigma_6 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

6.5 Permutations and the determinant

We can find the determinant of a square matrix of degree $A_{n\times n}$ as follows (Lankham et al., 2006; Zake Sheet, 2021):

$$|A| = \sum_{i=1}^{n} \bar{\sigma_i} a_1 \sigma_i(1) a_2 \sigma_i(2) ... a_n \sigma_i(n)$$

where σ_i is the permutation i, and $\sigma_i(k)$ is the value of the permutation at the point k in whech can be expressed ad follows:

$$\bar{\sigma_i} = \frac{[(\sigma_i(1) - \sigma_i(2))(\sigma_i(1) - \sigma_i(3)...(\sigma_i(1) - \sigma_i(n)))][(\sigma_i(2) - \sigma_(3))]...}{(1-2)(1-3)...(1-n)(2-3)...}$$

The amount can be written as follows:

$$\bar{\sigma_i} = \prod_{k<l} \frac{\sigma_i(k) - \sigma_i(1)}{k-1}$$

Example 6.7 Find $|A|$ of $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$.

Slution:
$|A| = \sum_{i=1}^{n} \bar{\sigma_i} a_1 \sigma_i(1) a_2 \sigma_i(2)$.
$\sigma_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. $\bar{\sigma_1} = \frac{1-2}{1-2} = 1, \bar{\sigma_2} = \frac{2-1}{1-2} = -1$.
Thus, $|A| = 1 \times a_{11}a_{22} + (-1)a_{12}a_{22} = a_{11}a_{22} - a_{12}a_{22}$.

Example 6.8 Find $|B|$ of $B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$.

Slution:

$$\begin{aligned} |B| &= \sum_{i=1}^{n} \bar{\sigma_i} b_1 \sigma_i(1) b_2 \sigma_i(2) b_3 \sigma_i(3) \\ &= \bar{\sigma_1} b_1 \sigma_1(1) b_2 \sigma_1(2) b_3 \sigma_1(3) + \bar{\sigma_2} b_1 \sigma_2(1) b_2 \sigma_2(2) b_3 \sigma_2(3) \\ &+ \bar{\sigma_3} b_1 \sigma_3(1) b_2 \sigma_3(2) b_3 \sigma_3(3) + \bar{\sigma_4} b_1 \sigma_4(1) b_2 \sigma_4(2) b_3 \sigma_4(3) \\ &+ \bar{\sigma_5} b_1 \sigma_5(1) b_2 \sigma_5(2) b_3 \sigma_5(3) + \bar{\sigma_6} b_1 \sigma_6(1) b_2 \sigma_6(2) b_3 \sigma_6(3) \end{aligned}$$

where,
$\sigma_1 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}, \sigma_2 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}, \sigma_3 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}$,
$\sigma_4 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}, \sigma_4 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}, \sigma_6 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}$.
and that,

$$\begin{aligned}
\bar{\sigma_1} &= \frac{(\sigma_1(1)-\sigma_1(2))(\sigma_1(1)-\sigma_1(3))(\sigma_1(2)-\sigma_1(3))}{(1-2)(1-3)(2-3)} \\
&= \frac{(1-2)(1-3)(2-3)}{(1-2)(1-3)(2-3)} = 1 \\
\bar{\sigma_2} &= \frac{(\sigma_2(1)-\sigma_2(2))(\sigma_2(1)-\sigma_2(3))(\sigma_2(2)-\sigma_2(3))}{(1-2)(1-3)(2-3)} \\
&= \frac{(2-1)(2-3)(1-3)}{(1-2)(1-3)(2-3)} = -1 \\
\bar{\sigma_3} &= \frac{(\sigma_3(1)-\sigma_3(2))(\sigma_3(1)-\sigma_3(3))(\sigma_3(2)-\sigma_3(3))}{(1-2)(1-3)(2-3)} \\
&= \frac{(1-3)(1-2)(3-2)}{(1-2)(1-3)(2-3)} = -1 \\
\bar{\sigma_4} &= \frac{(\sigma_4(1)-\sigma_4(2))(\sigma_4(1)-\sigma_4(3))(\sigma_4(2)-\sigma_4(3))}{(1-2)(1-3)(2-3)} \\
&= \frac{(3-2)(3-1)(2-1)}{(1-2)(1-3)(2-3)} = -1 \\
\bar{\sigma_5} &= \frac{(\sigma_5(1)-\sigma_5(2))(\sigma_5(1)-\sigma_5(3))(\sigma_5(2)-\sigma_5(3))}{(1-2)(1-3)(2-3)} \\
&= \frac{(2-3)(2-1)(3-1)}{(1-2)(1-3)(2-3)} = 1 \\
\bar{\sigma_6} &= \frac{(\sigma_6(1)-\sigma_6(2))(\sigma_6(1)-\sigma_6(3))(\sigma_6(2)-\sigma_6(3))}{(1-2)(1-3)(2-3)} \\
&= \frac{(3-1)(3-2)(1-2)}{(1-2)(1-3)(2-3)} = 1
\end{aligned}$$

Thus,

$$\begin{aligned}
|B| &= (1)b_{11}b_{22}b_{33} + (-1)b_{12}b_{21}b_{33} + (-1)b_{11}b_{23}b_{32} \\
&+ (-1)b_{13}b_{22}b_{31} + (+1)b_{12}b_{23}b_{31} + (+1)b_{13}b_{21}b_{32} \\
&= b_{11}b_{22}b_{33} - b_{12}b_{21}b_{33} - b_{11}b_{23}b_{32} \\
&- b_{13}b_{22}b_{31} + b_{12}b_{23}b_{31} + b_{13}b_{21}b_{32}.
\end{aligned}$$

Which is the same result if we try to find it in other ways.

6.6 Exercises

Q1: Find the determinants of the following matrices in three different ways:

(i) $\begin{bmatrix} 2 & 1 \\ 3 & 3 \end{bmatrix}$

(ii) $\begin{bmatrix} 1 & 2 & 5 \\ 0 & 2 & 0 \\ 1 & 3 & 2 \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 2 & 1 \end{bmatrix}$

(iv) $\begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 0 & 0 & 1 \\ 3 & 1 & 2 & 1 \\ 0 & 3 & 1 & 0 \end{bmatrix}$

Q2: Use the easiest method to find the determinant of the following matrix:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 2 & 0 & 1 & 0 & 2 \\ 1 & 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 2 & 1 \end{bmatrix}$$

6.7 Properties of determinants

A determinant is a unique number that can be ascertained from a square matrix. The determinants of a Matrix say K is represented as $|A|$. The determinants and its properties are useful as they enable us to obtain the same outcomes with distinct and simpler configurations of elements. The determinant is considered an important function as it satisfies some additional properties of determinants that are derived from the following conditions:

(i) Multiplicativity; $|AB| = |A|\,|B|\,; \forall A, B.$

(ii) Invariance under transpose; $|A| = \left|A^T\right|, \forall A.$

(iii) Invariance under row operations; if A' is a Matrix formed by summing up the multiple of any row to another row, then $|A| = |A'|\,, \forall A.$

(iv) There is a change of sign under row swap. If A' is a Matrix made by interchanging the positions of two rows, then $|A'| = -\,|A|\,, \forall A.$

There are some important properties of determinants that are widely used. These properties make calculations easier and also are helping in solving various kinds of problems. The description of each of the important properties of determinants is given below.

(i) All zero property. The determinants will be equivalent to zero if each term of rows and columns are zero.

Example 6.9 (1). $\begin{vmatrix} 2 & 0 & 1 \\ 1 & 0 & 1 \\ 3 & 0 & 2 \end{vmatrix} = 0.$ (2). $\begin{vmatrix} 2 & 1 & 2 & 1 & 3 \\ 2 & 4 & 3 & 2 & 1 \\ 1 & 3 & 2 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 1 & 2 & 2 \end{vmatrix} = 0$

(ii) If the Matrix A^T is the transpose of matrix A, then $|A| = \left|A^T\right|.$

Example 6.10 $|A| = \begin{vmatrix} 1 & k & 2 & 5 \\ 2 & l & 1 & 5 \\ 3 & m & 3 & 5 \\ 4 & n & 4 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ k & l & m & n \\ 2 & 1 & 3 & 4 \\ 5 & 5 & 5 & 5 \end{vmatrix} = \left|A^T\right|.$
Verifying the results is left to the reader as an exercise.

(iii) If the elements in the two rows (columns) of a square matrix are equal, then its determinant equals zero.

Example 6.11 $\begin{vmatrix} 1 & -3 & 5 & 7 & 11 \\ a & b & c & d & e \\ 6 & t & 7 & 8 & -3 \\ 1 & -3 & 5 & 7 & 11 \\ 4 & 3 & 5 & 7 & -3 \end{vmatrix} = 0.$ Because its first and fourth rows are equal, and verifying of the result is left to the reader as an exercise.

(iv) If the elements of a square matrix are complex numbers, then the determinant of the conjugate of the matrix is equal to the conjugate of the determinant of the matrix.

Example 6.12 If $A = \begin{bmatrix} 3i & 2+i \\ i-1 & 2i \end{bmatrix}$, then $\bar{A} = \begin{bmatrix} -3i & 2-i \\ -i-1 & -2i \end{bmatrix}$.

$|A| = (3i)(2i) - (i-1)(2+i) = -3 - i).$

$\left|\bar{A}\right| = (-3i)(-2i) - (-i-1)(2-i) = -(3+i) = -3 + i.$

(v) The determinant of the product of two square matrices is equal to the product of the determinants of those two matrices. Or, $|AB| = |A|\,|B|\,, \forall A, B.$

Example 6.13 $\left|\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x & y \\ p & q \end{bmatrix}\right| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \begin{vmatrix} x & y \\ p & q \end{vmatrix}.$

Solution: $\left|\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x & y \\ p & q \end{bmatrix}\right| = adxq + cdpy - (adpy + cdxq)$ (1).

$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \begin{vmatrix} x & y \\ p & q \end{vmatrix} = adxq + cdpy - (adpy + cdxq)$ (2).

Thus, (1)= (2).

(vi) Sum Property: If a few elements of a row or column are expressed as a sum of terms, then the determinant can be expressed as a sum of two or more determinants. Or,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

The elements of the first row represent the sum of terms, which can be split into two different determinants. Further, the new determinants also have the same second and third row, as the earlier determinant.

Example 6.14 $\begin{vmatrix} 1 & 3 & -5 \\ 1 & -7 & 2 \\ 7 & 3 & 8 \end{vmatrix} + \begin{vmatrix} 1 & 3 & -5 \\ -1 & 2 & 9 \\ 7 & 3 & 8 \end{vmatrix} = \begin{vmatrix} 1 & 3 & -5 \\ 1 & -7 & 2 \\ 7 & 3 & 8 \end{vmatrix}$.

Solution:

$$\begin{vmatrix} 1 & 3 & -5 \\ 1 & -7 & 2 \\ 7 & 3 & 8 \end{vmatrix} + \begin{vmatrix} 1 & 3 & -5 \\ -1 & 2 & 9 \\ 7 & 3 & 8 \end{vmatrix} = \begin{vmatrix} 1 & 3 & -5 \\ 1 & -7 & 2 \\ 7 & 3 & 8 \end{vmatrix}$$

$$-304 + 287 = -17$$

$$-17 = -17$$

(vii) Multiplication Property: The value of the determining becomes k times the earlier value of the determinant if each of the elements of a particular row or column is multiplied with a constant k. Or,

$$A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, B = \begin{vmatrix} ka_{11} & ka_{12} & ka_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$|B| = k\,|A|\,.$$

The elements of the first row are multiplied with a constant k, and the determinant value is also multiplied with the constant k. This property helps in taking a common factor from each row or a column of the determinant.

Example 6.15 $2\begin{vmatrix} -1 & 7 \\ \frac{-3}{4} & 2 \end{vmatrix} = \begin{vmatrix} -1 & 14 \\ \frac{-3}{4} & 4 \end{vmatrix}$.

Solution:

$$2\begin{vmatrix} -1 & 7 \\ \frac{-3}{4} & 2 \end{vmatrix} = \begin{vmatrix} -1 & 14 \\ \frac{-3}{4} & 4 \end{vmatrix}$$

$$2(\frac{13}{4}) = \frac{13}{2}$$

$$\frac{13}{2} = \frac{13}{2}.$$

(viii) Sign Property: The sign of the value of the determinant changes if any two rows or any two columns are interchanged. Or,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{vmatrix}.$$

Example 6.16 $\begin{vmatrix} 3 & 5 & -9 \\ -3 & 0 & 2 \\ 12 & 10 & 21 \end{vmatrix} = - \begin{vmatrix} 3 & 5 & -9 \\ 12 & 10 & 21 \\ -3 & 0 & 2 \end{vmatrix}.$

Solution:

$$\begin{vmatrix} 3 & 5 & -9 \\ -3 & 0 & 2 \\ 12 & 10 & 21 \end{vmatrix} = - \begin{vmatrix} 3 & 5 & -9 \\ 12 & 10 & 21 \\ -3 & 0 & 2 \end{vmatrix}$$

$$645 = -645.$$

(ix) Proportionality or Repetition: If the corresponding elements of two rows (columns) in a square matrix are proportional, then the determinant of that matrix is zero.

Example 6.17 $\begin{vmatrix} 1 & 2 & 3 \\ 1 & 7 & 2 \\ 7 & 14 & 21 \end{vmatrix} = 0.$

Since the ratio between the corresponding elements from the first row to the third row is $\frac{1}{7}$, thereby, the value of the determinant is equal to zero.

(x) Property Of Invariance: If each element of a row and column of a determinant is added with the equimultiples of the elements of another row or column of a determinant, then the value of the determinant remains unchanged. This can be expressed in the form of a formula as; $R_i \to R_i + kR_j$, or $C_i \to C_i + kC_j$.

Example 6.18

$$\begin{vmatrix} 1 & 2 & 3 \\ 1 & 7 & 2 \\ 7 & 14 & 21 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 1 & 7 & 2 \\ \frac{15}{2} & 15 & \frac{45}{2} \end{vmatrix}.$$

Solution:

$$\begin{vmatrix} 1 & 2 & 3 \\ 1 & 7 & 2 \\ 4 & 12 & 20 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 1 & 7 & 2 \\ \frac{9}{2} & 13 & \frac{43}{2} \end{vmatrix},$$

$$44 = 44.$$

We added to the fourth row half of the product of the elements of the first row.

(xi) Minors and cofactors:

(a) A determinant of order 3 will have 9 minors, and each minor will be a determinant of order 2, and a determinant of order 4 will have 16 minors, and each minor will be a determinant of order 3.

(b) $a_{11}C_{21} + a_{12}C_{22} + a_{13}C_{23}$. Or, cofactor multiplied to different row (column) elements results in zero value.

Example 6.19

$$\begin{aligned} |A| &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{21}(-1)^{3+1} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} + a_{22}(-1)^{3+2} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \\ &+ a_{23}(-1)^{3+3} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \\ &= 0. \end{aligned}$$

(xii) Triangular property: If the elements above or below the main diagonal are equal to zero, then the value of the determinant is equal to the product of the elements of the diagonal matrix. Or,

$$|A| = \begin{vmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}, |B| = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{vmatrix}.$$

$$|A| = |B| = \sum_{i=1}^{4} a_{ii} = a_{11} + a_{22} + a_{33} + a_{44}.$$

Note: The determinant of a unit matrix is always equal to $|I| = 1$. A unit matrix can be defined as a scalar matrix in which all the diagonal elements are equal to 1 and all the other elements are zero. Unit matrix is also called the identity matrix.

6.8 Various examples

Example 6.20 Evaluate $\begin{vmatrix} 1 & 2 & 3 & 4 \\ a+1 & b+2 & c+3 & d+4 \\ a & b & c & d \\ 4 & 2 & 1 & 3 \end{vmatrix}$.

Solution:

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ a+1 & b+2 & c+3 & d+4 \\ a & b & c & d \\ 4 & 2 & 1 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ a & b & c & d \\ a & b & c & d \\ 4 & 2 & 1 & 3 \end{vmatrix} + \begin{vmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ a & b & c & d \\ 4 & 2 & 1 & 3 \end{vmatrix}$$

$$= 0 + 0 = 0.$$

The reader must search for the reasons in the properties of the determinants.

Example 6.21 Evaluate $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \\ 3 & 2 & 1 & 4 \\ 4 & 2 & 3 & 1 \end{vmatrix}$.

Solution:

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \\ 3 & 2 & 1 & 4 \\ 4 & 2 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -3 & -3 & -4 \\ 0 & -4 & -8 & -8 \\ 0 & -6 & -9 & -15 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -3 & -3 & -4 \\ 0 & 0 & -4 & -\frac{8}{3} \\ 0 & 0 & -3 & -7 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -3 & -3 & -4 \\ 0 & 0 & -4 & -\frac{8}{3} \\ 0 & 0 & 0 & -5 \end{vmatrix} = (1)(-3)(-4)(-5) = -60.$$

The reader must search for the reasons in the properties of the determinants.

6.9 Exercises

Solve the following questions:

Q1: Find the value of the following determinants, each according to the required method.

(i) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & 1 & 3 \\ 0 & 1 & 0 & 2 \\ 1 & 4 & 2 & 1 \end{vmatrix}$. [Ans.: -38; via minors and cofactors.]

(ii) $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$. [Ans.: 0; via permutations.]

(iii) $\begin{vmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{vmatrix}$. [Ans.: -2; in any way.]

Q2: Prove that the following determinants are correct without decoding them:

(i) $\begin{vmatrix} 1 & a & a+c \\ 1 & b & c+a \\ 1 & c & a+b \end{vmatrix} = 0.$

(ii) $\begin{vmatrix} x_1+y_1 & x_2+y_2 & x_3+y_3 \\ y_1+z_1 & y_2+z_2 & y_3+z_3 \\ z_1+x_1 & z_2+x_2 & z_3+x_3 \end{vmatrix} = 2\begin{vmatrix} x_1 & y_2 & z_3 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}.$

(iii) $\begin{vmatrix} 2a_1+b_1 & 2b_1+c_1 & 2c_1+a_1 \\ 2a_2+b_2 & 2b_2+c_2 & 2c_2+a_2 \\ 2a_3+b_3 & 2b_3+c_3 & 2c_3+a_3 \end{vmatrix} = 9\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$

(iv) $\begin{vmatrix} yz & x^2 & x^2 \\ y^2 & x^2 & y^2 \\ z^2 & z^2 & xy \end{vmatrix} = \begin{vmatrix} yz & xy & xz \\ xy & xz & yz \\ xz & yz & xy \end{vmatrix} = xyz \neq 0.$

Q3: Prove that each element of the following matrix is equal to its cofactor:

$$\begin{vmatrix} -\frac{6}{7} & \frac{2}{7} & \frac{3}{7} \\ \frac{2}{7} & -\frac{3}{7} & \frac{6}{7} \\ \frac{3}{7} & \frac{6}{7} & \frac{2}{7} \end{vmatrix}.$$

Q4: Without decoding the determinant, show that the following equation in the second degree has a roots a, b, where $a \neq b$:

$$\begin{vmatrix} 1 & x & x^2 \\ 1 & a & a^2 \\ 1 & b & b^2 \end{vmatrix} = 0.$$

Q5: Without decoding the determinant, show that the following equation in the third degree has a root $x = (-a-b)$, then find the other roots:

$$\begin{vmatrix} c & a & b \\ b & x & a \\ a & b & x \end{vmatrix} = 0.$$

Q6: If the points $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ are on the straight line, then:

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0.$$

Q7: Prove that:

$$\begin{vmatrix} a+b & a+c & b+c \\ a+c & a+b & b+c \\ b+c & b+a & a+c \end{vmatrix} = 2(a+b+c) \begin{vmatrix} 1 & c & b \\ 1 & b & b \\ 1 & b & a \end{vmatrix}.$$

7

Inverse of A Matrix

7.1 Introduction

In any mathematical system that can be used to represent and solve real problems, it is a great advantage to have a multiplicative inverse. For the set of rational numbers ($\mathbb{Q}$),the multiplicative inverse is simply the reciprocal:

$$a \cdot \frac{1}{a} = \frac{1}{a} \cdot a = 1$$

Multiplying an element of our set by its inverse yields the identity element, 1. We would like to have such an inverse for square ($n \times n$) matrices. For matrix A, we will call the inverse A^{-1}. Then we have:

$$A \cdot A^{-1} = A^{-1} \cdot A = I$$

where I is the identity matrix, the matrix with 1's along the diagonal, zeros elsewhere. It is also true that;

$$A \cdot I = A, \;\; A^{-1} \cdot I = A^{-1}$$

One important case where the inverse matrix can be very helpful is in solving systems of equations. If we represent a system as:

$$A \cdot \vec{x} = \vec{b}$$

where A is the matrix of coefficients, the vector $\vec{x} = (x_1, x_2, ..., x_n)$ is the $n-$ dimensional vector of variables, and $\vec{b}$ is the vector of solutions of the same dimension. When we solve such a system, it is $\vec{x}$ that we looking for.

Now if we have the inverse of A, then we can also find $\vec{x}$ this way:

$$A \cdot \vec{x} = \vec{b}$$
$$A^{-1} \cdot A\vec{x} = A^{-1} \cdot \vec{b}$$
$$\vec{x} = A^{-1} \cdot \vec{b}$$

Let us discuss elementary transformations and equivalent matrices before delving into the matrix inverse, because such subjects are necessary conditions for the matrix inverse in the following sections.

7.2 Elementary transformations

Elementary transformations are bijective functions that transform important properties a matrix to another and make dealing with the transformed matrix simpler and easier. If a matrix contains n independent vectors, then the transformed matrix contains the same number of independent vectors. These elementary transforms include on six types, as described in the following sections.

7.2.1 The transform $h_i(R)$

This transformation involves multiplying the elements of a ith row in the matrix by an element k, and denoted by $h_i(R)$.

Example 7.1 $h_2(5)(\begin{bmatrix} 1 & 3 & -2 \\ 4 & -5 & \frac{1}{2} \\ -7 & 4 & 20 \end{bmatrix}) = \begin{bmatrix} 1 & 3 & -2 \\ 20 & -25 & \frac{5}{2} \\ -7 & 4 & 20 \end{bmatrix}.$

7.2.2 The transform $h_j(C)$

This transformation involves multiplying the elements of a jth column in the matrix by an element k, and denoted by $h_j(C)$.

Example 7.2 $k_3(-2)(\begin{bmatrix} 1 & 3 & -2 \\ 4 & -5 & \frac{1}{2} \\ -7 & 4 & 20 \end{bmatrix}) = \begin{bmatrix} 1 & 3 & 4 \\ 4 & -5 & -1 \\ -7 & 4 & -40 \end{bmatrix}$.

7.2.3 The transform h_{ij}

This transformation involves replace row i with row j, and denoted by h_{ij}.

Example 7.3 $h_{13}(\begin{bmatrix} 1 & 3 & -2 \\ 4 & -5 & -1 \\ -7 & 4 & 20 \end{bmatrix}) = \begin{bmatrix} -7 & 4 & 20 \\ 4 & -5 & -1 \\ 1 & 3 & -2 \end{bmatrix}$.

7.2.4 The transform k_{ij}

This transformation involves replace column i with column j, and denoted by k_{ij}.

Example 7.4 $k_{23}(\begin{bmatrix} 1 & 3 & -2 \\ 4 & -5 & -1 \\ -7 & 4 & 20 \end{bmatrix}) = \begin{bmatrix} 1 & -2 & 3 \\ 4 & -1 & -5 \\ -7 & 20 & 4 \end{bmatrix}$.

7.2.5 The transform $h_{ij}(k)$

The symmetric elements in row i are added to the elements of row j after multiplying them by the element k, and this transformation denoted by $h_{ij}(k)$.

Example 7.5 $h_{21}(-\frac{1}{2})(\begin{bmatrix} 2 & 4 & 6 & 10 \\ 1 & 2 & 3 & 2 \\ 1 & 1 & 0 & 0 \\ 11 & 0 & 3 & -2 \end{bmatrix}) = \begin{bmatrix} 2 & 4 & 6 & 10 \\ 0 & 0 & 0 & -3 \\ 1 & 1 & 0 & 0 \\ 11 & 0 & 3 & -2 \end{bmatrix}$.

7.2.6 The transform $k_{ij}(k)$

The symmetric elements in column i are added to the elements of column j after multiplying them by the element k, and this transformation denoted by $k_{ij}(k)$.

Example 7.6 $k_{15}(\sqrt{2})(\begin{bmatrix} 2 & 1 & 1 & 3 & \sqrt{2} \\ 1 & 0 & 1 & 2 & \sqrt{2} \\ 3 & 0 & 2 & 1 & -\sqrt{2} \end{bmatrix}) = \begin{bmatrix} 4 & 1 & 1 & 3 & \sqrt{2} \\ 3 & 0 & 1 & 2 & \sqrt{2} \\ 1 & 0 & 2 & 1 & \sqrt{2} \end{bmatrix}$.

7.3 Inversion transformation

Inversion transformations are a natural extension of Poincaré transformations to include all conformal, one-to-one transformations on coordinate space-time (Poincaré, 1913; Minkowski, 1908; Minkowski, 1988). The inverse of each transformation is defined according to its original transformation as follows:

(i) $h_i^{-1}(k) = h_i(\frac{1}{k})$.

(ii) $k_i^{-1}(k) = k_i(\frac{1}{k})$.

(iii) $h_{ij}^{-1} = h_{ij}$.

(iv) $k_{ij}^{-1} = k_{ij}$.

(v) $h_{ij}^{-1}(k) = h_{ij}(-k)$.

(vi) $k_{ij}^{-1}(k) = k_{ij}(-k)$.

We must take into account the inverse transformations in terms of sequence as they do in the inverse combination of functions.

Example 7.7

$$[h_{23}k_{32}k_{21}(3)h_1(2)]^{-1} = h_1^{-1}(2)k_{21}^{-1}(3)k_{32}^{-1}h_{23}^{-1} = h_1(\frac{1}{2})k_{21}(-3)k_{32}h_{23}.$$

Example 7.8

$$h_{21}(1)h_{32}(2)h_2^{-1}(\frac{1}{2})(\begin{bmatrix}1&2&3&5\\2&1&2&6\\1&2&1&7\end{bmatrix})$$
$$=h_{21}(1)h_{32}(2)(\begin{bmatrix}1&2&3&5\\4&2&4&12\\1&2&1&7\end{bmatrix})$$
$$=h_{21}(1)(\begin{bmatrix}1&2&3&5\\4&2&4&12\\9&6&9&31\end{bmatrix})$$
$$=\begin{bmatrix}1&2&3&5\\5&4&7&17\\9&6&9&31\end{bmatrix}.$$

Example 7.9

$$h_2(\frac{1}{2})h_{32}(-2)h_{21}-1(\begin{bmatrix}1&2&3&1\\4&5&7&1\\9&6&9&1\end{bmatrix})$$
$$=h_2(\frac{1}{2})h_{32}(-2)(\begin{bmatrix}1&2&3&1\\4&2&4&0\\9&6&9&1\end{bmatrix})$$
$$=h_2(\frac{1}{2})(\begin{bmatrix}1&2&3&1\\4&2&4&0\\1&2&1&0\end{bmatrix})$$
$$=\begin{bmatrix}1&2&3&1\\2&1&2&0\\1&2&1&0\end{bmatrix}.$$

7.4 Elementary matrices

A square matrix is called an elementary matrix if the result of transforming an identity matrix is one of the elementary transformations, and then the row transformations are equal to the column transformations, that is,

(i) The effect of $h_i(k)$ on I_n is the same as the effect of $k_i(k)$ on I_n. Or, $h_i(k) = h_i(k)(I_n) = k_i(k)(I_n) = k_i(k)$.

(ii) The effect of h_{ij} on I_n is the same as the effect of k_{ij} on I_n. Or, $h_{ij} = h_{ij}(I_n) = k_{ij}(I_n) = k_{ij}$.

(iii) The effect of $h_{ij}(k)$ on I_n is the same as the effect of $k_{ij}(k)$ on I_n. Or, $h_{ij}(k) = h_{ij}(k)(I_n) = k_{ij}(I_n) = k_{ij}(k)$.

From the above, we arrive at the following definition in brief:

Definition 7.1 An elementary matrix is a matrix which differs from the identity matrix by one single elementary row operation (Axler, 2015).

Example 7.10 Let $n = 3$ then:

$$h_{12} = h_{12}(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = k_{12}.$$

$$h_2(2) = h_2(2)(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = k_2(2).$$

$$h_{23}(4) = h_{23}(4)(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} = k_{32}(4).$$

Note:

(i) It is noted that the elementary matrices reflect the operation of transformations on matrices.

(ii) Although the number of elementary matrices is three, the number of transformations is six, and this is because the multiplication of these matrices varies from the left or from the right.

(iii) Multiplying the elementary matrix by another matrix A on the right expresses a transformation in the columns, while Multiplying the elementary matrix by another matrix A on the left expresses a transformation in the rows.

(iv) The product of the elementary matrices on the right is represented by the letter Q, while the product of the elementary matrices on the left is represented by the letter P.

(v) P represents the sum of the transformations of rows, and Q represents the sum of the transformations of columns, as follows:

$$\underbrace{h_n...h_2h_1}_{P}\ A\ \underbrace{k_1k_2...k_m}_{Q} = PAQ.$$

(vi) If the matrix B is a matrix resulting from the transformations of columns and rows of another matrix A, then the relation between A and B is as follows:

$$PAQ = B$$

Or,

$$P^{-1}BQ^{-1} = A$$

whereas,

$$P^{-1} = (h_1^{-1}.h_2^{-1}...h_n^{-1})$$

when,

$$P = h_n...h_2h_1.$$

Example 7.11

$$\underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}_{P}\underbrace{\begin{bmatrix} 1 & 2 & 2 \\ 2 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}}_{A}\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}}_{Q}$$

$$= \begin{bmatrix} 2 & 4 & 4 \\ 1 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 10 & 4 & 12 \\ 3 & 1 & 6 \\ 6 & 2 & 3 \end{bmatrix}$$

$$= B.$$

7.5 Equivalent matrices

Definition 7.2 Two matrices A and B are said to be equivalent if they are of the same order and $B(A)$ can be obtained from $A(B)$ by a sequence of elementary row and column operations, and denoted by $A \sim B$ or $A \equiv B$.

Or, for all A, B be $m \times n$ matrices over the ring $\mathbb{R}$ with identity, there exist; an invertible square matrix P of order n over $\mathbb{R}$ and an invertible square matrix Q of order m over $\mathbb{R}$, such that $B = Q^{-1}AP$ then $A \equiv B$ (Hefferon, 2017; Gradshteyn and Ryzhik, 1980).

Example 7.12 $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 \\ 4 & 2 & 4 \\ 9 & 6 & 9 \end{bmatrix}.$

Note: Two matrices are said to be equivalent if they satisfy the conditions shown below:

(i) Each matrix has the same number of rows.

(ii) Each matrix has the same number of columns.

(iii) The corresponding elements (entries) of each matrix are equal to each other.

Example 7.13 Consider $A = \begin{bmatrix} 3 & -1 \\ 6 & 5 \end{bmatrix}, B = \begin{bmatrix} 3 & -1 \\ 6 & 3 \end{bmatrix}.$

$A \not\equiv B$, because $A_{22} = 5 \neq 3 = B_{22}$.

7.6 Normal form of a matrix

Definition 7.3 If the matrix A can be divided so that in the upper left corner there is an identity matrix, and the rest of the matrices are zero matrices, then it is said to be in the normal form (Lancaster and Tismenetsky, 1985; Roman et al., 2005; Gantmakher, 2000).

Example 7.14 The matrix $A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ is in the norm form, because we can rewrite it in form of $A = \left[\begin{array}{ccc:cc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 \end{array}\right] = \left[\begin{array}{c:c} I_3 & 0 \\ \hdashline 0 & 0 \end{array}\right]$.

7.7 Rank of matrix

Definition 7.4 If A is a matrix, then the degree of the largest determinant of a square submatrix of A that is not equal to zero is the rank of the matrix, and denoted by $rank(A)$ (Axler, 2015; Roman et al., 2005; Bourbaki, 2013).

Example 7.15 Consider $A = \begin{bmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 3 & 5 & 0 \end{bmatrix}$. It can be put in reduced row-echelon form by using the following elementary row operations:

$$\begin{bmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 3 & 5 & 0 \end{bmatrix} \underset{\sim}{\xrightarrow{2R_1+R_2 \to R_2}} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 3 & 5 & 0 \end{bmatrix} \underset{\sim}{\xrightarrow{-3R_1+R_3 \to R_3}} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & -1 & -3 \end{bmatrix}$$
$$\underset{\sim}{\xrightarrow{R_2+R_3 \to R_3}} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \underset{\sim}{\xrightarrow{-2R_2+R_1 \to R_1}} \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}.$$

The final matrix (in reduced row echelon form) has two non-zero rows, and thus the $rank(A) = 2$.

Example 7.16 Find $rank(A) = \begin{bmatrix} 0 & 2 & 3 & 4 \\ 1 & 3 & 5 & 4 \\ 2 & 8 & 13 & 12 \end{bmatrix}$. It can be put in reduced row-echelon form by using the following elementary row and column operations:

$$\begin{bmatrix} 0 & 2 & 3 & 4 \\ 1 & 3 & 5 & 4 \\ 2 & 8 & 13 & 12 \end{bmatrix} \underset{\sim}{\xrightarrow{h_{21}}} \begin{bmatrix} 1 & 3 & 5 & 4 \\ 0 & 2 & 3 & 4 \\ 2 & 8 & 13 & 12 \end{bmatrix} \underset{\sim}{\xrightarrow{h_{31}(-2)}} \begin{bmatrix} 1 & 3 & 5 & 4 \\ 0 & 2 & 3 & 4 \\ 2 & 8 & 13 & 12 \end{bmatrix}$$

$$\underset{\sim}{\xrightarrow{k_{21}(-3)}} \begin{bmatrix} 1 & 0 & 5 & 4 \\ 0 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \end{bmatrix} \underset{\sim}{\xrightarrow{k_{21}(-3)}} \begin{bmatrix} 1 & 0 & 5 & 4 \\ 0 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \end{bmatrix} \underset{\sim}{\xrightarrow{k_{31}(-5)}} \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \end{bmatrix}$$

$$\underset{\sim}{\xrightarrow{k_{41}(-4)}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \end{bmatrix} \underset{\sim}{\xrightarrow{k_2(\frac{1}{2})}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 4 \\ 0 & 1 & 3 & 4 \end{bmatrix} \underset{\sim}{\xrightarrow{k_{32}(-3)}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 4 \\ 0 & 1 & 0 & 4 \end{bmatrix}$$

$$\underset{\sim}{\xrightarrow{k_{42}(-4)}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \underset{\sim}{\xrightarrow{h_{32}(-1)}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= \left[\begin{array}{cc:cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 \end{array}\right] = \left[\begin{array}{c:c} I_2 & 0 \\ \hdashline 0 & 0 \end{array}\right].$$

The final matrix (in reduced row and column echelon form) has two non-zero rows, and thus the $rank(A) = 2$.

Example 7.17 Consider $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \end{bmatrix}$.

It is in the degree 3×4, while $rank(A) = 1$. Because the determinant of the matrix of degree 3×3 is equal to zero, likewise the determinant of the matrix of degree 2×2 is equal to zero.

Note: The best way to find the rank of a matrix is by elementary transformations in order to make the matrix in normal form, from which it is possible to judge its rank.

7.8 Exercises

Q1: Find the rank of each of the following matrices:

(i) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

(ii) $\begin{bmatrix} 2 & 1 & 3 \\ 0 & 2 & -1 \\ 0 & 0 & 7 \end{bmatrix}$.

(iii) $\begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 3 & -5 & 7 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 6 \end{bmatrix}$.

(iv) $\begin{bmatrix} -1 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ 17 & 5 & 2 & 0 \\ 8 & 3 & -1 & 0 \end{bmatrix}$.

(v) $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 3 & 4 & 5 & 6 & 7 \end{bmatrix}$.

(vi) $\begin{bmatrix} 2 & -1 & 0 & 5 & 3 \\ 1 & 5 & -2 & 4 & 7 \\ -1 & 17 & -6 & 2 & 15 \\ 3 & 4 & -2 & 9 & 11 \end{bmatrix}$.

Q2: Show how the rank changes with respect to δ change in the matrix $\begin{bmatrix} 1 & 1 & \delta \\ 1 & \delta & 1 \\ \delta & 1 & 1 \end{bmatrix}$.

Q3: Consider $A = \begin{bmatrix} 2 & -1 & 4 \\ 3 & 0 & 1 \end{bmatrix}$. What are the elementary matrices B, D whose determinants are non-zero and which make the product BAD a matrix in regular form?

Q4: Factor the following matrices into the product of elementary matrices:

(i) $\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}$.

(ii) $\begin{bmatrix} a & 1 & 0 \\ 1 & 0 & 0 \\ b & c & 2 \end{bmatrix}$.

(iii) $\begin{bmatrix} 1 & a & b & c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

Q5: Give an example to illustrate the following relations:

(i) $rank(A+B) = rank(A) + rank(B)$.

(ii) $rank(A-B) = rank(A) - rank(B)$.

(iii) $rank(AB) = rank(A)$, but $rank(AB) < rank(B)$.

(iv) $rank(AB) < rank(A)$, so as $rank(AB) < rank(B)$.

(v) $rank(AB) = rank(A) = rank(B)$.

(vi) $rank(AB) < rank(A)$, but $rank(AB) = rank(B)$.

Q6: Consider $A_{n\times m} \times B_{m\times p} \times C_{p\times n} = I_{n\times n}$.

(i) What can you say about rank A, B, C?

(ii) What can you say about the value of p, m?

Q7: Prove that the rank of the product of two matrices does not exceed the rank of either one of them.

Q8: Prove that the sum of two matrices does not exceed the sum of their ranks.

Q9: If A is a matrix of degree 3×3, and B is the same matrix of degree A with the addition of another column of degree 3×4, is $rank(A)$ can be different from $rank(b)$? Explain with an example.

Q10: Prove that the necessary and sufficient condition for the vectors $A_1, A_2, ..., A_n$ to be independent is the rank of the vectors is equal to m.

7.9 Inverse of a matrix

If two square matrices $A_{n\times n}$ and $B_{n\times n}$ have the property that $AB = I_{n\times n}$ then A and B are said to be inverses of one another and we write $(A = B^{-1}) \wedge (B = A^{-1})$.

A wonderful feature of row reduction is that when we have a matrix equation $AB = C$, we can apply the reduction operations for the matrix A to the rows of A and C simultaneously and ignore B, and what we get will be as true as what we started with.

Let, us start with the matrix equation $AA^{-1} = I$. If we row reduce A so it becomes the identity matrix I, then the left hand side here becomes IA^{-1} which is A^{-1}, the matrix inverse to A. The right hand side however is what we obtain if we apply the row operations necessary to reduce A to the identity, starting with the identity matrix$I$.

Thus, it can be concluded that the inverse matrix, A^{-1} can be obtained by applying the row reduction operations that make A into I starting with I.

After this introduction, we are ready to begin the precise mathematical definition of the concept of matrix inverse as follows:

Definition 7.5 A square matrix $A_{n\times n}$ is called invertible (nonsingular), if there exists a square matrix $B_{n\times n}$ such that $AB = BA = I_n$, where I_n denotes the $n \times n$ identity matrix and the multiplication used is ordinary matrix multiplication. The matrix B is uniquely determined by A, and is called the multiplicative inverse of A, denoted by A^{-1} (Axler, 2015; Weisstein, 2014).

7.10 Matrix inverse methods

There are different ways to find the inverse of a square matrix that its determinant is not equal to zero. Below we review the most important and common of these methods.

7.10.1 The method of adjoint matrix

Consider $A_{n\times n}$, and $|A| \neq 0$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

The matrix of cofactors associated with A is an $n \times n$ matrix, where each element is replaced by its associated cofactor of this matrix, denoted by $C(A)$. Let α_{ij} is a cofactor of the element a_{ij}, then $C(A) = \alpha_{ij}$ (Gantmacher, 1959; Strang, 2006; Householder, 2013).

The transpose of the cofactors of A denoted by:

$$Adj(A) = [C(A)]^T = [C(A)]' = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \dots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \dots & \alpha_{n2} \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{nn} \end{bmatrix}.$$

Thus, the inverse of A will be defined as follows:

$$A^{-1} = \frac{Adj(A)}{|A|} = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \dots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \dots & \alpha_{n2} \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{nn} \end{bmatrix}.$$

Example 7.18 Find the inverse of the matrix $A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & 2 \\ -1 & 2 & 1 \end{bmatrix}$

Solution:

$$\text{Minor matrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{31} \end{bmatrix}$$

$$= \begin{bmatrix} (-1)^{1+1}M_{11} & (-1)^{1+2}M_{12} & (-1)^{1+3}M_{13} \\ (-1)^{2+1}M_{21} & (-1)^{2+2}M_{22} & (-1)^{2+3}M_{23} \\ (-1)^{3+1}M_{31} & (-1)^{3+2}M_{32} & (-1)^{3+3}M_{31} \end{bmatrix}$$

$$= \begin{bmatrix} +M_{11} & -M_{12} & +M_{13} \\ -M_{21} & +M_{22} & -M_{23} \\ +M_{31} & -M_{32} & +M_{31} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{31} \end{bmatrix}$$

$$= \begin{bmatrix} \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & 2 \\ -1 & 2 \end{vmatrix} & \begin{vmatrix} 2 & 1 \\ -1 & 2 \end{vmatrix} \\ -\begin{vmatrix} 2 & -1 \\ 2 & 1 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ -1 & 2 \end{vmatrix} \\ \begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 2 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} 1-4 & -(2+2) & 4+1 \\ -(2+2) & 1-1 & -(2+2) \\ 4+1 & -(2+2) & 1-4 \end{bmatrix} = \begin{bmatrix} -3 & -4 & 5 \\ -4 & 0 & -4 \\ 5 & -4 & -3 \end{bmatrix}$$

$$\therefore Adj(A) = \begin{bmatrix} -3 & -4 & 5 \\ -4 & 0 & -4 \\ 5 & -4 & -3 \end{bmatrix}$$

$$\because |A| = (1)M_{11} - 2M_{12} + (-1)M_{13} = -3 - 8 - 5 = -16$$

$$\therefore A^{-1} = \frac{1}{|A|} \times Adj(A)$$

$$= \frac{1}{-16} \begin{bmatrix} -3 & -4 & 5 \\ -4 & 0 & -4 \\ 5 & -4 & -3 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{3}{16} & \frac{4}{16} & \frac{-5}{16} \\ \frac{1}{4} & 0 & \frac{1}{4} \\ \frac{-5}{16} & \frac{1}{4} & \frac{3}{16} \end{bmatrix}$$

7.10.2 The method of elementary transformations

Let us consider the matrix $A_{n\times n}$. It is possible to find the elementary matrices that give the normal form when multiplied by the right and left of the matrix A (Strang, 2022). Or, it is possible to find P, Q such that:

$$\begin{aligned} &PAQ = I \\ &\therefore A = P^{-1}Q^{-1} \\ &\therefore A^{-1} = (P^{-1}Q^{-1})^{-1} \\ &\therefore A^{-1} = QP \end{aligned}$$

Example 7.19 Find the inverse of the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & 2 & 1 \end{bmatrix}$.

Solution:

$$A = IA$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & 2 & 1 \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{-1}{4} \end{bmatrix} \begin{bmatrix} 1 & 0 & -0 \\ 0 & \frac{-1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{P}$$

$$\times \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & 2 & 1 \end{bmatrix}}_{A}$$

$$= \underbrace{\begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}}_{Q}$$

$$= I_{3\times 3}$$

$$\therefore A^{-1} = \underbrace{\begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}}_{Q}$$

$$\times \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{-1}{4} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{-1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{P}$$

$$= \begin{bmatrix} \frac{5}{12} & \frac{1}{3} & \frac{1}{4} \\ \frac{7}{12} & \frac{2}{3} & \frac{1}{4} \\ \frac{1}{12} & \frac{1}{3} & \frac{1}{4} \end{bmatrix}$$

7.10.3 The method of transformations on rows: Jacobian method

There are six transformations on a matrix, including the three transformations due to rows and three others due to columns. These operations are known as elementary operations. These operations are performed on square matrices only. These elementary operations are:

(i) Interchanging any two rows (columns).

(ii) Multiplication of the elements of any row (column) by a positive integer.

(iii) Addition (subtraction) of multiples of one row (column) to another.

Let us assume the matrix $A_{n\times n}$. The method of transformations on rows is to place the identity matrix beside the matrix whose inverse is to be found and perform operations on the rows in order to make them in normal form, while performing each operation on the identity matrix at the same time. As a result, the identity matrix is transformed into the form of the inverse of the matrix A. Or;

$$PA = I \wedge PI = P$$
$$\therefore A = P^{-1}$$
$$A^{-1} = P$$

Example 7.20 Find the inverse of the following matrix using elementary operations;

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}.$$

Solution:

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}$$

$$A = IA$$

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot A$$

$$h_{12} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot A$$

$$h31(-3) \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & -5 & -8 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & -3 & 1 \end{bmatrix} \cdot A$$

$$h_{12}(-2)h_{32}(5) \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 0 \\ 5 & -3 & 1 \end{bmatrix} \cdot A$$

$$h_{13}(1)h_{23}(-2) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{-1}{2} & \frac{1}{2} \\ -4 & 3 & 1 \\ \frac{5}{2} & \frac{-3}{2} & \frac{1}{2} \end{bmatrix} \cdot A$$

$$\therefore I = B \cdot A$$

$$\therefore A^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{-1}{2} & \frac{1}{2} \\ -4 & 3 & 1 \\ \frac{5}{2} & \frac{-3}{2} & \frac{1}{2} \end{bmatrix}$$

7.10.4 The method of Triangularization

Let us consider the square matrix A, and we like to find its inverse by triangularization method. The method is to decompose this matrix into submatrices. L, D, R.

L is a lower triangular matrix with its diagonal elements are one, D is a diagonal matrix, and R is an upper triangular matrix with its diagonal elements are one. These matrices are easy to find their inverse because;

$$A = LDR$$
$$A^{-1} = R^{-1}D^{-1}L^{-1}$$

Now we explain the algorithm of the method. We are trying to decompose the matrix A into form R. That is, we transform A into an upper triangular matrix, each element of its diagonal is one. In this process, we form L such that all its diagonal elements are zeros; the elements of ${}_{is}a$ in the lower triangle are the number multiplied by a_{ij} making them equal to a_{ij}, starting by a_{11}, a_{22}, ..., etc (Axler, 2010; Axler, 2010; Herstein, 1991).

Whenever we obtain the lower and upper triangular matrix with the required conditions, we have; R, D, where the diagonal of D is simultaneously the diagonal of A. Thereby, the resulting matrix R is such that obtained by the transformation in which the elements of each row are divided by the diagonal element in that row.

Example 7.21 Find the inverse of $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix}$.

Solution: We perform a transformation on the matrix to transform it into a superior triangular matrix as follows:

$$h_{21}(-2)h_{31}(-1)\begin{bmatrix}1&2&3\\2&1&2\\1&1&2\end{bmatrix}=\begin{bmatrix}1&2&3\\0&-3&-4\\1&-1&-1\end{bmatrix}$$

$$h_{32}(\frac{-1}{3})\begin{bmatrix}1&2&3\\0&-3&-4\\1&-1&-1\end{bmatrix}=\underset{DR}{\begin{bmatrix}1&2&3\\0&-3&-4\\0&0&\frac{1}{3}\end{bmatrix}}$$

$$\therefore L=\begin{bmatrix}1&0&0\\2&1&0\\\frac{10}{3}&\frac{1}{3}&1\end{bmatrix},D=\begin{bmatrix}1&0&0\\0&-3&0\\0&0&\frac{1}{3}\end{bmatrix},$$

$$R=\begin{bmatrix}1&2&3\\0&1&\frac{4}{3}\\0&0&1\end{bmatrix},L^{-1}=\begin{bmatrix}1&0&0\\-2&1&0\\\frac{-10}{3}&-10&1\end{bmatrix},$$

$$D^{-1}=\begin{bmatrix}1&0&0\\0&\frac{-1}{3}&0\\0&0&3\end{bmatrix},R^{-1}=\begin{bmatrix}1&-2&\frac{-1}{3}\\0&1&\frac{-4}{3}\\0&0&1\end{bmatrix}.$$

$$\therefore A^{-1}=\begin{bmatrix}1&-2&\frac{-1}{3}\\0&1&\frac{-4}{3}\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&\frac{-1}{3}&0\\0&0&3\end{bmatrix}\begin{bmatrix}1&0&0\\-2&1&0\\\frac{-10}{3}&-10&1\end{bmatrix}$$

$$=\begin{bmatrix}0&1&-1\\\frac{2}{2}&\frac{-1}{3}&0\\\frac{-1}{3}&\frac{-1}{3}0&1\end{bmatrix}.$$

7.10.5 The method of Escalator

This method is of particular importance for finding the inverse of a matrix of degree $(n+1)\times(n+1)$. It is, also named Partition method. The idea is to partition a matrix into smaller submatrices and then calculate the inverse from the given inverse of one of the smaller submatrices (Kosko, 1957).

If the matrix whose inverse is to be found is of degree $(n+1)\times(n+1)$, if we know the inverse of the matrix of degree $n\times n$, which is formed from the first matrix by adding a row and a column such that the rank of the new matrix is not less than the degree of the first matrix

(Pease, 1969).

Let us consider, $A_{(n+1)\times(n+1)}$. We can represent the matrix as follows:

$$A_{(n+1)\times(n+1)} = \left[\begin{array}{c:c} A_{n\times n} & A_{n1} \\ \hdashline A_{1n} & a \end{array}\right]$$

$$A = \left[\begin{array}{c:c} A_1 & A_2 \\ \hdashline A_3 & a \end{array}\right] \text{(For easiness and simplicity)}$$

$$\therefore A^{-1} = \left[\begin{array}{c:c} B_1 & B_2 \\ \hdashline B_3 & b \end{array}\right]$$

$$\therefore AA^{-1} = \begin{bmatrix} A_1 & A_2 \\ A_3 & a \end{bmatrix} \begin{bmatrix} B_1 & B_2 \\ B_3 & b \end{bmatrix}$$

$$= I_{(n+1)\times(n+1)}$$

$$\therefore \begin{bmatrix} A_1B_1 + A_2B_2 & A_1B_2 + A_2b \\ A_3B_1 + aB_3 & A_3B_2 + ab \end{bmatrix} = \begin{bmatrix} I_{nn} & 0 \\ 0 & 1 \end{bmatrix}$$

$\therefore A_1B_1 + A_2B_3 = 1...(1)$

$A_1B_2 + A_2b = 0...(2)$

$A_3B_1 + aB_3 = 0...(3)$

$A_3B_2 + ab = 1...(4)$

From (2), we have: $B_2 = -A_1^{-1}A_2b$.

Substituting this result into (4) we get:$(a - A_3A_1^{-1}A_2)b = 1$

Here, there is only a variable b, which we find it

We find B_b, and also from (1) we get:

$B_1 = A_1^{-1}(I - A_2B_3)$

Substituting this result into (3), we get:$(a - A_3A_1^{-1}A_2)B_3 = -A_3A_1^{-1}$

Thus, we have obtained $B3$, from which we obtain $B1$.

After arranging the matrix, we will obtainA^{-1}.

Example 7.22 Find the inverse of the following matrix by escalator method: $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix}$

Solution:

Divide the matrix into submatrices of a lower degree and the result is:

$$\begin{bmatrix}1&2&3\\2&1&2\\1&1&2\end{bmatrix}=\left[\begin{array}{cc:c}1&2&3\\2&1&2\\ \hdashline 1&1&2\end{array}\right]=\left[\begin{array}{c:c}A_1&A_2\\ \hdashline A_3&a\end{array}\right]$$

$$\therefore A_1=\begin{bmatrix}1&2\\2&1\end{bmatrix},A_2=\begin{bmatrix}3\\2\end{bmatrix},A_3=\begin{bmatrix}1&1\end{bmatrix},a=2$$

$$\therefore A_1^{-1}=\frac{1}{3}\begin{bmatrix}-1&2\\2&-1\end{bmatrix}=\begin{bmatrix}\frac{-1}{3}&\frac{2}{3}\\\frac{2}{3}&\frac{-1}{3}\end{bmatrix}$$

By applying the escalator algorithm as mentioned earlier

, and substituting in,

$(a-A_3A_1^{-1}A_2)b=I$

$$\therefore ([2]-\begin{bmatrix}1&1\end{bmatrix})\cdot\begin{bmatrix}\frac{-1}{3}&\frac{2}{3}\\\frac{2}{3}&\frac{-1}{3}\end{bmatrix}\begin{bmatrix}3\\2\end{bmatrix})[b=I]$$

$$\therefore [\frac{1}{3}]b=I\Rightarrow[b]=[3]$$

After substituting in:$B_2=A_1^{-1}A_2b$ it became;

$$B_2=\begin{bmatrix}\frac{-1}{3}&\frac{2}{3}\\\frac{2}{3}&\frac{-1}{3}\end{bmatrix}\begin{bmatrix}3\\2\end{bmatrix}[3]=\begin{bmatrix}-1\\-4\end{bmatrix}$$

After substituting in: $([a]-A_3A_1^{-1}A_2)B_3=-A_3A_1^{-1}$

$$([a]-A_3A_1^{-1}A_2)B_3=[\frac{1}{3}]$$

$$\therefore [\frac{1}{3}]B_3=-\begin{bmatrix}1&1\end{bmatrix}\begin{bmatrix}\frac{-1}{3}&\frac{2}{3}\\\frac{2}{3}&\frac{-1}{3}\end{bmatrix}\Rightarrow B_3=\begin{bmatrix}-1&-1\end{bmatrix}$$

$$\because B_1=A_1^{-1}(I-A_2B_2)=\begin{bmatrix}\frac{-1}{3}&\frac{2}{3}\\\frac{2}{3}&\frac{-1}{3}\end{bmatrix}(I-\begin{bmatrix}3\\2\end{bmatrix}\begin{bmatrix}-1&-1\end{bmatrix})$$

$$\therefore B_1=\begin{bmatrix}0&1\\2&1\end{bmatrix}$$

$$\therefore A^{-1}=\left[\begin{array}{cc:c}0&1&-1\\2&1&-4\\ \hdashline -1&-1&3\end{array}\right]=\begin{bmatrix}0&1&-1\\2&1&-4\\-1&-1&3\end{bmatrix}.$$

7.11 Exercises

Q1: Find the inverse of the following matrices:

(i) $\begin{bmatrix} -2 & 5 \\ 3 & 4 \end{bmatrix}$.

(ii) $\begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}$.

(iii) $\begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -3 \\ 3 & -2 & 4 \end{bmatrix}$.

(iv) $\begin{bmatrix} 2 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix}$.

(v) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$.

(vi) $\begin{bmatrix} 1 & -1 & 2 \\ 1 & 9 & 3 \\ 1 & 4 & 2 \end{bmatrix}$.

(vii) $\frac{1}{11}\begin{bmatrix} 9 & 6 & -2 \\ -6 & 7 & -6 \\ -2 & 6 & 9 \end{bmatrix}$.

(viii) $\begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

(ix) $\begin{bmatrix} k & c & -b \\ -c & k & a \\ b & -a & k \end{bmatrix}$.

(x) $\begin{bmatrix} 3 & 2 & -1 & 4 \\ 4 & 3 & -1 & 4 \\ -1 & 2 & 4 & 4 \\ 1 & 2 & 0 & 0 \end{bmatrix}$.

Q2: Find the inverse of the following matrices by the method of triangularization:

(i) $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \\ 3 & 2 & 1 & 3 \\ 4 & 4 & 1 & 2 \end{bmatrix}$.

(ii) $C = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 2 & 1 & 2 \\ 3 & 0 & 2 & 1 \\ 2 & 1 & 0 & 3 \end{bmatrix}$.

Q3: Prove the following properties for an invertible matrix A:

(i) $(A^{-1})^{-1} = A$.

(ii) $(kA)^{-1} = k^{-1}A^{-1}, k \neq 0$.

(iii) $(A^T)^{-1} = (A^{-1})^T$.

(iv) $|A^{-1}| = |A|^{-1}$.

(v) For any invertible matrices $n \times n$, A, B, $(AB)^{-1} = B^{-1}A^{-1}$.

(vi) If $A_1, A_2, ..., A_{n-1}, A_n$ are invertible $n \times n$ matrices, then: $(A_1A_2..., A_{n-1}A_n)^{-1} = A_n^{-1}A_{n-1}^{-1}...A_2^{-1}A_1^{-1}$.

Q4: Prove that if A is a symmetric matrix, and $|A| \neq 0$ then A^{-1} is a symmetric matrix.

7.12 Inverse of a complex matrix

If we like to deal with matrices with complex entries, systems of linear equations with complex coefficients, complex variables, complex solutions, determinants of complex matrices, inverse of complex matrices, and vector spaces with scalar multiplication by any complex number allowed. Moreover, the proofs of most facts, and theorems about the real version of these concepts extend easily to the complex case (Nicholson, 2020).

In what follows, we focus our attention on the precise definition of a complex matrix and how to find its inverse, given its utmost importance in all fields in general and in mathematics in particular.

Definition 7.6 An $m \times n$ complex matrix is a rectangular array of complex numbers arranged in m rows and n columns. The set of all $m \times n$ complex matrices is denoted as $A^{\mathbb{C}}_{m\times n}$, or complex $A_{m\times n}$ (Andrilli and Hecker, 2022).

7.13 Method to find the inverse of a complex matrix

Let $A = iB$ be a complex matrix, where A, B are real matrices. Let its inverse be $C + iD$. Then we have to find C and D such that:

$$(A + iB)(C + iD) = I \tag{7.1}$$

(i) At least one of the matrices A or B is nonsingular.

(a) Suppose A is nonsingular so that A^{-1} exists. Then, from (7.1),

$$(AC - BD) + i(AD + BC) = I$$

Comparing real and imaginary parts, we have:

$$AC - BD = I \tag{7.2}$$

$$AD + BC = O \tag{7.3}$$

Premultiplying (7.2) by A^{-1}, we get;

$$C_A^{-1}BD = A^{-1}I \tag{7.4}$$

Premultiplying (7.3) by A^{-1}, we get;

$$D + A^{-1}BC = O$$

$$\therefore D = -A^{-1}BC \tag{7.5}$$

From (7.4) and (7.5);

$$\begin{aligned}
&C - A^{-1}B(-A^{-1}BC) = A^{-1}I\\
&C + A^{-1}BA^{-1}BC = A^{-1}I\\
&AC + BA^{-1}BC = I \quad \text{(Premultiplying by } A)\\
&(A + BA^{-1}B)C = I
\end{aligned}$$

$$\therefore C = (A + BA^{-1}B)^{-1} \tag{7.6}$$

From (7.5) and (7.6),

$$D = -A^{-1}B(A + BA^{-1}B)^{-1} \tag{7.7}$$

(b) Suppose B is nonsingular so that B^{-1} exists. Then we will have:

$$C = B^{-1}A(AB^{-1}A + B)^{-1} \tag{7.8}$$

$$D = -(AB^{-1}A + B)^{-1} \tag{7.9}$$

(c) Suppose A, B both are nonsingular so that A^{-1}, B^{-1} both exist then the above results will be, of course, identical.

$$\begin{aligned}
C &= B^{-1}A(AB^{-1}A + B)^{-1}\\
&= (A^{-1}B)^{-1}(AB^{-1}A + B)^{-1}\\
&= [(AB^{-1}A + B)(A^{-1}B)]^{-1}\\
&= (AB^{-1}AA^{-1}B + BA^{-1}B)^{-1}
\end{aligned}$$

$$\therefore C = (A + BA^{-1}B)^{-1} \tag{7.10}$$

Similarly, we have;

$$\begin{aligned} D &= -A^{-1}B(A+BA^{-1}B)^{-1} \\ &= -(B^{-1}A)^{-1}(A+BA^{-1}B)^{-1} \\ &= -[(A+BA^{-1}B)(B^{-1}A)]^{-1} \end{aligned}$$

$$\therefore D = -(AB^{-1}A+B)^{-1} \tag{7.11}$$

(ii) Suppose both A, B are singular but $A + iB$ is not. Then we use the following method.

Let

$$\begin{aligned} F &= A + rB \\ G &= B - rA \end{aligned} \tag{7.12}$$

where r is a real number such that F or G becomes nonsingular.

$$\begin{aligned} &\because F + iG = (A+rB) + i(B-rA) \\ &= (A+iB) - ir(A+iB) \\ &= (1-ir)(A+iB) \\ &\therefore A + iB = \frac{1}{1-ir}(F+iG) \\ &\because (1-ir)^{-1} = \frac{1}{1-ir} \end{aligned}$$

$$\therefore (A+iB)^{-1} = (1-ir)(F+iG)^{-1} \tag{7.13}$$

Example 7.23 Find the inverse of the following complex matrix: $M = \begin{bmatrix} 5+i & 4+2i \\ 10+3i & 8+6i \end{bmatrix}$.

Solution:

$$\text{Let } M = A + iB = \begin{bmatrix} 5 & 4 \\ 10 & 8 \end{bmatrix} + i \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \tag{7.14}$$

Now, $|A| = 0, |B| = 0$. Therefore A, B are singular.

Let,

$$\begin{aligned} F &= A + rB \\ G &= B - rA \end{aligned}$$

where r is a real.

$$\therefore (A+iB)^{-1=} = (1-ir)(F+iG)^{-1} \tag{7.15}$$

Again,

$$F = \begin{bmatrix} 5 & 4 \\ 10 & 8 \end{bmatrix} + r\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 5+r & 4+2r \\ 10+3r & 8+6r \end{bmatrix}$$
$$G = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} - r\begin{bmatrix} 5 & 4 \\ 10 & 8 \end{bmatrix} = \begin{bmatrix} 1-5r & 2-4r \\ 3-10r & 6-8r \end{bmatrix}$$

Put $r = 1$ so that

$$F = \begin{bmatrix} 6 & 6 \\ 13 & 14 \end{bmatrix}$$
$$G = \begin{bmatrix} -4 & -2 \\ -7 & -2 \end{bmatrix}$$

Because, $|A| \neq 0, |B| \neq 0$. Therefore A, B are nonsingular.
Now, we are going to find the inverse of $F + iG$.
Let

$$(F+iG)^{-1} = X + iY$$

Assume that $(F + iG)^{-1}$ exists then,

$$X = (F + GF^{-1}G)^{-1}$$
$$Y = -F^{-1}G(F + GF^{-1}G)^{-1}$$

Now,

$$F^{-1} = \frac{adj.F}{|F|}\begin{bmatrix} \frac{7}{3} & -1 \\ \frac{-13}{6} & 1 \end{bmatrix}$$
$$\therefore F^{-1}G = \begin{bmatrix} \frac{7}{3} & -1 \\ \frac{-13}{6} & 1 \end{bmatrix}\begin{bmatrix} -4 & -2 \\ -7 & -2 \end{bmatrix} = \begin{bmatrix} \frac{-7}{3} & \frac{-8}{3} \\ \frac{5}{3} & \frac{7}{3} \end{bmatrix}$$

and

$$GF^{-1}G = \begin{bmatrix} -4 & -2 \\ -7 & -2 \end{bmatrix} \begin{bmatrix} \frac{-7}{3} & \frac{-8}{3} \\ \frac{5}{3} & \frac{7}{3} \end{bmatrix} = \begin{bmatrix} 6 & 6 \\ 13 & 14 \end{bmatrix}$$

$$F + GF^{-1}G = \begin{bmatrix} 6 & 6 \\ 13 & 14 \end{bmatrix} + \begin{bmatrix} 6 & 6 \\ 13 & 14 \end{bmatrix} = \begin{bmatrix} 12 & 12 \\ 26 & 28 \end{bmatrix}$$

$$\therefore (F + GF^{-1}G)^{-1} = \begin{bmatrix} 12 & 12 \\ 26 & 28 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{7}{6} & \frac{-1}{2} \\ \frac{-13}{12} & \frac{1}{2} \end{bmatrix}$$

$$= X$$

$$\because Y = -F^{-1}GX$$

$$= -\begin{bmatrix} \frac{-7}{3} & \frac{-8}{3} \\ \frac{5}{3} & \frac{7}{3} \end{bmatrix} \begin{bmatrix} \frac{7}{6} & \frac{-1}{2} \\ \frac{-13}{12} & \frac{1}{2} \end{bmatrix}$$

$$= -\begin{bmatrix} \frac{1}{6} & \frac{-1}{6} \\ \frac{-7}{12} & \frac{1}{3} \end{bmatrix}$$

$$\therefore X + iY = \begin{bmatrix} \frac{7}{6} & \frac{-1}{2} \\ \frac{-13}{12} & \frac{1}{2} \end{bmatrix} - i \begin{bmatrix} \frac{1}{6} & \frac{-1}{6} \\ \frac{-7}{12} & \frac{1}{3} \end{bmatrix}$$

$$= \frac{1}{6} \begin{bmatrix} 7 - i & -3 + i \\ -6.5 + 3.5i & 3 - 2i \end{bmatrix}$$

$$\therefore (A + iB)^{-1} = (1 - ir)(X + iB)$$

$$= \frac{1 - i}{6} \begin{bmatrix} 7 - i & -3 + i \\ -6.5 + 3.5i & 3 - 2i \end{bmatrix}$$

$$= \frac{1}{6} \begin{bmatrix} 6 - 8i & -2 + 4i \\ -3 + 10i & 1 - 5i \end{bmatrix}$$

Example 7.24 Find the inverse of the matrix

$$M = \begin{bmatrix} 3 + 3i & 1 + 4i \\ 4i & 2 - 3i \end{bmatrix}$$

Solution: Let $M = A + iB = \begin{bmatrix} 3 + 3i & 1 + 4i \\ 4i & 2 - 3i \end{bmatrix}$

where $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}, B = \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}$.

Obviously, $|A| = 6 \neq 0, |B| = -25 \neq 0$. Therefore A^{-1}, B^{-1} are exist.

Hence, $A^{-1} = \frac{1}{6} \begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix}, B^{-1} = \frac{-1}{25} \begin{bmatrix} -3 & -4 \\ -4 & 3 \end{bmatrix}$.

Now, let us $(A+iB)^{-1} = C + iD$ then,

$$C = (A + BA^{-1}B)^{-1} \tag{7.16}$$

$$D = -(AB^{-1}A + B)^{-1} \tag{7.17}$$

$$A^{-1}B = \frac{1}{6}\begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 2 & 11 \\ 12 & -9 \end{bmatrix}$$

$$BA^{-1}B = \frac{1}{6}\begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix}\begin{bmatrix} 2 & 11 \\ 12 & -9 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 54 & -3 \\ -28 & 71 \end{bmatrix}$$

$$A + BA^{-1}B = \frac{1}{6}\begin{bmatrix} 72 & 3 \\ -28 & 83 \end{bmatrix}$$

$$\therefore C = (A + BA^{-1}B)^{-1} = \frac{1}{1010}\begin{bmatrix} 83 & -3 \\ 28 & 72 \end{bmatrix}$$

Similarly,

$$D = \frac{1}{6060}\begin{bmatrix} -474 & -786 \\ -744 & 684 \end{bmatrix}$$

$$\therefore (A+iB)^{-1} = C + iD = \frac{1}{6060}\begin{bmatrix} 498 - 474i & -18 - 786i \\ 168 - 744i & 432 + 684i \end{bmatrix}$$

7.14 Exercises

Q1: Find the enverse of the following:

(i) $M = \begin{bmatrix} 1 & 0 \\ 1+i & -i \end{bmatrix}$

(ii) $M = \begin{bmatrix} 2-i & -i \\ -2i & 1+i \end{bmatrix}$

(iii) $M = \begin{bmatrix} 2i & 4 & 5+6i \\ 3+i & i & 2-4i \\ 2 & 1-i & 5+i \end{bmatrix}$

(iv) $M = \begin{bmatrix} 1+i & 2-i & 3+2i \\ 1 & i & 1+i \\ 2 & 1-i & 1+i \end{bmatrix}$

Q2: Prove that the inverse of product of two matrices is the product of the inverse taken in the reverse order. Or, if $M = A+iB, W = E+iF$ are invertible matrices, prove that $(MW)^{-1} = W^{-1}W^{-1}$.

8

Numerical Solution of A System of Linear Equations

8.1 Introduction

Linear systems are the basis and a fundamental part of linear algebra (Anton, 1987), a subject used in most modern mathematics. Computational algorithms for finding the solutions are an important part of numerical linear algebra (Hartmanis and Stearns, 1965; Gill et al., 2021), and play a prominent role in engineering, physics, chemistry, computer science, and economics (Callier and Desoer, 2012). A system of non-linear equations can often be approximated by a linear system (Morozov et al., 2007; Pshenichnyj, 1987), it is a technique when making a mathematical model or computer simulation of a relatively complex system. There are methods to solve a system of linear equation such that; a solution to it is an assignment of values to the variables such that all the equations are simultaneously satisfied (Axelsson, 2007).

8.2 Fundamental of linear equations

8.2.1 Definition of linear equation

Definition 8.1 A system of linear equations is a collection of one or more linear equations involving the same variables (Anton, 1987).

Example 8.1

$$\begin{aligned}&x_1 + 4x_2 + 2x_3 + 3x_4 = 5\\&x_2 + 4x_3 + 4x_4 = 0\\&-x_1 + x_3 = -2\\&2x_1 + 4x_3 + x_4 = 3\end{aligned}$$

8.2.2 Mathematical formulation of linear system

A general system of m linear equations with n variables and coefficients can be written as:

$$\begin{aligned}&a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1\\&a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2\\&.\\&.\\&.\\&a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m\end{aligned} \tag{8.1}$$

where $x_1, x_2, ..., x_n$ are the variables (unknowns), $a_{11}, a_{12}, ..., a_{mn}$ are the coefficients of the system, and $b_1, b_2, ..., b_m$ are the constant terms (Beauregard and Fraleigh, 1973).

It can be expressed of (8.1) in the matrices form as in (8.2) bellow:

$$AX = B \tag{8.2}$$

where, $A = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ . & . & ... & . \\ . & . & ... & . \\ . & . & ... & . \\ a_{m1} & a_{m2} & ... & a_{mn} \end{bmatrix}$ is a matrix coefficients, $X =$

$\begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$ is a column of variables, and $B = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_m \end{bmatrix}$ is a column of quantities.

Also, the matrix $M = \left[\begin{array}{cccc:c} a_{11} & a_{12} & ... & a_{1n} & b_1 \\ a_{21} & a_{22} & ... & a_{2n} & b_2 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{m1} & a_{m2} & ... & a_{mn} & b_m \end{array}\right]$ is called Augmented matrix (Marcus and Minc, 1992). $\left[\begin{array}{c:c} A & B \end{array}\right]$. This matrix has an important role in our field of the study.

The solution to (8.1) is to find the values of $x_1, x_2, ...x_n$ so that all equations in the system are fulfilled, and these depend on the relation between m, n so as depend on the value of the column B.

8.2.3 Basic definitions

Definition 8.2 An augmented matrix $\left[\begin{array}{c:c} A & B \end{array}\right]$ is a $k \times (n+1)$ matrix obtained by appending a $k-$ dimensional row vector B, on the right, as a further column to a $k \times n-$ dimensional matrix A (Marcus and Minc, 1992).

Definition 8.3 The system in (8.1) is called homogeneous if all elements in B equal to zero. Or, $b_1 = b_2 = ... = b_m = 0$ (Anton, 1987).

Definition 8.4 The solution of a system is called trivial (zero) solution if all values of X are zero. Or, $x_1 = x_2 = ... = x_n = 0$ (Anton, 1987; Ralston and Rabinowitz, 2001)

Definition 8.5 A matrix whose number of columns and rows are equal to$n$ is called of degree n (Strang, 2006).

Theorem 8.1 *If W is a solution to the homogeneous system $AX = 0$, and Y is a solution to (8.1) then any other solution C to (8.1) will be as the form of $C = W + Y$.*

Proof

$$\because Y, C \text{ are solutions to (8.1)}$$
$$\therefore AC = B, AY = B$$
$$\therefore AW = A(C - Y)$$
$$= AC - AY$$
$$= B - B$$
$$= 0. \blacklozenge$$

Corollary *The system (8.1) has a unique solution if and only if the trivial solution is a unique solution to the homogeneous system.*

Definition 8.6 A system of equations is called consistent if there is at least one set of values for the variables that satisfies each equation in the system (Goult, 1974; Hoffman and Kunze, 1967; Hohn, 1972).

Definition 8.7 A system of equations is called inconsistent if there is no set of values for the variables that satisfies all of the equations (Goult, 1974; Hoffman and Kunze, 1967; Hohn, 1972).

Theorem 8.2 *A system of m equations and n variables has always a nontrivial solution if $m \leq n$.*

Proof Suppose the system is

$$AX = 0 \tag{8.3}$$

Now, we have to prove that it is possible to find a solution to $W \neq 0$ where

where A is $m \times n$ matrix.

$$AW = 0$$

We are going to prove the solution by mathematical induction as follows:

- If $n = 2$, we have the equation;

$$a_{11}x_1 + a_{12}x_2 = 0$$

where can not be $a_{11} = a_{12} = 0$ at the same time.

If one of them is zero, let us consider $a_{12} \neq 0$. The solution will be $x_1 = 0, x_2 = \lambda$, where λ is a constant, and put $\lambda = 1$.

If $a_{11} \neq 0$ the solution will be as:

$$x_1 = -\frac{a_{12}}{a_{11}}k, x_2 = \lambda$$

Thus, there is the intrivial solution, where $n = 2$.

- Suppose that $n > 2$. The theorem stay true for the system its number of equations less the variables in which there are less than n variables.

  Thus, (8.3) has intrivial solution if all elements of the column n in the matrix A are equal to zero. Or,

$$A_{in} = 0, i = 1, 2, ..., m$$
$$\therefore \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ . \\ . \\ . \\ 1 \end{bmatrix}$$

  But, if some elements column n are not equal to zero, suppose $a_{pn} \neq 0$.

  In this case, we find a matrix $B_{m\times(n-1)}$ such that:

$$b_{ij} = a_{ij} - \frac{a_{pj}}{a_{pn}}a_{in}, \quad j = 1, 2, ..., m; j = 1, 2, ..., n-1$$

  Or, all elements of row p in B are equal zero. In other words the equation p in the system:

$$BX = 0$$
$$0x_1 + 0x_2 + ... + 0x_{n-1} = 0$$

Thus any selection of the values for $x_1, x_2, ..., x_{n-1}$ satisfy the equation. Thereby, W is a solution for;

$$\begin{aligned} &BX = 0 \Leftrightarrow W \ \text{ is a solution for the system} \\ &B'X = 0 \end{aligned}$$

where B' is a matrix B deleted of it the row p. Or, the system;

$$B'X = 0$$

is consisted of $(m-1), (n-1)$ equations and variables respectively.

Thus, there is nontrivial solution for the system its equations less than its n variables.

Thus, by mathematical induction we have found a nontrivial solution for the system;

$$B'X = 0$$

Therefore, the theorem emphasized that can found nontrivial solution for the system;

$BX = 0$

$\exists AX = 0$

$\therefore$ from the definition of the elements B we have

$$x_1 a_{i1} + x_2 a_{i2} + ... + x_{n-1} a_{i(n-1)} + (\sum_{j=1}^{n-1} x_j \frac{a_{ij}}{a_{in}}) a_{in} = 0, i = 1, 2, ..., m$$

It means we we have found a nontrivial solution for the system: $AX = 0$.♦

Theorem 8.3 *If the square matrix A in the degree n then:*

(i) *The homogeneous system $AX = 0$ has the trivial solution.*

(ii) *The system $AX = B$ has the unique solution, for each different value of column B.*

(iii) *The matrix A is invertible.*

8.3 Solutions for systems with equal equations and variables

In this chapter, we will try to solve the system of equations (8.1) when the number of equations is equal to the number of variables. There are three main methods:

8.3.1 Cramer's Method

This method depends directly and completely on the determinant of the matrix and is called Cramer's method (Cramer, 1750; Kosinski, 2001), as it is based on the following theorem (Gong et al., 2002):

Theorem 8.4 *considers the matrix equation $AX = B$, where the $n \times n$ matrix A has a nonzero determinant, and X, B are $n \times m$ matrices. Given sequences; $1 \leq i_1 < i_2 < ... < i_k \leq n$, and $1 \leq j_1 < j_2 < ... < j_k \leq m$, let $X_{I,J}$ be the $k \times k$ submatrix of X with rows in $I = (i_1, ..., i_k)$ and and columns in $J = (j_1, ..., j_k)$. Let $A_B(I, J)$ be the $n \times n$ matrix formed by replacing the i_s rows of A by the j_s columns of B, for all $s = 1, ..., k$. Then:*
$X_{I,J} = \frac{det(A_B(I,J))}{det(A)}$.

Proof The proof for Cramer's rule depends of the properties of the determinants; in which a linearity with respect to any given column the determinant is zero whenever two columns are equal.

Fix the index j of a column, and consider that the entries of the other columns have fixed values. This makes the determinant a function of the entries of the jth column. Linearity with respect of this column means that this function has the form:

$$D_j(a_{1,j}, ..., a_{n,j}) = C_{1,j}a_{1,j} + ... + C_{n,j}a_{n,j}$$

where the $C_{i,j}$ are coefficients that depend on the entries of A that are not in column j. So, one has

$$det(A) = D_j(a_{1,j}, ..., a_{n,j}) = C_{1,j}a_{1,j} + ... + C_{n,j}a_{n,j}$$

If the function D_j is applied to any other column k of A, then the result is the determinant of the matrix obtained from A by replacing column j by a copy of column k, so the resulting determinant is 0 (the case of two equal columns).

Now consider a system of n linear equations in n variables $x_1, ..., x_n$, whose coefficient matrix is A, with $det(A)$ assumed to be nonzero:

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\
&a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\
&. \\
&. \\
&. \\
&a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n = b_n
\end{aligned}$$

If one combines these equations by taking $C_{1,j}$ times the first equation, plus $C_{2,j}$ times the second, and so forth until $C_{n,j}$ times the last, then for every k the resulting coefficient of x_k becomes:

$$D_j(a_{1,k}, ..., a_{n,k}).$$

So, all coefficients become zero, except the coefficient of x_j that becomes $det(A)$. Similarly, the constant coefficient becomes $D_j(b_1, ..., b_n)$, and the resulting equation is thus;

$$det(A)x_j = D_j(a_{1,k}, ..., a_{n,k}).$$

which gives the value of x_j as;

$$x_j = \frac{1}{det(A)} D_j(a_{1,k}, ..., a_{n,k})$$

As, by construction, the numerator is the determinant of the matrix obtained from A by replacing column j by $\mathbf{b}$, we get the expression of Cramer's rule as a necessary condition for a solution.

It remains to prove that these values for the unknowns form a solution. Let M be the $n \times n$ matrix that has the coefficients of D_j as jth row, for $j = 1, ..., n$ (this is the adjugate matrix for A). Expressed in matrix terms, we have thus to prove that:

$$\mathbf{x} = \frac{1}{det(A)} M\mathbf{b}$$

And, is a solution that;

$$A(\frac{1}{det(A)} M)\mathbf{b} = \mathbf{b}$$

In which, it suffices to prove that;

$$A(\frac{1}{det(A)} M) = I_n$$

where I_n is the identity matrix.

The above properties of the functions D_j show that one has $MA = det(A)In$, and therefore,

$$(\frac{1}{det(A)} M)A = I_n$$

This completes the proof, since a left inverse of a square matrix is also a right-inverse (From invertible matrix theorem). ♦

Theorem 8.5 (An alternative formulation of Theorem 8.4) *If the determinant of the coefficient matrix in a system of n variables is not equal to zero, then the equations have a unique solution, which is:*

$x_i = \frac{D_i}{D}, \forall i = 1, 2, ..., n,$ *where* $D = |A|$, *and*

$$D_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} & \dots & a_{1n} \\ b_2 & a_{22} & a_{23} & \dots & a_{2n} \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ b_n & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix}, \quad D_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} & \dots & a_{1n} \\ a_{21} & b_2 & a_{23} & \dots & a_{2n} \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ a_{n1} & b_n & a_{n3} & \dots & a_{nn} \end{vmatrix}, \dots,$$

$$D_n = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & b_1 \\ a_{21} & a_{22} & a_{23} & \dots & b_2 \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ a_{n1} & a_{n2} & a_{n3} & \dots & b_n \end{vmatrix}.$$

Proof We are going to prove the theorem in case where $n = 3$, and the same method can be followed when $n \geq 4$.

Let us consider the following system:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}$$

To solve the system, the determinant is

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11}x_1 & a_{12}x_2 & a_{13}x_3 & a_{12} & a_{13} \\ a_{21}x_1 & a_{22}x_2 & a_{23}x_3 & a_{22} & a_{23} \\ a_{31}x_1 & a_{32}x_2 & a_{33}x_3 & a_{32} & a_{33} \end{vmatrix}$$

Based on the properties of the determinants, it can be rewritten as ta sum of three determinants as follows:

$$\begin{aligned} &\begin{vmatrix} a_{11}x_1 & a_{12} & a_{13} \\ a_{21}x_2 & a_{22} & a_{23} \\ a_{31}x_3 & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{12}x_2 & a_{12} & a_{13} \\ a_{22}x_2 & a_{22} & a_{23} \\ a_{32}x_2 & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{13}x_3 & a_{12} & a_{13} \\ a_{23}x_3 & a_{22} & a_{23} \\ a_{33}x_3 & a_{32} & a_{33} \end{vmatrix} \\ &= x_1 \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + x_2 \begin{vmatrix} a_{12} & a_{12} & a_{13} \\ a_{22} & a_{22} & a_{23} \\ a_{32} & a_{32} & a_{33} \end{vmatrix} + x_3 \begin{vmatrix} a_{13} & a_{12} & a_{13} \\ a_{23} & a_{22} & a_{23} \\ a_{33} & a_{32} & a_{33} \end{vmatrix} \end{aligned}$$

$$x_1 \left| A \right| = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}$$

Provided $|A| \neq 0$ implies that;

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{|A|} = \frac{D_1}{|A|}$$

By the same method $x_2 = \frac{D_2}{|A|}$, and $x_3 = \frac{D_3}{|A|}$. ♦

Example 8.2 Use Cramer's method to solve the system:

$$0.3x_1 - 0.2x_2 = 0.6$$
$$0.2x_1 + x_2 = 0.5$$

Solution:

$$A = \begin{bmatrix} 0.3 & -0.2 \\ 0.2 & 1 \end{bmatrix}, B = \begin{bmatrix} 0.6 \\ 0.5 \end{bmatrix}$$

$$x_1 = \frac{|D_1|}{|A|} = \frac{\begin{vmatrix} 0.6 & -0.2 \\ 0.5 & 1 \end{vmatrix}}{\begin{vmatrix} 0.3 & -0.2 \\ 0.2 & 1 \end{vmatrix}} = \frac{\frac{7}{10}}{\frac{17}{50}} = \frac{35}{17}$$

$$x_2 = \frac{|D_2|}{|A|} = \frac{\begin{vmatrix} 0.6 & -0.2 \\ 0.5 & 1 \end{vmatrix}}{\begin{vmatrix} 0.3 & 0.6 \\ 0.2 & 0.5 \end{vmatrix}} = \frac{\frac{3}{100}}{\frac{17}{50}} = \frac{3}{34}$$

$\therefore X = \{(x_1, x_2)\} = \left\{(\frac{35}{17}, \frac{3}{34})\right\}$ is a set solution of the system

.

Example 8.3 Use Cramer's method to solve the system:

$$0.5x_1 + 0.2x_2 + x_3 = 0.7$$
$$0.3x_1 - x_2 - 0.2x_3 = 0.9$$
$$0.4x_1 + 0.3x_2 - 0.3x_3 = 0.3$$

Solution:

$$A = \begin{bmatrix} 0.5 & 0.2 & 1 \\ 0.3 & -1 & -0.2 \\ 0.4 & 0.3 & -0.3 \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, B = \begin{bmatrix} 0.7 \\ 0.9 \\ 0.3 \end{bmatrix}$$

$$x_1 = \frac{|D_1|}{|A|} = \frac{\begin{vmatrix} 0.7 & 0.2 & 1 \\ 0.9 & -1 & -0.2 \\ 0.3 & 0.3 & -0.3 \end{vmatrix}}{\begin{vmatrix} 0.5 & 0.2 & 1 \\ 0.3 & -1 & -0.2 \\ 0.4 & 0.3 & -0.3 \end{vmatrix}} = \frac{\frac{108}{125}}{\frac{84}{125}} = \frac{9}{7}$$

$$x_2 = \frac{|D_2|}{|A|} = \frac{\begin{vmatrix} 0.5 & 0.7 & 1 \\ 0.3 & 0.9 & -0.2 \\ 0.4 & 0.3 & -0.3 \end{vmatrix}}{\begin{vmatrix} 0.5 & 0.2 & 1 \\ 0.3 & -1 & -0.2 \\ 0.4 & 0.3 & -0.3 \end{vmatrix}} = \frac{\frac{-46}{125}}{\frac{84}{125}} = \frac{-24}{42}$$

$$x_3 = \frac{|D_3|}{|A|} = \frac{\begin{vmatrix} 0.5 & 0.2 & 0.7 \\ 0.3 & -1 & 0.9 \\ 0.4 & 0.3 & 0.3 \end{vmatrix}}{\begin{vmatrix} 0.5 & 0.2 & 1 \\ 0.3 & -1 & -0.2 \\ 0.4 & 0.3 & -0.3 \end{vmatrix}} = \frac{\frac{14}{125}}{\frac{84}{125}} = \frac{1}{6}$$

$\therefore X = \{(x_1, x_2, x_3)\} = \left\{ (\frac{9}{7}, \frac{-24}{42}, \frac{1}{6}) \right\}$ is a set solution of the system.

8.3.2 Gauss's Method

The aim in this section is to describe how the solutions to a linear system are actually found. The essential idea is to add multiples of one equation to the others in order to eliminate a variable and to continue this process until only one variable is left. Once this final variable is determined, its value is substituted back into the other equations in order to evaluate the remaining variables.

Definition 8.8 Gaussian elimination is an algorithm for solving systems of linear equations. It consists of a sequence of row-wise operations performed on the corresponding matrix of coefficients. Or, the method, characterized by step-by-step elimination of the variables, is called Gaussian elimination (Grötschel, 2012; Calinger et al., 1999; Gowers et al., 2010).

8.3.3 Gauss's Method and row echelon form

To obtain row reduction in a matrix, a series of elementary row operations must be used to modify the matrix until the bottom-left corner of the matrix is filled with zeros, as much as possible. There are three types of elementary row operations:

(i) Swapping two rows.

(ii) Multiplying a row by a nonzero number.

(iii) Adding a multiple of one row to another row.

Now, let us describe the solution algorithm step by step. Consider the system of (8.1). After substituting the value of x_1 in the first equation in the system, in the second, third, ...etc. equations, we obtain of a new system of the form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n &= b_1 \\ a'_{22}x_2 + a'_{23}x_3 + ... + a'_{2n}x_n &= b'_2 \\ a'_{32}x_2 + a'_{33}x_3 + ... + a'_{3n}x_n &= b'_3 \\ . \\ . \\ . \\ a'_{n2}x_2 + a'_{n3}x_3 + ... + a'_{nn}x_n &= b'_n \end{aligned} \tag{8.4}$$

After that substituting the value of x_2 in the second equation in the system of (8.4) in the third, fourth, ...etc. equations, we obtain of a new system of the form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n &= b_1 \\ a'_{22}x_2 + a'_{23}x_3 + ... + a'_{2n}x_n &= b'_2 \\ a''_{33}x_3 + ... + a''_{3n}x_n &= b''_3 \\ . \\ . \\ . \\ a''_{n3}x_3 + ... + a''_{nn}x_n &= b''_n \end{aligned} \tag{8.5}$$

And so on until the original system converted into the following formula:

$$\begin{aligned} c_{11}x_1 + c_{12}x_2 + a_{13}x_3 + ... + c_{1n}x_n &= d_1 \\ c_{22}x_2 + c_{23}x_3 + ... + c_{2n}x_n &= d_2 \\ c_{33}x_3 + ... + a_{3n}x_n &= d_3 \\ . \\ . \\ . \\ c_{(n-1)(n-1)}x_{n-1} + c_{(n-1)n}x_n &= d_{n-1} \\ c_{nn}x_n &= d_n \end{aligned} \tag{8.6}$$

where

$$x_n = \frac{d_n}{c_{nn}}$$

$$c_{(n-1)(n-1)}x_{n-1} + c_{(n-1)n}x_n = d_{n-1}$$

Or, we obtain:

$$x_{n-1} = \frac{1}{c_{(n-1)(n-1)}}(d_{n-1} - c_{(n-1)n}\frac{d_n}{c_{nn}})$$

Thus, by substituting in the equations $1, 2, ..., n-3, n-2$ in the system (8.5) we obtain on the values of:

$$x, x_2, ..., x_{n-3}, x_{n-2}$$

And we obtain on the same system (8.5), if we write the system in the form of limited matrix:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & ... & a_{1n} & b_1 \\ a_{21} & a_{22} & ... & a_{2n} & b_2 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{n1} & a_{n2} & ... & a_{nn} & b_n \end{array}\right] \tag{8.7}$$

Now, we have to get rid of the items $a_{21}, a_{31}, ..., a_{n1}$ by multiplying items of the first row by the item $\frac{a_{21}}{a_n}$ and subtraction of the items of the items of the second row. Again, multiplying items of the first row by the item $\frac{a_{31}}{a_{11}}$ and subtraction of the items of the items of the third row and so on to obtain:

$$\left[\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & ... & a_{1n} & b_1 \\ 0 & a_{22} & a'_{23} & ... & a'_{2n} & b'_2 \\ 0 & a'_{32} & a'_{33} & ... & a'_{3n} & b'_3 \\ . & & & & & \\ . & & & & & \\ . & & & & & \\ 0 & a'_{n2} & a'_{n3} & ... & a'_{nn} & b'_n \end{array}\right]$$

After that, we get rid of the terms $a'_{32}, a'_{12}, ..., a'_{n2}$ by subtracting terns of second row by the term $\frac{a'_{32}}{a'_{22}}$ of the third row, and multiply of terms of second row by the term $\frac{a'_{42}}{a'_{22}}$ from the fourth row and so on till the nth row. until we get the following limited matrix:

$$\left[\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & ... & a_{1n} & b_1 \\ 0 & a'_{22} & a'_{23} & ... & a'_{2n} & b'_2 \\ 0 & 0 & a''_{33} & ... & a''_{3n} & b''_3 \\ . & & & & & \\ . & & & & & \\ . & & & & & \\ 0 & 0 & a''_{n3} & ... & a''_{nn} & b''_n \end{array}\right]$$

Repeating this processes on the remain rows, we obtain the limited matrix:

$$\left[\begin{array}{cccccc:c} c_{11} & c_{12} & c_{13} & ... & c_{1n} & d_1 \\ 0 & c_{22} & c_{23} & ... & c_{2n} & d_2 \\ 0 & 0 & c_{33} & ... & c_{3n} & d_3 \\ . & & & & & \\ . & & & & & \\ . & & & & & \\ 0 & 0 & 0 & ... & c_{nn} & d_n \end{array}\right]$$

where $x_n = \frac{d_n}{c_{nn}}$, and substituting this value in the equations $n_1, n_2, 1$ in the last system, in which it is equivalence to the system (8.1), we can find the remain variables.

Example 8.4 Solve the given set of equations by using Gauss elimination method:

$$x - y + z = 4$$
$$x - 4y + 2z = 8$$
$$x + 2y + 8z = 12$$

Solution: After converting the equations in matrix form, the system will be as follows:

$$\left[\begin{array}{ccc:c} 1 & -1 & 1 & 4 \\ 1 & -4 & 2 & 8 \\ 1 & 2 & 8 & 12 \end{array}\right]$$

By $(R_2 - R_1 \to R_2) \wedge (R_3 - R_1 \to R - 3)$, we get:

$$\left[\begin{array}{ccc:c} 1 & -1 & 1 & 4 \\ 0 & -3 & 1 & 4 \\ 0 & 3 & 7 & 7 \end{array}\right]$$

By $R_3 + R_2 \to R_3$, the system became to echelon form:

$$\left[\begin{array}{ccc:c} 1 & -1 & 1 & 4 \\ 0 & -3 & 1 & 4 \\ 0 & 0 & 8 & 12 \end{array}\right]$$

Thus,

$$
\begin{aligned}
x - y + z &= 4 \\
-\,3y + z &= 4 \\
8z &= 12
\end{aligned}
$$

From the third equation $z = \frac{3}{2}$, substituting the value of z in the second equation implies $y = \frac{-5}{6}$, and substituting the values of y, z in the first equation, the value of $x = \frac{5}{3}$. Thus, the set solution of the system is $\{(x, y, z)\} = \left\{(\frac{5}{3}, \frac{-5}{6}, \frac{3}{2})\right\}$.

Example 8.5 Consider the following linear system and solve it by Gaussian eliminations:

$$
\begin{aligned}
&2x_1 + x_2 - x_3 + 2x_4 = 5 \\
&4x_1 + 5x_2 - 3x_3 + 6x_4 = 9 \\
&-2x_1 + 5x_2 - 2x_3 + 6x_4 = 4 \\
&4x_1 + 11x_2 - 4x_3 + 8x_4 = 2
\end{aligned}
$$

Solution:

$$AX = B$$

$$\left[\begin{array}{cccc:c} 2 & 1 & -1 & 2 & 5 \\ 4 & 5 & -3 & 6 & 9 \\ -2 & 5 & -2 & 6 & 4 \\ 4 & 11 & -4 & 8 & 2 \end{array}\right] \xrightarrow[\sim]{h_{21}(-2),h_{31}(1),h_{41}(2)}$$

$$\left[\begin{array}{cccc:c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 6 & -3 & 8 & 9 \\ 0 & 9 & -2 & 4 & -8 \end{array}\right] \xrightarrow[\sim]{h_{32}(-2),h_{42}(-3)}$$

$$\left[\begin{array}{cccc:c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 0 & -1 & 4 & 11 \\ 0 & 0 & 1 & -2 & -5 \end{array}\right] \xrightarrow[\sim]{h_{43}(1)}$$

$$\left[\begin{array}{cccc:c} 2 & 1 & -1 & 2 & 5 \\ 0 & 3 & -1 & 2 & -1 \\ 0 & 0 & -1 & 4 & 11 \\ 0 & 0 & 0 & 2 & 6 \end{array}\right] \equiv$$

$$\begin{aligned} 2x_1 + x_2 - x_3 + 2x_4 &= 5 \\ 3x_2 - x_3 + 2x_4 &= -1 \\ - x_3 + 4x_4 &= 11 \\ 2x_4 &= 6 \end{aligned}$$

Solving by back substitution, we obtain, the set solution $X = \{(x_1, x_2, x_3, x_4)\} = \{(1, -2, 1, 3)\}$.

Note: It is likely to be $a_{11} = 0$. To avoid this situation, we choose the equation in which a_{i1} is as large as possible in which we need to divide by a_{i1}. This operation in the Gaussian method with replacing between rows is called partial pivoting.

Example 8.6 Use Gaussian elimination method to solve the following linear system:

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 0 \\ 3x_1 - x_2 - 2x_3 &= 9 \\ 4x_1 + 3x_2 - 3x_3 &= 2 \end{aligned}$$

Solution: After replacing between the first and third grows, we have:

$$\left[\begin{array}{ccc:c} 4 & 3 & -3 & 3 \\ 1 & 2 & 1 & 0 \\ 3 & -1 & -2 & 9 \end{array}\right]$$

And utilizing Gaussian eliminations, the obtained solution will be:
$X = \{(x_1, x_2, x_3)\} = \{(3, -2, 1)\}$.

8.3.4 Coefficient matrix partition (LU) method

To solve the system:

$$AX = B \tag{8.8}$$

We partition the coefficient matrix A into an upper triangular matrix U, and a lower triangular matrix L in which $A = LU$ and $L_{ii} = 1, i = 1, 2, ..., n$ (Lang, 1984; Lang, 2002; Fraleigh, 2003; Nering, 1970). When substituting in A into the system (8.8), we get on the equivalent system:

$$LUX = B \tag{8.9}$$

By putting $UX = Y$ in (8.9), the resulted equation will be:

$$LY = B \tag{8.10}$$

By solving system (8.10), we find the value of Y, and after finding it, we can solve the system:

$$UX = Y \tag{8.11}$$

It is equivalent to solving the system (8.8).

Example 8.7 Solve the following system by coefficient matrix partition method:

$$
\begin{aligned}
2x_1 - 2x_2 + 3x_3 &= 3 \\
-2x_1 + 5x_2 + x_3 &= 5 \\
2x_1 - 8x_2 + 2x_3 &= 7
\end{aligned}
$$

Solution:

$$\begin{bmatrix} 2 & -2 & 3 \\ -2 & 5 & 1 \\ 2 & -8 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}$$

$$\therefore A = \begin{bmatrix} 2 & -2 & 3 \\ -2 & 5 & 1 \\ 2 & -8 & 2 \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, B = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}.$$

After partition of A to L, U lower and upper triangular matrix respectively, where $L = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix}, U = \begin{bmatrix} 2 & -2 & 3 \\ 0 & 3 & 4 \\ 0 & 0 & 7 \end{bmatrix}$, we solve the system:

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}$$

$$y_1 = 3, y_2 = 8, y_3 = 20$$

$$\therefore Y = \begin{bmatrix} 3 \\ 8 \\ 20 \end{bmatrix}.$$

Now, we have to solve the following System:

$$\begin{bmatrix} 2 & -2 & 3 \\ 0 & 3 & 4 \\ 0 & 0 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 8 \\ 20 \end{bmatrix}$$

$$x_3 = \frac{20}{7}, x_2 = \frac{-8}{7}, x_1 = \frac{-55}{14}$$

$$\therefore X = \begin{bmatrix} \frac{-55}{14} \\ \frac{-8}{7} \\ \frac{20}{7} \end{bmatrix}.$$

Thus, the set solution to the original problem is $X = \left\{\left(\frac{-55}{14}, \frac{-8}{7}, \frac{20}{7}\right)\right\}$.

Note: The two methods Gauss and LU are equivalent.

8.3.5 Matrix Inverse Method

Solving a system of linear equations using the inverse of a matrix requires the definition of two new matrices: X is the matrix representing the variables of the system, and B is the matrix representing the constants. Using matrix multiplication, we may define a system of equations with the same number of equations as variables as:

$$AX = B \tag{8.12}$$

To solve the system (8.12) of linear equations using an inverse matrix, let A be the coefficient matrix, let X be the variable matrix, and let B be the constant matrix.

$$\because AX = B$$
$$\therefore X = A^{-1}B$$

That is, if we find the matrix that is the inverse of matrix A by one of the methods explained in Chapter Seven, then the product of multiplying the inverse by column B gives a column with the values of the variables.

Example 8.8 Solve the following system by using inverse of matrix:

$$\begin{bmatrix} 1 & 3 & 2 & 1 \\ 2 & 1 & 7 & 3 \\ -1 & 2 & -3 & 4 \\ 2 & -3 & -3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ -4 \\ 7 \end{bmatrix}.$$

Solution: Using one of the legitimate methods in Chapter Seven, we find that:

$$\begin{bmatrix} 1 & 3 & 2 & 1 \\ 2 & 1 & 7 & 3 \\ -1 & 2 & -3 & 4 \\ 2 & -3 & -3 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{68}{89} & \frac{-52}{89} & \frac{-27}{89} & \frac{49}{89} \\ \frac{45}{178} & \frac{-3}{178} & \frac{7}{178} & \frac{-8}{89} \\ \frac{-19}{89} & \frac{25}{89} & \frac{1}{89} & \frac{-15}{89} \\ \frac{-17}{178} & \frac{13}{178} & \frac{29}{178} & \frac{5}{89} \end{bmatrix},$$

$$X = \begin{bmatrix} \frac{68}{89} & \frac{-52}{89} & \frac{-27}{89} & \frac{49}{89} \\ \frac{45}{178} & \frac{-3}{178} & \frac{7}{178} & \frac{-8}{89} \\ \frac{-19}{89} & \frac{25}{89} & \frac{1}{89} & \frac{-15}{89} \\ \frac{-17}{178} & \frac{13}{178} & \frac{29}{178} & \frac{5}{89} \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ -4 \\ 7 \end{bmatrix} = \begin{bmatrix} \frac{395}{89} \\ \frac{-10}{89} \\ \frac{-41}{89} \\ \frac{-16}{89} \end{bmatrix}.$$

That is, $X = \{(x_1, x_2, x_3, x_4)\} = \left\{\left(\frac{395}{89}, \frac{-10}{89}, \frac{-41}{89}, \frac{-16}{89}\right)\right\}$.

8.4 The methods and their operations on the computer

In this chapter, we explained several methods for solving linear equations so that we can choose the optimal method when solving systems of linear equations.

Each method has its advantages because some methods are easier and faster than others, and we noticed that:

(i) The solution to finding the inverse is the product of multiplying the inverse by the column on the right side of the system.

(ii) The solution of using the Gaussian method is to multiply rows by fixed quantities, then subtract them from another row, ..., and so on.

(iii) These methods are designed primarily for use on computers.

Table 8.1: The methods and their operations on the computer

The method	Number of multiplication operations
Cramer	$\frac{n^3(n+1)}{3}$
Gauss	$\frac{n^3}{3}$
LU	$\frac{n^3}{3}$
Matrix inverse	n^3

(iv) It is very important to take into account the number of algebraic operations that are actually used to find solutions by each method.

(v) The most time-consuming operations are multiplication and division.

(vi) Comparing these methods, we find that the approximate number of multiplication operations used in each method for a system of degree n is as shown in Table 8.1.

8.5 Exercises

Q1: Solve the following systems by Cramer's method:

(i)

$$\begin{aligned} 2x_1 + 3x_2 &= 0 \\ 3x_1 + 4x_2 &= 1 \end{aligned}$$

(ii)

$$\begin{aligned} 8x_1 - 4x_2 + x_3 &= 8 \\ 3x_1 + x_2 - x_3 &= 0 \\ 2x_1 + 7x_2 - 4x_3 &= 0 \end{aligned}$$

(iii)

$$\begin{aligned} 4x_1 - 6x_2 + 8x_3 + 2x_4 &= 10 \\ -2x_1 + 3x_2 - 4x_3 + 4x_4 &= 12 \\ 12x_1 + 18x_2 - 24x_3 + 12x_4 &= 11 \\ 7x_1 - 5x_2 + 7x_3 + 9x_4 &= 9 \end{aligned}$$

Q2: Solve each of the following equations in two different methods:

(i)

$$\begin{aligned} x_1 + 2x_2 - 4x_3 &= -1 \\ 7x_1 + 3x_2 + 5x_3 &= 26 \\ -2x_1 - 6x_2 + 7x_3 &= -6 \end{aligned}$$

(ii)

$$\begin{aligned} 4x_1 + 5x_2 - x_3 &= -7 \\ -2x_1 - 9x_2 + 2x_3 &= 8 \\ 5x_2 + 7x_3 &= 21 \end{aligned}$$

Q3: Solve the following equations in three different methods:

$$\begin{aligned} 14x_1 + 3x_2 + 7x_3 + 3x_4 &= 7 \\ 13x_1 + x_2 + 3x_3 + 5x_4 &= 3 \\ 18x_1 + 15x_2 + 12x_3 + 10x_4 &= 12 \\ 15x_1 - 12x_2 + 2x_3 - 10x_4 &= 4 \end{aligned}$$

Q4: Solve the following equations by LU method:

$$\begin{aligned} 2x_1 - 3x_2 + x_3 &= 5 \\ 2x_1 - x_2 + 4x_3 &= 3 \\ x_1 + 4x_2 + 2x_3 &= 6 \end{aligned}$$

Q5: Solve the following equations:

$$\begin{aligned} a_1x_1 + a_2x_2 + a_3x_3 &= b_1 \\ a_4x_1 + a_5x_2 + a_6x_3 &= b_2 \\ a_7x_1 + a_8x_2 + a_9x_3 &= b_3 \end{aligned}$$

where $a_i, b_j \in \mathbb{R}, i = 1, ..., 9; j = 1, 2, 3.$

9

Eigenvalues and Eigenvectors

9.1 Introduction

Eigenvalues are associated with eigenvectors in linear algebra. Both terms are used in the analysis of linear transformations. Eigenvalues are the special set of scalar values that is associated with the set of linear equations especially in the matrix equations. The eigenvectors are also termed as characteristic roots. It is a non-zero vector that can be changed at most by its scalar factor after the application of linear transformations. The corresponding factor which scales the eigenvectors is called an eigenvalue.

The mathematical interpretation is that eigenvalues are values that determine how separate a square matrix is from the basis vectors. When solving the eigenvalue equation for a square matrix A , the eigenvalues are found using the equation $|A - \lambda I| = 0$, where I is the identity matrix. These values are the eigenvalues of matrix A. Whenever, we obtain eigenvalues of matrix A, it can calculate eigenvectors of each eigenvalue. Eigenvectors are vectors that determinate the directions in which the matrix A propagates when multiplied by the eigenvector corresponding to the corresponding eigenvalue.

9.2 Eigenvalues and eigenvectors feature

In mathematics, especially in linear algebra, linear transformations specify the direction of the vector, it is important to know which vectors have their directions unchanged by a given linear transformation. An eigenvector is such a vector. Precisely, an eigenvector v of a linear transformation T is scaled by a constant factor λ when the linear transformation is applied to it: $Tv = \lambda v$.

The corresponding eigenvalue, or characteristic root is the multiplying factor λ. Geometrically, vectors are multi-dimensional quantities with magnitude and direction. A linear transformation rotates, stretches, or shears the vectors upon which it acts. Its eigenvectors are those vectors that are only stretched, with no rotation or shear. The corresponding eigenvalue is the factor by which an eigenvector is stretched or squished. If the eigenvalue is negative, the eigenvector's direction is reversed (Faires and Burden, 2012; Hamadameen, 2022).

9.3 Eigenvalue and eigenvector through changes direction

Let us consider Figure 9.2. In the mapping, the blue arrow changes direction, whereas the pink arrow does not. Here, the pink arrow is an eigenvector because it does not change direction. Also, the length of this arrow is not changed; its eigenvalue is 1.

9.4 Transformations and eigenvalues

Let us consider the following equation:

$$Y = AX \tag{9.1}$$

where the square matrix $A_{n\times n} \in \mathbb{F}$, $X = [x_1, x_2, ..., x_n]^T, Y = [y_1, y_2, ..., y_n]^T \in \mathbb{F}$, and $\mathbb{F}$ is a real number field. The matrix A would transform the column vector X into the column vector Y. This transformation can take many different forms, the simplest of which is

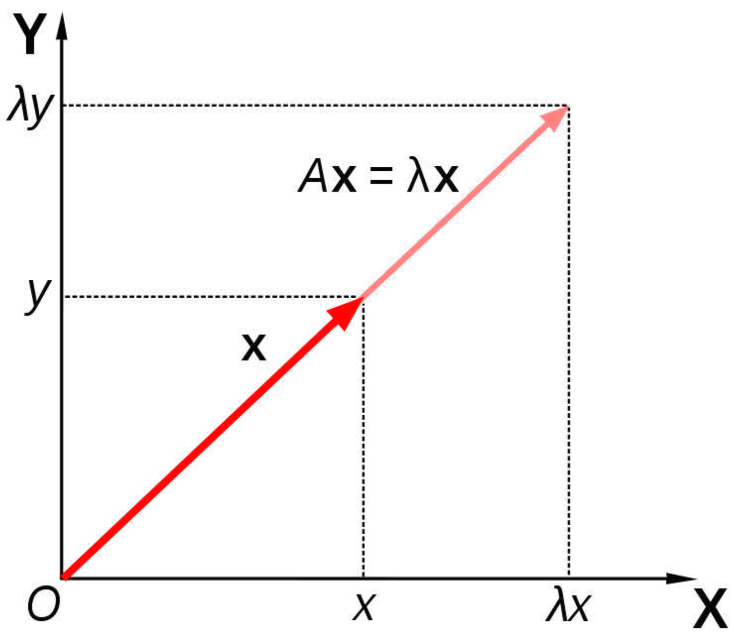


Figure 9.1: Eigenvalues and eigenvectors feature

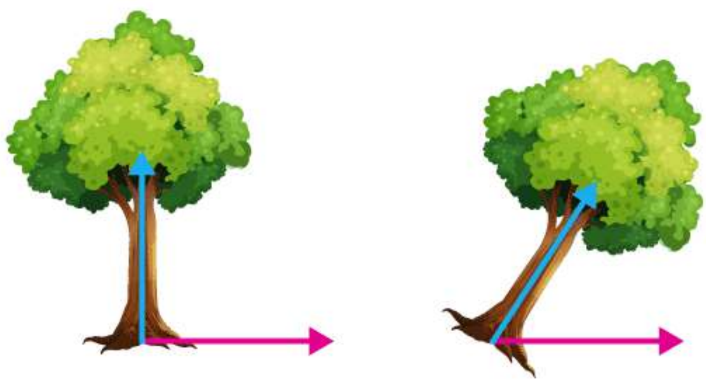


Figure 9.2: Eigenvalues and eigenvectors through changes direction

the linear transformation, which transforms X into itself by a number say λ of the field $\mathbb{F}$, which can be expressed as follows:

$$AX = \lambda X \tag{9.2}$$

Definition 9.1 The value λ in which satisfies the equation:

$$T(\lambda) = |A - \lambda I| = 0 \tag{9.3}$$

it is called the eigenvalue of the matrix A (Richard, 1993; Roman et al., 2005).

Note: The equation (9.3) is called the characteristic of eigenvalue equation and it can be simplified as follows in the form of a polynomial:

$$T(\lambda) = \sum_{i=0}^{i=n} a_i \lambda^i \tag{9.4}$$

where a_i are constants $\forall i$.

Definition 9.2 Based on some studies (Herstein, 1991; Herstein, 1964; Nering, 1970; Wu, 2005), Eigenvectors can defined as following: The vector X is defined as a corresponding eigenvector to the eigenvalue λ, and satisfied the equation:

$$AX = \lambda X \equiv (A - \lambda I)X = 0 \tag{9.5}$$

9.5 Polynomial equation of degree n in eigenvalue

Equation (9.5) can be represents as a set of homogeneous linear equations as follows:

$$\begin{aligned}
&(a_{11} - \lambda)x_1 + a_{12}x_2 + ... + a_{1n}x_n = 0 \\
&a_{21}x_1 + (a_{22} - \lambda)x_2 + ... + a_{2n}x_n = 0 \\
&. \\
&. \\
&. \\
&a_{n1}x_1 + a_{n2}x_2 + ... + (a_{nn} - \lambda)x_n = 0
\end{aligned} \tag{9.6}$$

A nontrivial solution to system (9.6) exists if;

$$|A - \lambda I| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ . & & & \\ . & & & \\ . & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0 \tag{9.7}$$

The system (9.7) is a polynomial equation of degree n in λ (Sarkar, 2008).

9.6 Applications of eigenvalues and eigenvectors

Applications of eigenvalues and eigenvectors enter almost all fields of science and have a pivotal and effective role for practical and applied conclusions. Most modern sciences use them to reach tangible practical results. Below is the list of fields in which the applications of them are carried out.

(i) Geometric transformations. Eigenvalues of geometric transformations includes; scaling, unequal scaling, rotation, horizontal shear, and hyperbolic rotation in each of; matrix, characteristic polynomial, eigenvalues, algebraic mult, geometric mult, and eigenvectors.

(ii) Principal component analysis.

(iii) Graphs.

(iv) Markov chains.

(v) Vibration analysis.

(vi) Tensor of moment of inertia.

(vii) Stress tensor.

(viii) Schrödinger equation.

(ix) Wave transport.

(x) Molecular orbitals.

(xi) Geology and glaciology.

(xii) Basic reproduction number.

(xiii) Eigenfaces.

(xiv) Eigenvoices.

(xv) Antieigenvalue theory.

(xvi) Eigenoperator.

(xvii) Eigenplane.

(xviii) Eigenmoments.

(xix) Eigenvalue algorithm.

(xx) Quantum states.

(xxi) Jordan normal form.

(xxii) List of numerical-analysis software.

(xxiii) Nonlinear eigenproblem.

(xxiv) Normal eigenvalue.

(xxv) Quadratic eigenvalue problem.

(xxvi) Singular value.

(xxvii) Spectrum of a matrix

For more information, the reader can review (Trefethen and Bau, 2022; Vellekoop and Mosk, 2007; Rotter, 2017; Bender et al., 2020; Graham and Midgley, 2000; Sneed and Folk, 1958; Knox-Robinson and Gardoll, 1998; Evans and Benn, 2014; Diekmann et al., 1990; Diekmann and Heesterbeek, 2000; Xirouhakis et al., 1999).

9.7 Properties of eigenvalues

Eigenvalues have the following properties (Hartman, 2011):

(i) Eigenvectors with distinct eigenvalues are linearly independent.

(ii) Singular matrices have zero eigenvalues.

(iii) If A is a square matrix, then $\lambda = 0$ is not an eigenvalue of A.

(iv) For a scalar multiple of a matrix: If A is a square matrix and λ is an eigenvalue of A, then, $a\lambda$ is an eigenvalue of aA.

(v) For matrix powers: If A is square matrix and λ is an eigenvalue of A, and $n \geq 0$ is an integer, then λ^n is an eigenvalue of A^n.

(vi) For polynomials of matrix: If A is a square matrix, λ is an eigenvalue of A, and $p(x)$ is a polynomial in variable x, then $p(\lambda)$ is the eigenvalue of matrix $p(A)$.

(vii) Inverse matrix: If A is a square matrix, and λ is an eigenvalue of A, then λ^{-1} is an eigenvalue of A^{-1}.

(viii) Transpose matrix: If A is a square matrix, and λ is an eigenvalue of A, then λ is an eigenvalue of A^T.

9.8 Eigenvalues and eigenvectors of matrices

To illustrate and explore how the eigenvalues and eigenvectors of a matrix relate to other properties of that matrix. The goal here is to be amazed at the many connections between mathematical concepts (Hartman, 2011). The aim here is to show that the linear transformations over a finite-dimensional vector space can be represented using matrices, which is especially common in numerical and computational applications (Jones, 2011; Herstein, 1991; Herstein, 1964; Nering, 1970; Press et al., 2007).

9.9 Definitions of various characteristics

In this section, we will list some definitions of the various characteristics in which they are important and necessary for the subject of our study in this chapter, based on some reliable references (Sarkar, 2008; Hartman, 2011; Jones, 2011; Herstein, 1991; Herstein, 1964; Nering, 1970; Press et al., 2007).

Definition 9.3 The matrix $A - \lambda I$ is called characteristic matrix of given matrix A which is obtained by subtracting λ from diagonal elements of A.

Definition 9.4 The $|A - \lambda I|$ when expanded will give a polynomial of degree n in λ which is called characteristic polynomial of matrix A.

Definition 9.5 The equation $|A - \lambda I| = 0$ is called characteristic equation of matrix A.

Definition 9.6 The roots $\lambda_1, \lambda_2, ..., \lambda_n$ of the characteristic equation are called characteristic roots or eigenvalues.

Definition 9.7 Corresponding to each characteristic root λ there corresponds nonzero vector X which satisfies the equation:
$(A - \lambda I)X = 0$. The X are characteristic vectors or eigenvectors.

9.10 Eigenvalues and eigenvectors of matrices

This section deals with how finds an eigenvalues and eigenvectors of matrices.This section deals with how finds an eigenvalues and eigenvectors of matrices. We try to submit three examples in squar matrices of degree $2 \times 2, 3 \times 3$, and 4×4. And we left matrices of degree $n \times n, n \geq 5$ as an exercises to the reader.

Example 9.1 Find the eigenvalues and eigenvectors of the following matrix: $\begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}$.

Solution: When solve the equation:

$$\begin{bmatrix} 3-\lambda & 2 \\ 1 & 2-\lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$(3-\lambda)(2-\lambda) - 3 = 0$$

$$\Rightarrow \lambda^2 - 5\lambda + 4 = 0$$

$$\Rightarrow \lambda = 1, \lambda = 4.$$

$$\therefore \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

$$If \lambda = 1 \Rightarrow \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow k \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

By the same way if $\lambda = 4$, we get $k \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Thus $X = k \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$, where k is constant.

Example 9.2 Find the eigenvalues of the following matrix: $\begin{bmatrix} 4 & 6 & 10 \\ 3 & 10 & 13 \\ -2 & -6 & -8 \end{bmatrix}$.

Solution: When solve the equation:

$$\begin{bmatrix} 4-\lambda & 6 & 10 \\ 3 & 10-\lambda & 13 \\ -2 & -6 & -8-\lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

To find eigenvalues $\lambda_i, i = 1, 2, 3$, we know that λ_i are the roots of $|A - \lambda I|$.

$$\therefore |A - \lambda I| = \left| \begin{bmatrix} 6 & 6 & 10 \\ 3 & 10 & 13 \\ -2 & -6 & -8 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right|$$

$$= \left| \begin{bmatrix} 6 & 6 & 10 \\ 3 & 10 & 13 \\ -2 & -6 & -8 \end{bmatrix} - \lambda \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \right| = \begin{vmatrix} 4-\lambda & 6 & 10 \\ 3 & 10-\lambda & 13 \\ -2 & -6 & -8-\lambda \end{vmatrix}$$

$$= (4-\lambda) \begin{vmatrix} 10-\lambda & 13 \\ -6 & -8-\lambda \end{vmatrix} - (6) \begin{vmatrix} 3 & 13 \\ -2 & -8-\lambda \end{vmatrix} + 10 \begin{vmatrix} 3 & 10-\lambda \\ -2 & -6 \end{vmatrix}$$

$$= \underbrace{(4-\lambda)[(10-\lambda)(-8-\lambda) - 13(-6)]}_{1} + \underbrace{(-6)[(3)(-8-\lambda) - 13(-2)]}_{2}$$

$$+ \underbrace{10[(3)(-6) - (10-\lambda)(-2)]}_{3} = 0...(1)$$

Now, from (1), the first term:

$$= -\lambda^3 + 6\lambda^2 - 6\lambda - 8...(1a).$$

Similarly, from the second term of (1), we get:

$$= -12 + 18\lambda...(1b).$$

Again, similarly, from the third term of (1), we get:

$$= 20 - 20\lambda...(1c).$$

Hence,

$$\therefore |A - \lambda I| = -\lambda^3 + 6\lambda^2 - 8\lambda = 0$$
$$\therefore \lambda(\lambda^2 - 6\lambda + 8) = 0$$
$$\therefore \lambda_1 = 0, \lambda_2 = 2, \lambda_3 = 4$$
$$\therefore \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

Example 9.3 Find the eigenvalues of the following matrix:

$$A = \begin{bmatrix} -4 & \frac{1}{3} & \frac{13}{3} \\ \frac{-3}{2} & \frac{1}{2} & \frac{3}{2} \\ -3 & \frac{1}{3} & \frac{10}{3} \end{bmatrix}.$$

Solution: For the matrix A, we have eigenvalues:

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} -1 \\ \frac{1}{2} \\ \frac{1}{3} \end{bmatrix}.$$

Now, we age going to find eigenvectors corresponds with eigenvalues. For $\lambda_1 = -1$, we have:

$$\begin{bmatrix} -4-(-1) & \frac{1}{3} & \frac{13}{3} \\ \frac{-3}{2} & \frac{1}{2}-(-1) & \frac{3}{2} \\ -3 & \frac{1}{3} & \frac{10}{3}-(-1) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -3 & \frac{1}{3} & \frac{13}{3} \\ \frac{-3}{2} & \frac{3}{2} & \frac{3}{2} \\ -3 & \frac{1}{3} & \frac{13}{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow -3x_1 + \frac{1}{3}x_2 + \frac{13}{3}x_3 = 0$$

$$\frac{-3}{2}x_1 + \frac{3}{2}x_2 + \frac{3}{2}x_3 = 0$$

$$-3x_1 + \frac{1}{3}x_2 + \frac{13}{3}x_3 = 0$$

$$\Rightarrow \vec{x_1} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$$

By the same way, for $\lambda_2 = \frac{1}{2}$, we get $\vec{x_2} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$, and for $\lambda_3 = \frac{1}{3}$, we get $\vec{x_3} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$. Thus, the eigenvector matrix corresponding to the eigenvalues $\begin{bmatrix} -1 \\ \frac{1}{2} \\ \frac{1}{3} \end{bmatrix}$ is $\begin{bmatrix} 3 & 2 & 1 \\ 1 & 1 & 0 \\ 2 & 2 & 1 \end{bmatrix}$.

Example 9.4 Find the complex eigenvalues and eigenvectors of the matrix: $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

Solution:

$$|A - \lambda I| = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$= \begin{bmatrix} 1-\lambda & -1 \\ 1 & 1-\lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\Rightarrow \lambda^2 - 2\lambda + 2 = 0$$
$$\Rightarrow \lambda_{1,2} = \frac{2 \mp \sqrt{4-8}}{2} = 1 \mp i$$

For $\lambda = 1 + i$, we have:

$$A - (1+i)I_2 = \begin{bmatrix} 1-(1+i) & -1 \\ 1 & 1-(1+i) \end{bmatrix} = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix}.$$

Now we row reduce, noting that the second row is i times the first:

$$\begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \xrightarrow{R_2 = R_2 - iR_1} \begin{bmatrix} -i & -1 \\ 0 & 0 \end{bmatrix} \xrightarrow{R_1 = R_1 \div -i} \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix}$$

The parametric form $x = iy$, so an eigenvalue:

$$\vec{v_1} = \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

For $\lambda = 1 - i$, we have:

$$A - (1-i)I_2 = \begin{bmatrix} 1-(1-i) & -1 \\ 1 & 1-(1-i) \end{bmatrix} = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix}.$$

Now we row reduce, noting that the second row is $-i$ times the first:

$$\begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \xrightarrow{R_2 = R_2 + iR_1} \begin{bmatrix} i & -1 \\ 0 & 0 \end{bmatrix} \xrightarrow{R_1 = R_1 \div i} \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}$$

The parametric form $x = -iy$, so an eigenvalue:

$$\vec{v_2} = \begin{bmatrix} -i \\ 1 \end{bmatrix}.$$

Thus, the verify to the answer will be:

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} i-1 \\ i+1 \end{bmatrix} = (1+i) \begin{bmatrix} i \\ 1 \end{bmatrix}.$$
$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -i \\ 1 \end{bmatrix} = \begin{bmatrix} -i-1 \\ -i+1 \end{bmatrix} = (1+i) \begin{bmatrix} -i \\ 1 \end{bmatrix}.$$

Example 9.5 Find the eigenvalues and their corresponding eigenvectors of the following matrix: $\begin{bmatrix} 5 & 1 & -1 & 0 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$.

Solution: By utilizing (9.2), we can find the eigenvalues: $\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 2 \\ 3 \end{bmatrix}$. And, their corresponding eigenvector matrix is

$$\vec{x} = \begin{bmatrix} 1 & \frac{-228}{721} & \frac{228}{721} & \frac{-390}{1351} \\ 0 & \frac{684}{721} & 0 & \frac{1170}{1351} \\ 0 & 0 & \frac{684}{721} & \frac{390}{1351} \\ 0 & 0 & 0 & \frac{390}{1351} \end{bmatrix}.$$

9.11 Conclusions from eigenvalues, eigenvectors and Traces

(i) The eigenvalues of the upper (lower) and diagonal triangular matrix are the same as the eigenvalues of the matrix.

(ii) The sum of the eigenvalues of the matrix is equal to the sum of the elements on the main diagonal of the matrix. This sum is called the Trace of the matrix, and is expressed as: as: $T(A) = \sum_{i=1}^{n} \lambda_i$.

(iii) The product of the eigenvalues is the determinant path of the matrix. Or:

$|A| = \prod_{i=1}^{n} \lambda_i$.

9.12 Additional properties of eigenvalues

We are going to show additional properties of the eigenvalues that have been presented in the previous sections, through inserting and submitting the following theorems in what follows.

Theorem 9.1 *If $\lambda_1, \lambda_2, ..., \lambda_n$ are the eigenvalues of the matrix A, then $\lambda_1^k, \lambda_2^k, ..., \lambda_n^k$ are the eigenvalues of the matrix A^k. Furthermore, each eigenvalue of the matrix is an eigenvalue of the matrix A^k.*

Proof Suppose that λ, X are eigenvalue and eigenvector of the matrix respectively, then:

$$AX = \lambda X$$

Now, by multiplying each side of the equation from left by the matrix A, we get:

$$\begin{aligned} A^2X &= \lambda AX \\ &= \lambda(\lambda X) \\ &= \lambda^2 X \end{aligned}$$

That means λ^2 is the eigenvalue of the matrix A^2, and the corresponding eigenvalue is the same λ.

Now, by using the mathematical induction. We assume that the theorem is true for $r = k$. Or,

$$A^r X = \lambda^r X$$

Again, by multiplying both sides from the left of the matrix A, we get:

$$AA^rX = A\lambda^rX$$
$$\therefore A^{r+1}X = \lambda^rAX$$
$$= \lambda^r(\lambda X)$$
$$= \lambda^{r+1}X$$

That means λ^{r+1} is the eigenvalue of the matrix A^{r+1}. In addition, X is the eigenvector of the matrix A^{r+1}.

Thus, and based on the mathematical induction, the theorem is true for all value of k. ♦

Corollary (Cayley-Hamilton theorem)

Each square matrix satisfies its characteristic equation.

Note: Necessary clarification before proof

The Cayley-Hamilton theorem states that every square matrix over a commutative ring (the real or complex numbers) satisfies its own characteristic equation. If A is a given $n \times n$ matrix and I_n is the $n \times n$ identity matrix, then the characteristic polynomial of A is defined as: $p_A(\lambda) = |\lambda I_n - A|$, where $||$ is the deterministic operation, and λ is a variable for a scalar element of the base ring F (Atiyah, 2018). Since the entries of the matrix $|\lambda I_n - A|$ are (linear or constant) polynomials in λ, the determinant is also a degree-n monic polynomial in λ, so that:

$$p_A(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + ... + c_1\lambda + c_0.$$

Or,

$$p_A(A) = A^n + c_{n-1}A^{n-1} + ... + c_1A + c_0I_n.$$

The Cayley–Hamilton theorem states that this polynomial expression (Hamilton, 2008; Cayley, 1858b) is equal to the zero matrix, which is to say that; $p_A(A) = 0$ that is, the polynomial p_A is an annihilating polynomial for A. The theorem allows A^n to be expressed as a linear combination of the lower matrix powers of A. When the ring is a field F, the Cayley–Hamilton theorem is equivalent to the statement that the minimal polynomial of a square matrix divides its characteristic polynomial.

Proof For a generic 2×2 matrix, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ the characteristic polynomial is given by:

$$p_\lambda = \lambda^2 - (a+d)\lambda + (ad - bc),$$

so the Cayley-Hamilton theorem states that:

$$p(A) = A^2 - (a+d)A + (ad-bc)I_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

$$\begin{aligned}
&\because A^2 - (a+d)A + (ad-bc)I_2 \\
&= \begin{bmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{bmatrix} - \begin{bmatrix} a(a+d) & b(a+d) \\ c(a+d) & d(a+d) \end{bmatrix} + (ad-bc)I_2 \\
&= \begin{bmatrix} bc-ad & 0 \\ 0 & bc-ad+(ad-bc)I_2 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \blacklozenge
\end{aligned}$$

Example 9.6 For 1×1 matrix $A = (a)$ apply the Cayley-Hamilton theorem.

Solution:

$$p(\lambda) = \lambda - a$$

$$\therefore p(A) = a - a(1) = 0.$$

It is trivial solution.

Example 9.7 For 2×2 matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ apply the Cayley-Hamilton theorem.

Solution:

$$
\begin{aligned}
&|\lambda I_2 - A| \\
&= \begin{vmatrix} \lambda - 1 & -2 \\ -3 & \lambda - 4 \end{vmatrix} \\
&= (\lambda - 1)(\lambda - 4) - (-2)(-3) \\
&= \lambda^2 - 5\lambda - 2 \\
&\therefore p(X^2) - 5X - 2I_2 \\
&p(A) = A^2 - 5A - 2I_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}
\end{aligned}
$$

For verifying;

$$
\begin{aligned}
&A^2 - 5A - 2I_2 \\
&= \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} - \begin{bmatrix} 5 & 10 \\ 15 & 22 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}
\end{aligned}
$$

Theorem 9.2 *If matrix A is real and symmetric, then all of its eigenvalues and their corresponding eigenvectors are real as well.*

Proof Let A be a real symmetric matrix and λ be a complex eigenvalue of A. Here, there exists a complex vector x such that;

$$Ax = \lambda x, x \neq 0. \tag{9.8}$$

Based on the definition of eigenvalues. We take the complex conjugates of both sides. Since A is real matrix, then $\bar{A} = A$. Thus;

$$A\bar{x} = \bar{\lambda}\bar{x} \tag{9.9}$$

The transpose of both sides of (9.9). As A is symmetric, we get:

$$\bar{x}^T A = (\bar{\lambda}\bar{x})^T \tag{9.10}$$

$$\therefore (\bar{\lambda}\bar{x})^T x = \bar{x}^T Ax = \bar{x}^T \lambda x = \lambda \bar{x}^T x \tag{9.11}$$

$$\bar{x}^T x = \bar{x}_1 x_1 + \bar{x}_2 x_2 + ... + \bar{x}_n x_n + > 0, x \neq 0 \tag{9.12}$$

Thus, if we divide both sides of (9.11) by $\bar{x}^T x$, we have $\lambda = \bar{\lambda}$. That means, λ is real. ♦

Theorem 9.3 *All eigenvalues of Hermitian matrix are real.*

Proof Suppose that λ is an eigenvalue of a Hermitian matrix A, and x is an eigenvector corresponding to the eigenvalue λ.

$$Ax = \lambda x \tag{9.13}$$

Multiplying both sides of (9.13) by $\bar{x}^T$ from the left, we obtain:

$$\begin{aligned} &\bar{x}^T(Ax) = \bar{x}^T(\lambda x) \\ &= \lambda \bar{x}^T x \\ &= \lambda \|x\| \end{aligned} \tag{9.14}$$

On the other hand, we have;

$$\bar{x}^T(Ax) = (Ax)^T \bar{x} = x^T A^T \bar{x} \tag{9.15}$$

The first equality of (9.15) follows because $u \cdot v$ of the vectors u, v is commutative. Thereby, we have;

$$u \cdot v = u^T v = v^T u = v \cdot u \tag{9.16}$$

When applying the fact of (9.16) with $u = \bar{x}, v = Ax$ we get;

$$x^T A^T \bar{x} = \lambda \|x\| \tag{9.17}$$

Taking the complex conjugate of (9.17), we obtain:

$$\bar{x}^T \bar{A}^T x = \bar{\lambda} \|x\| \tag{9.18}$$

Keep in the mind that; since $\|x\|$ is a real number, hence $\bar{x} = x, \|\bar{x}\| = \|x\|$. Besides, since A is a Hermitian matrix, hence $\bar{A}^T = A$.

$$\begin{aligned} \bar{\lambda}\,\|x\| &= \bar{x}^{T}[\text{from}(9.18)] \\ &= \bar{x}^{T}\lambda x[\text{from}(9.13)] \\ &= \lambda\,\|x\| \end{aligned} \tag{9.19}$$

Since x is an eigenvector, hence x is not the zero vector and the length $\|x\| \neq 0$. Thus, when we divide by the length $\|x\|$, the result will be:

$$\lambda = \bar{\lambda} \tag{9.20}$$

From (9.20), we have proved that the λ is a real number. ♦

In the 1930s, the mathematician Semyon Aranovich Gerschgorin stated a theorem has been known by the Gerschgorin circle theorem may be used to bound the spectrum of a square matrix (Li and Zhang, 2019; Varga, 2010; Varga, 1962; Horn and Johnson, 2012; Gloub and Van Loan, 1996). But before starting the theorem, we need the following definition:

Definition 9.8 Let A be a complex $n \times n$ matrix, with entries a_{ii}. For $i \in \{1, 2, ..., n\}$, and let R_i be the sum of the absolute values of the non-diagonal entries in the ith row:

$R_i = \sum_{j \neq i} |a_{ij}|$. Let $D(a_{ij}, R_i) \subseteq \mathbb{C}$ be a closed disc centered at a_{ii} with radius R_i. Such a disc is called a Gerschgorin disc (Li and Zhang, 2019; Varga, 2010).

Theorem 9.4 (Gerschgorin theorem) *Every eigenvalue of A lies within at least one of the Gerschgorin discs $D(a_{ii}, R_i)$. Or an alternative: Every eigenvalue of the matrix A lies in at least one of the circles whose center is a_{ii} and whose radii are $r_i = \sum_{j=1}^{n} |a_{ij}|\,; j \neq i, i = 1, 2, ..., n$.*

Proof If λ is an eigenvalue of the matrix A then there is at least one its corresponding nonzero eigenvector x, such that:

$$Ax = \lambda x \tag{9.21}$$

Now, suppose that $x_p \in x$ is the maximum absolute value, such the x can be expressed as:

$$x = \begin{bmatrix} x_1 & x_2 & \dots & x_{p-1} & \dots & x_{p+1} & \dots & x_n \end{bmatrix}^T \tag{9.22}$$

Dividing (9.22) by x_p where $|x_i| \leq 1, \forall i \wedge i \neq p$. Thus in (9.21), we notice p:

$$\sum_{j=1}^{n} a_{pj} x_j = \lambda x_p = \lambda \tag{9.23}$$

$$\begin{aligned} \therefore \left|\lambda^{-1} pp\right| &= \left|\sum_{j=1}^{n} a_{pj} x_{j-1} pp\right| \\ &\leq \left|\sum_{j=1}^{n} a_{pj} x_j\right|, j \neq p \\ &\leq \left|\sum_{j=1}^{n} a_{pj}\right| |x_j| \\ &\leq \left|\sum_{j=1}^{n} a_{pj}\right|, j \neq p. \ \blacklozenge \end{aligned} \tag{9.24}$$

Theorem 9.5 *For any real symmetric matrix there exists a matrix Q such that $Q^{-1}AQ$ is a diagonal matrix and has the same eigenvalues of matrix A (The matrix $Q^{-1}AQ$ is called a matrix similar to A).*

Proof If a matrix A is similar to a diagonal matrix $Q^{-1}AQ$, then the eigenvalues are known, and the theorem has been proved according to studies (Horn, 1985; Horn and Johnson, 2012; Srivastava and Sahami, 2009; Nearing et al., 2003). ⧫

Theorem 9.6 *Any similar transformation $P^{-1}AP$ applied to the matrix A does not change its eigenvalue.*

Proof Suppose that λ, x are eigenvalue and eigenvector of the matrix A respectively. Thus, we get:

$$\begin{aligned} Ax &= \lambda x \\ PAx &= \lambda P x \end{aligned} \tag{9.25}$$

If, we assume that $Px = y$ then we have:

$$x = P^{-1}y \tag{9.26}$$

Now, substituting (9.26) in (9.25), we get:

$$PAP^{-1}y = \lambda y \tag{9.27}$$

Thus, we conclude that λ, y are the eigenvalue and eigenvector of the matrix $P^{-1}AP$ correspondingly. ♦

Theorem 9.7 *For any square matrix A there is a similarity matrix P where the matrix A is transformed into a triangle similarity matrix $T = P^{-1}AP$.*

Proof Based on Theorems (9.5 & 9.6). Besides, by using some algebraic steps on the matrix A, we can find a similar matrix P to get $P^{-1}AP = T$. ♦

9.13 Methods for finding eigenvalues

There are two main methods for finding the eigenvalues of square matrices, and in this section we will deal them in some detail with real examples.

9.13.1 LU method

It is a method of converting any square matrix whose eigenvalues are real numbers into a triangular matrix as defined and described in chapter 5, as we can see in (5.5.11& 5.5.12).

The matrix A can be converted into submatrices LU as follows:

Let us assume that;

$$A_1 = UL \tag{9.28}$$

It should be noticed that the eigenvalue of A_1 in (9.28) is the same of the matrix A, because;

$$A_1 = UR = RAR^{-1} \tag{9.29}$$

The matrix in (9.29) is a similar of the matrix A. Again, after some steps we can convert A_1 into LU, and the result will be as follows;

$$A_2 = UL \tag{9.30}$$

We continue these successive steps until obtain on triangular matrices $A_3, A_4, ...A_n$ in which their eigenvalues are the diagonal elements, where they are the same eigenvalues of A, as we emphasized that in Section 9.11.

Example 9.8 Use LU method to find the eigenvalues of the following matrix:

$$\begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & 6 \\ -1 & -6 & 0 \end{bmatrix}.$$

Solution:

$$\because \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & 6 \\ -1 & -6 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ \frac{1}{2} & -1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 2 & -3 \\ 0 & 3 & -9 \\ 0 & 0 & \frac{15}{2} \end{bmatrix},$$

$$\therefore L = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ \frac{1}{2} & -1 & 1 \end{bmatrix}, U = \begin{bmatrix} -2 & 2 & -3 \\ 0 & 3 & -9 \\ 0 & 0 & \frac{15}{2} \end{bmatrix}. \tag{9.31}$$

$$\therefore A_1 = UR = \begin{bmatrix} -\frac{11}{2} & 5 & -3 \\ \frac{15}{2} & 12 & 9 \\ -\frac{15}{4} & \frac{15}{2} & -\frac{15}{2} \end{bmatrix}.$$

After we partition A_1 into L and U, and repeat the same steps in (9.31) we have:

$$A_2 = \begin{bmatrix} -\frac{73}{10} & \frac{263}{100} & -3 \\ \frac{93}{25} & \frac{31}{100} & -\frac{491}{100} \\ -\frac{27}{25} & frac54 & -\frac{79}{50} \end{bmatrix}.$$

.

.

.

$$\therefore A_{10} = \begin{bmatrix} \frac{116}{25} & \frac{319}{100} & -3 \\ \frac{71}{100} & -\frac{27}{10} & -\frac{7}{25} \\ -\frac{37}{500} & frac307100 & -\frac{297}{100} \end{bmatrix}. \tag{9.32}$$

.

.

.

$$\therefore A_{30} \cong \begin{bmatrix} 5 & \frac{16}{5} & -3 \\ 0 & -3 & 0 \\ 0 & 0 & --3 \end{bmatrix}.$$

Thus, in (9.32), we find that:

$$\therefore \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \\ -3 \end{bmatrix}.$$

9.13.2 Jacobi method

The Jacobi eigenvalue algorithm is an iterative method for the calculation of the eigenvalues and eigenvectors of a real symmetric matrix, namely in the numerical lineal algebra, and the process known as diagonalization (Jacobi, 1846; Golub and Van de Vorst, 2000).

The Jacobi eigenvalue method repeatedly performs rotations until the matrix becomes almost diagonal. Then the elements in the diagonal are approximations of the real eigenvalues of the undertaken matrix (Schönhage, 1964).

In this method, we consider a real symmetric matrix A, and transform it into a diagonal matrix $Q^T AQ$ (Press et al., 1992; Rutishauser, 1966; Saad, 2023) by eliminating all the non-diagonal elements one by one through the following algorithm (Sameh, 1971; Shroff, 1990; Veselić, 1979):

(i) To get rid of the element a_{pq}, we find the matrix Q, in which $Q_{ii} = 1, i \neq p, q$. And for the remained elements $Q_{ij} = 0$ except $Q^T Q = I$.

(ii) Put $Q_{pp} = Q_{qq} = cos(Q), Q_{pq} = -sin(Q), q_{qp} = sin(Q)$ where $Q = \frac{1}{2}tan^{-1}(\frac{2a_{pq}}{a_{pp}-a_{qq}})$.

(iii) Computing the matrix $A_1 = Q^T AQ$ at the zero position of (p, q).

(iv) Change $p = p+1, q = q+1$ and find the matrix Q_1 with the same properties of the above matrix Q.

(v) Computing the matrix $A_2 = Q_1^T A_1 Q_1$ at the zero position of $(p+1, q+1)$.

(vi) It should be noticed that the value of zero position at (p, q) in the matrix A_1 ,ay be changed its value from zero in the matrix A_2.

(vii) Repeating the same steps for the other positions to get rid of the non-diagonal elements till obtain on the matrix $A_{n\frac{n-1}{2}}$.

(viii) Keep in mind may be we have to repeat all the above steps to obtain the last matrix $A_{n\frac{n-1}{2}}$.

Example 9.9 Find the eigenvalues of the following symmetric matrix by Jacobi method: $\begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix}$.

Solution: By applying Jacobi method, we want to rid of $a_{21} = 7$, we have:

$$\begin{aligned} Q &= \frac{1}{2}tan^{-1}(\frac{2a_{pq}}{a_{pp}-a_{qq}}) \\ &= \frac{1}{2}tan^{-1}(\frac{(2)(7)}{5-10}) \\ &= -\frac{1829}{52} \\ &\therefore \sin(Q) = \frac{317}{549}, \cos(Q) = -\frac{1201}{1471} \end{aligned} \tag{9.33}$$

Thereby, the matrix will be:

$$Q = \begin{bmatrix} \cos(\frac{1829}{52}) & -\sin(\frac{-1829}{52}) & 0 & 0 \\ \sin(-\frac{1829}{52}) & \cos(\frac{1829}{52}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1201}{1471} & \frac{317}{549} & 0 & 0 \\ -\frac{317}{549} & \frac{1201}{1471} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{aligned} &\therefore A_1 = QAQ^{-1} \\ &= \begin{bmatrix} \frac{1201}{1471} & \frac{317}{549} & 0 & 0 \\ -\frac{317}{549} & \frac{1201}{1471} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix} \begin{bmatrix} \frac{1201}{1471} & -\frac{317}{549} & 0 & 0 \\ \frac{317}{549} & \frac{1201}{1471} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{2897}{194} & 0 & \frac{2559}{256} & \frac{2314}{269} \\ 0 & \frac{538}{8029} & \frac{373}{1335} & \frac{187}{4633} \\ \frac{2559}{256} & \frac{373}{1335} & 10 & 9 \\ \frac{2314}{269} & \frac{187}{4633} & 9 & 10 \end{bmatrix}. \end{aligned} \tag{9.34}$$

Now, we are going to rid from $a_{31} = \frac{2559}{256} \in A_1$. By the same way we obtain:

$$A_2 = \begin{bmatrix} \frac{11381}{500} & \frac{552}{3023} & 0 & \frac{3179}{258} \\ \frac{552}{3023} & \frac{67}{1000} & \frac{2331}{10000} & \frac{107}{1949} \\ 0 & \frac{2331}{10000} & \frac{3020}{1391} & \frac{1099}{617} \\ \frac{3179}{258} & \frac{103}{1883} & \frac{1099}{617} & 1 \end{bmatrix}$$

$$\vdots \qquad (9.35)$$

$$\therefore A_{18} = \begin{bmatrix} \frac{7239}{239} & 0 & 0 & 0 \\ 0 & \frac{51}{5000} & 0 & 0 \\ 0 & 0 & \frac{446}{529} & 0 \\ 0 & 0 & 0 & \frac{3453}{895} \end{bmatrix}$$

Thus, from (9.35), we find that, the eigenvalues of the matrix A are:

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} = \begin{bmatrix} \frac{7239}{239} \\ \frac{51}{5000} \\ \frac{446}{529} \\ \frac{3453}{895} \end{bmatrix} \cong \begin{bmatrix} 30.2887 \\ 0.0102 \\ 0.8431 \\ 3.8581 \end{bmatrix}.$$

9.14 Exercises

Solve the following questions:

Q1: Find the eigenvalues of the following matrices and what do you notice? Which observation do I mention?

(i) $\begin{bmatrix} a & b \\ d & b \end{bmatrix}, \forall a, b, c, d \in \mathbb{R}.$

(ii) $\begin{bmatrix} a & b \\ d & b \end{bmatrix}, \forall a, b, c, d \in \mathbb{C}.$

(iii) $\begin{bmatrix} a & b \\ d & b \end{bmatrix}, \forall a, b, c, d \in \mathbb{C}.$

(iv) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \\ 4 & 0 & 0 \end{bmatrix}.$

(v) $\begin{bmatrix} -\sin\alpha & \cos\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}$.

(vi) $\begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}$.

(vii) $\begin{bmatrix} 1 & 7 & 3 \\ 7 & 4 & 5 \\ 3 & 5 & 2 \end{bmatrix}$.

(viii) $\begin{bmatrix} 1 & 7 \\ 7 & 1 \end{bmatrix}$.

(ix) $\begin{bmatrix} 2 & -2 & 3 \\ 1 & 4 & 5 \\ 2 & 1 & -3 \end{bmatrix}$.

(x) $\begin{bmatrix} 1 & -2 & 1 & -1 \\ 1 & 1 & -2 & 3 \\ 4 & 1 & -5 & 8 \\ 5 & -7 & 2 & -1 \end{bmatrix}$.

Q2: Consider a matrix A and its eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$. Prove that:

(i) A^T has the same eigenvalues of A.

(ii) The eigenvalues of kA are $k\lambda_i, \forall i$, and k is constant.

(iii) A^{-1} has the eigenvalues $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, ..., \frac{1}{\lambda_n}$.

(iv) The eigenvalues of an nonsingular matrix are not zeros.

Q3: Prove that $|(A - \lambda I)^{-1}| = \frac{1}{|A-\lambda I|}$.

Q4: If $\lambda_i, i = 1, 2, ..., k$ is an eigenvalue of the matrix $A_{n\times n}$ repeated k times, then rank of the $A - \lambda_i I \geq n - k$.

Q5: Prove that the characteristic equation of the orthogonal matrix $(Q^T Q = I)Q$ is the inverse equation $\gamma(\lambda) = |\lambda I - Q| = \mp\lambda^n \gamma(\frac{1}{\lambda})$.

Q6: Assume that $Q^T = Q^{-1}$, where Q^T is the transpose of the matrix Q, and Q^{-1} is the inverse of Q, where $Q^T = Q^{-1}$. Prove that $Q^T Q = QQ^T = I$.

10

Permutations and Combinations

10.1 Introduction

Concepts of permutations and combinations are of a great importance in the various sciences in general and in mathematics in particular.

Their study requires a systematic, academic, and mathematical study to show their effective role in several fields. For example, and not limited, whenever studying statistics; it is imperative for the researcher to be equipped with them in order to be a productive scientific, theoretical, and practical researcher in order to produce tangible and realistic results. Because in the study of statistics, we want to find the number of methods in which an event can occur, and the number of ways in which the elements of the set taken in the space of possibilities can be arranged. In addition to the number of different sets that can be composed of a certain number of things.

Due to the importance of rolling the permutations and combination, we can identify some fields in which these two concepts have their roles, which will be discussed in the following sections.

10.2 Permutations and combinations and their formulations

In this section, we delve into the basic concepts of permutation and combination, their role in mathematics and the fields in which they appear, in addition to the basic principle of the arithmetic.

10.2.1 Basic principle of the arithmetic

The basic principle of arithmetic is based on non-empty sets, and depending on the elements in these sets. We can find the number of ways in which we select the elements in the sets taken for arithmetic in the process. To illustrate, let us assume the following hypothesis:

Assume that there are $A_1, A_2, ..., A_k$ sets under taken and each one contains of $n_1, n_2, ..., n_k$ elements respectively. The number of ways to select an element in $A_i, i = 1, 2, ..., k$ is $n_1.n_2...n_k$.

Example 10.1 There are eight departments in a faculty, and each department chooses a representative in a committee. If the number of teaching staffs in these departments is $33, 23, 22, 28, 13, 6, 5, 4$, respectively. In how many ways can the representatives in the committee for the departments be chosen?

Solution: It can be chosen the representative in a committee for the first department in 33 ways, in 23 ways for the second department, and so on for the remain departments respectively. Thereby, the number of chosen of the representatives of a committee for the departments is;

$33 \cdot 23 \cdot 22 \cdot 28 \cdot 13 \cdot 6 \cdot 5 \cdot 4 = 729368640$ ways.

Now, let us ask this following question. Suppose we have a set containing of n elements, and we want to order the elements of this set. In how many ways can we order the elements of the set mathematically? Of course, to order these elements, we choose an element to be the first, then we choose another element to be the second, ...etc. And each order in this kind is called permutation. Thus, in what follows we have to dive of the concept of permutation.

10.2.2 Permutation

Definition 10.1 A permutation of a set is an arrangement of its members into a sequence or linear order, or if the set is already ordered, a rearrangement of its elements. Or, it is refers to the act or process of changing the linear order of an ordered set (Landau, 1994; Gold, 1984; McCoy, 1968; Nering, 1970; Heath, 2013).

Example 10.2 A string of length n has $n!$ permutation. What is a permutation of letters A, B, C from a group of letters takes all of them, regardless of their meaning in the logic of the language?

Solution: The permutation of string A, B, C is, $ABC, ACB, BAC, BCA, CBA, CAB$. Or, the permutation is $3! = 3 \cdot 2 \cdot 1 = 6$, as shown in Figure 10.1.

Example 10.3 Write the different permutations of the numbers $1, 2, 3, 4$ taken all of them at every permutation.

Solution: The number of permutations are $4 \cdot 3 \cdot 2 \cdot 1 = 24$.
As shown below:

$$4123, 4132, 4231, 4312, 4321, 1423,$$
$$1432, 1243, 1234, 1342, 1324, 2413,$$
$$2431, 2143, 2134, 2341, 4312, 3412,$$
$$3421, 3142, 3124, 3241, 3214, 2314.$$

Example 10.4 Write the different permutations of the letters x, y, z, w, in which two letters are taken from it at a time.

Solution: The number of permutations are $4 \cdot 3 = 12$.
As shown below:

$$xw, yw, zw, xy,$$
$$xz, xw, xy, yz,$$
$$wy, xz, yz, wz.$$

Definition 10.2 The factorial of a $n \in \mathbb{N}$, denoted by $n!$ is the product of all positive integers less than or equal to n. Or;
$n! = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot 3 \cdot 2 \cdot 1 = \prod_{k=1}^{n} k$ (Graham et al., 1994).

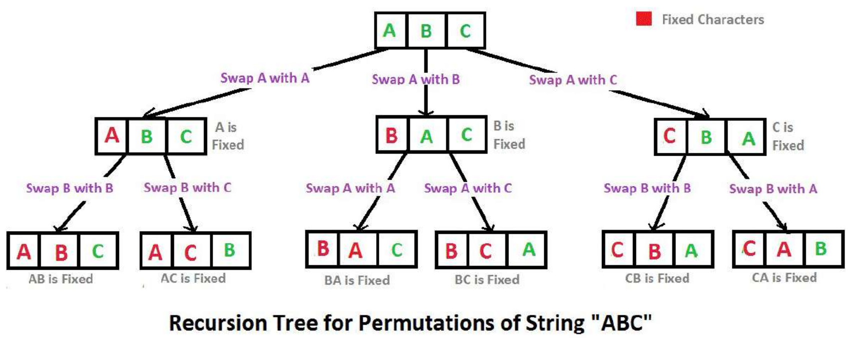


Figure 10.1: Permutations

Example 10.5

$$\begin{aligned}
&1! = 1\\
&2! = 2 \cdot 1 = 2\\
&3! = 3 \cdot 2 \cdot 1 = 6\\
&.\\
&.\\
&.\\
&k! = k \cdot (k-1) \cdot (k-3) \cdot ... \cdot 3 \cdot 2 \cdot 1
\end{aligned}$$

Definition 10.3 The factorial of zero is one, and expressed mathematically as; $0! = 1$ (Goldenberg and Carter, 2017; Dorf, 2003; Hamkins, 2020).

Definition 10.4 For all non-positive n, it can be defined a function Γ in which $\Gamma(n+1) = n\Gamma(n) = n! = \lim_{m\to\infty} \frac{m^n m!}{(n+1)(n+2)...(n+m)}$. The function called Gamma function (Farrell and Ross, 2013; Beals and Wong, 2010).

Example 10.6 If $(\frac{1}{2})! = \frac{\sqrt{\pi}}{2}$ then find $\Gamma(\frac{1}{2}), \Gamma(\frac{-1}{2})$.
Solution: Based on Definition 10.4, we have;

$$\begin{aligned}
&(\frac{1}{2})! = \Gamma(\frac{1}{2}+1) = \frac{1}{2}\Gamma(\frac{1}{2})\\
&= \frac{\sqrt{\pi}}{2}\\
&\therefore \Gamma(\frac{1}{2}) = \sqrt{\pi}.\\
&\because \Gamma(n+1) = n\Gamma(n),\\
&\therefore \Gamma(\frac{-1}{2}+1) = -\frac{1}{2}\Gamma(\frac{-1}{2})\\
&= \sqrt{\pi}\\
&\therefore \Gamma(\frac{-1}{2}) = -2\sqrt{\pi}.
\end{aligned}$$

Theorem 10.1 *The number of permutations of n different objects taken at a time is given by $n!$.*

Proof Permutation of n elements in a systematic and specific way. We choose an element to be the first, and thus there are n ways to choose it.

Now, there are $n-1$ remained of objects and there are $(n-1)$ ways to select it.

Repeating this procedure for the element k of the objectives, it can be $(n-k+1)$ of ways to chosen it.

Thus, we concussion based on the basic principle of the arithmetic there are;

$n! = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot 3 \cdot 2 \cdot 1$ permutations for n of elements. ♦

Example 10.7 There is a family of eight people who want to know the number of ways to arrange themselves in a:

(i) Straight line.

(ii) Round table.

Solution: (i) The number of arrange of eight people on a straight line is;

$8! = 8 \cdot 7 \cdot ... \cdot 3 \cdot 2 \cdot 1 = 40320.$

(ii) To arrange the family in a circular manner, it must fixed one of its members and arrange the rest of the members with respect to him. Thus, the process can be accomplished as follows;

$7! = 5040.$

Theorem 10.2 *The number of permutations of n different objects taken k at a time is given by:*

$P_k^n = \frac{n!}{(n-k)!} = n(n-1)(n-2)...(n-k+1).$

Proof

$$P_k^n = nP_{k-1}^{n-1}$$
$$= n\frac{(n-1)!}{((n-1)-(k-1))!}$$
$$= \frac{n!}{(n-k)!}$$
$$= P_k^n$$

Continuing this process we get;

$$P_k^n = nP_{k-1}^{n-1} = n(n-1)P_{k-2}^{n-2}$$
$$= n(n-1)(n-2)P_{k-3}^{n-3}...(n-(k-1))P_0^{n-k}$$
$$= n(n-1)(n-2)...(n-k+1)$$
$$\therefore P_k^n = P_k^{n-1} + kP_{k-1}^{n-1}$$
$$= \frac{(n-1)!}{((n-1)-k)!} + k\frac{(n-1)!}{(n-k)!}$$
$$= \frac{(n-1)!}{(n-1-k)!} + k\frac{(n-1)!}{(n-k)!}$$
$$= \frac{(n-1)!(n-k)}{(n-1-k)!(n-k)} + k\frac{(n-1)!}{(n-k)!}$$
$$= \frac{(n-1)!(n-k)}{(n-k)!}k\frac{(n-1)!}{(n-k)!}$$
$$= \frac{(n-1)!((n-k)+1)}{(n-k)!}$$
$$= \frac{n(n-1)!}{(n-k)!}$$
$$= \frac{n!}{(n-k)!}$$
$$= P_k^n. \blacklozenge$$

It can generalize Theorem 10.2 to the general form (Brualdi, 2004) as follows:

Theorem 10.3 (Generalized theorem) *A set consisting of n elements is partition to k subsets containing n_1 of similar elements, n_2*

of others similar elements, ..., n_k of others similar elements respectively such that: $n = \sum_{i=1}^{k} n_i$. *Then, the number of permutations given by:*
$P^{n}_{n_1,n_2,\ldots,n_k} = \frac{n!}{n_1!n_2!\ldots n_k!} = \frac{(\sum_{i=1}^{k} n_1)!}{\prod_{i=1}^{k} n_i!}$.

Proof It is possible to prove this generalized theorem as a corollary of Theorem 10.2, and based on the aforementioned reference. ♦

Example 10.8 If you are given an eight-letter word, in how many ways can it be a five-letter word, regardless of the meanings of the component words?

Solution: Due to the requirements of the problem, the matter is permutation of five words out of eight words. Thereby;

$$P^8_5 = \frac{8!}{(8-5)!} = \frac{8!}{3!} = 6720.$$

Example 10.9 How many numbers can you make of $0, 1, 2, 3, \ldots, 9$ such that start from 5 and have four digits?

Solution: We exclude the number 5 because we will put it with every number we make, so what remains is the formation of numbers of 4 digits. Thus;

$$P^9_4 = \frac{9!}{(9-4)!} = \frac{9!}{5!} = 3024.$$

Example 10.10 How many distinct permutations can be formed from the word "committee"?

Solution: The ninth permutations which consist of letters c, o, m, i, t, and e can be obtained as follows:

$$P^9_{1!,1!,2!,1!,2!,2!} = \frac{9!}{1!1!2!1!2!2!} = 45360.$$

Note: $P^n_k = \binom{n}{k} = \frac{n!}{(n-k)}$.

10.2.3 Basic properties of permutations

Permutations refer to the arrangement of objects, where the order of arrangement is important. The following are some basic properties of permutations:

(i) **Identity Permutation**: This is a permutation which leaves all elements in their original place. For example, for the set 1, 2, 3, the identity permutation is (1, 2, 3).

(ii) **Permutation of n different things taken k at a time**: The number of permutations of n different things taken k at a time, denoted by P_k^n, is given by $P_k^n = \frac{n!}{(n-k)!}$, where ! denotes factorial.

(iii) **Permutation of n things not all different**: If there are p objects of one kind, q objects of second kind, r objects of third kind, and so on, then the number of permutations of these objects given by $P_{p,q,r,\ldots}^n = \frac{n!}{p!q!r!\ldots}$.

(iv) **Permutation with Repetition**: If we are allowed to repeat objects, then the number of permutations of n objects taken k at a time is n^k.

(v) **Number of Circular Permutations**: If there are n different things to be arranged in a circle, then the number of permutations is $(n-1)!$.

10.2.4 Combination

Definition 10.5 A combination is a selection of items from a set that has distinct members, such that the order of selection does not matter. Mathematically, If a set has n elements, the number of $k-$combinations, denoted by $C(n,k) = C_k^n = \binom{n}{k} = \begin{cases} \frac{n!}{(n-k)!k!} = \frac{n(n-1)\ldots(n-k+1)}{k(k-1)\ldots3\cdot2\cdot1}, \forall k \leq n \\ 0, \forall n < k \end{cases}$ (Mazur, 2022; Ryser, 1963; Brualdi, 2004).

Note: The formulation of the combination can be derived from the fact that each $k-$ combination of a set A of n members has $k!$ permutations. So, $P_k^n = C_k^n \cdot k!$. Or $C_k^n = \frac{P_k^n}{k!}$ (Reichl, 2016). Thereby, the set of all $k-$combinations of a set A is often denoted by $\binom{A}{k}$.

Example 10.11 The number of ways in which 3 square objects can be selected from a set of 5 distinct square objects, without regard to their order.

Solution: The number of combinations are 10 ways, as shown in Figure 10.2.

$$C_3^5 = \frac{5!}{3!(5-3)!} = \frac{5 \cdot 4 \cdot 3!}{3!2!} = 5 \cdot 2 = 10.$$

Theorem 10.4 *Consider a set A contains n of elements. The number of combinations that can be formed from n elements taken k each time, regardless order is $C_k^n = \frac{n!}{k!(n-k)!}$.*

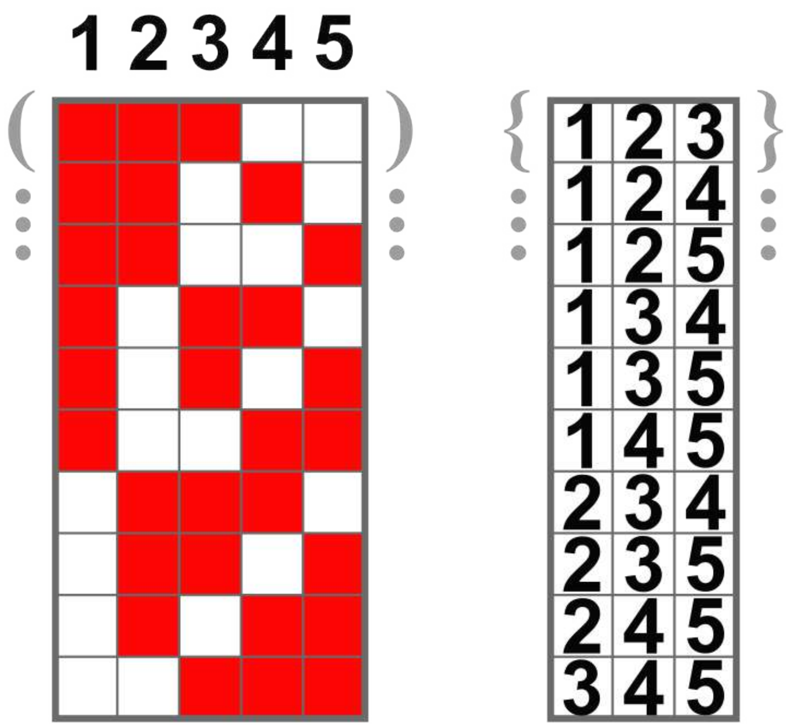


Figure 10.2: Combinations

Proof

$$C_k^n = \frac{n(n-1)(n-2)...(n-k+1)}{k!}$$
$$= x...(1)$$
$$C_{k-1}^n = \frac{n(n-1)(n-2)...(n+1)}{(k-1)!}$$
$$= x...(2)$$
$$\because (1) + (2)$$
$$= \frac{n(n-1)(n-2)...(n-k+1)}{k!}$$
$$+ \frac{n(n-1)(n-2)...(n+1)}{(k-1)!}$$
$$= 2x$$
$$= \frac{k![n(n-1)(n-2)...(k+1)]}{k!(n-k)!}$$
$$+ \frac{(n-k)![n(n-1)(n-2)...(n-k+1)]}{k!(n-k)!}$$
$$\because k![n(n-1)(n-2)...(k+1)] = 1 \cdot 2 \cdot 3 \cdot ... \cdot k$$
$$\therefore k![(k+1)(k+2)...(n-2)(n-1)n]$$
$$= n!$$
$$\because [n(n-1)(n-2)...(n+1)](k-1)!$$
$$= 1 \cdot 2 \cdot 2 \cdot ... \cdot (n-k)$$
$$\therefore (n-k)![(n-k+1)(n-k+2)...(n-2)(n-1)n]$$
$$= n!$$
$$\therefore 2x = \frac{n!}{(n-k)!k!} + \frac{n!}{k!(n-k)!}$$
$$= \frac{2n!}{(n-k)!k!}$$
$$\therefore C_k^n = \frac{n!}{k!(n-k)!}. \blacklozenge$$

It can generalize Theorem 10.4 to the general form (Mazur, 2022; Brualdi, 2004; Ryser, 1963) as follows:

Theorem 10.5 (Generalized theorem) *Assume that a set A consists of n elements is partition to k subsets $A_1, A_2, ..., A_k$ containing of $n_1, n_2, ..., n_k$ elements respectively such that: $n = \sum_{i=1}^{k} n_i$. Then, the number of combinations given by:* $\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}...\binom{n-n_1-n_2-n_3-...n_k}{n_k} = \binom{n}{n_1,n_2,...,n_k} = C^n_{n_1,n_2,...,n_k} = \frac{n!}{n_1!n_2!...n_k!}$.

Proof It is possible to prove this generalized theorem as a corollary of Theorem 10.4, and based on the aforementioned references, as follows:

By the product rule the total number of combinations is;

$C(n, n_1)C(n - n_1, n_2)C(n - n_1 - n_2, n_3)...C(n - n_1 - n_2 - ... - n_k, n_k)$

$\because$ the n_1 objects of type one can be placed in the n positions in $C(n, n_1)$ ways leaving $n - n_1$ positions.

$\because$ then n_2 objects of type two can be placed in the $n - n_1$ positions in $C(n - n_1, n_2)$ ways leaving $n - n_1 - n_2$ positions.

Continuous in this fashion, until n_k objects of type k can be placed in the $C(n - n_1 - n_2 - ... - n_k, n_k)$ ways, the product can be obtained as follows:

$$\frac{n!}{n_1!(n-n_1)!} \cdot \frac{(n-n_1)!}{n_2!(n-n_1-n_2)!} \cdot ... \cdot \frac{(n-n_1-...-n_{k-1})!}{n_k!0!}$$
$$= \frac{n!}{n_1!n_2!...n_k!} = C^n_{n_1,n_2,...,n_k}. \blacklozenge$$

Example 10.12 Prove that $\binom{n}{k} = \binom{n}{n-k}$.

Solution:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$
$$= \frac{n(n-1)(n-2)...(k+1)k(k-1)...3 \cdot 2 \cdot 1}{k(k-1)(k-2)...3 \cdot 2 \cdot 1(n-1)!}$$
$$= \frac{n(n-1)(n-2)...(k+1)}{(n-k)!} = \frac{P^n_{n-k}}{n-k}$$
$$= \binom{n}{n-k}.$$

Example 10.13 In how many ways can three committees, each of which consists of five people, be elected from twenty-five people, so that no person is on more than one committee?

Solution: The number of ways to elect the first committee:

$$\binom{25}{5} = \frac{25!}{5!(25-5)!} = \frac{25!}{5!20!}.$$

If five people are elected, twenty people remain, and from them the second committee consisting of five people is elected as follows:

$$\binom{20}{5} = \frac{20!}{5!15!}.$$

As for the third committee, it will be elected as follows:

$$\binom{15}{5} = \frac{15!}{5!10!}.$$

Thus, the number of electing three committees are:

$$\binom{25}{5}\binom{20}{5}\binom{15}{5} = \frac{25!}{5!20!}\cdot\frac{20!}{5!15!}\cdot\frac{15!}{5!10!} = \frac{25!}{(5!)^3 10!} = 2473653742560.$$

Or, the committees can be elected in 2 trillions, 473 billions, 653 millions, 742 thousands, and 560 ways.

Example 10.14 In how many ways can a set of children's toy balls be chosen that contains eight balls from among thirty red balls and ten black balls, so that the set formed contains:

(i) Only three red balls.

(ii) At least on five red balls and two black balls.

Solution:

(i) The number of ways to choose three red balls from thirty is $\binom{30}{3}$. The number of ways to choose five black balls from ten is $\binom{10}{5}$. Thus, the number of ways to chosen the balls is:

$$\binom{30}{3}\binom{10}{5} = \frac{30!}{3!(30-3)!}\cdot\frac{10!}{5!(10-5)!}.$$

(ii) The chosen set balls contains of five red balls and three black balls, or contained of six red balls and two black balls is:

(a) In the first case: $\binom{10}{3}\binom{30}{5} = \frac{10!}{3!7!} \cdot \frac{30!}{5!25!}$.

(b) In the second case: $\binom{10}{2}\binom{30}{6} = \frac{10!}{2!8!} \cdot \frac{30!}{6!24!}$.

Thus, the number of possible ways are:

$$\binom{10}{3}\binom{30}{5} + \binom{10}{2}\binom{30}{6} = \frac{10!}{3!7!} \cdot \frac{30!}{5!25!} + \frac{10!}{2!8!} \cdot \frac{30!}{6!24!}.$$

Example 10.15 It is intended to form four committees of twelve people, such that the first committee includes seven people, the second includes five people, the third includes three people, and the fourth includes three people. What are the number of ways in which these committees can be selected?

Solution: Assume that X is the set of twelve people such that;

$X = \{X_1, X_2, X_3, X_4\}$, where $X_i, i = 1, 2, 3, 4$ contains of $7, 5, 3, 3$ people respectively. thus, the number of ways to form the committees are:

$$\binom{12}{7, 5, 3, 3} = \frac{12!}{7!5!3!31} = 22.$$

10.2.5 Basic properties of combinations

(i) **The relationship between combinations and permutations is direct relationship**:

$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{P_n^k}{k!}, \forall k \leq n.$

(ii) **Identity: Symmetry of Combinations**: $\binom{n}{k} = \binom{n}{n-k}, \forall k \leq n.$

(iii) **Recursive Relationship in Combinations**: $\binom{n-1}{k} + \binom{n-1}{k-1} = \binom{n}{k}, 0 < k < n.$

(iv) **Identity: Summation of Combinations**: $\sum_{k=0}^{n} \binom{n}{k} = 2^n.$

(v) **Identity: Alternating Sum of Combinations**: $\sum_{k=0}^{n}(-1)^k\binom{n}{k} = 0$.

(vi) $\binom{n}{0} = \binom{n}{n} = 1$.

10.2.6 Difference between combination and permutations

In general terms, where combination means selection, permutation means arrangement. Let us check out some of the general points of differences between combination and permutation, as shown in Table 10.1.

Table 10.1: Difference between combination and permutations

(i)	Combination	Permutation
(1)	It is a way of selecting items from a large group of available objects, where the order of selection is not considered.	Permutation refers to the different ways of arranging objects, in a sequential order.
(2)	Order in combination is irrelevant.	Order in permutation is relevant.
(3)	It does not signify the arrangement of objects.	It signifies the arrangement of objects.
(4)	From a single permutation we can derive a single combination.	From a single combination we can derive multiple permutations.
(5)	These can be considered as unordered set.	These are the ordered elements.

10.3 Binomial theorem

The binomial theorem in its special cases was known since at least the 4th century BC when Greek mathematician Euclid mentioned the special case of the binomial theorem for exponent $n = 2, 3$ (Coolidge, 1949; Martzloff, 2007). In 6th century AD, the Indian mathematicians probably knew how to express this as a combinations, and a clear statement of this rule can be found in the 12th century text Lilavati by Bhaskara (Biggs, 1979).

Binomial theorem had been known in its formulation for the first time, and its table in a book by Al-Karaji, as quoted in the book named Al-Buhair published by Al-Samawal(Yadegari, 1980; Katz, 1998a; Rashed, 2013). Al-Karaji described the triangular pattern of the binomial coefficients, and provided a mathematical proof of both the binomial theorem and Pascal's triangle, using a form of mathematical induction (O'Connor and Robertson, 1999).

Many researchers and scientists left traces in formulating and developing the binomial theory (Coolidge, 1949; Landau, 2007; Martzloff, 2007; Kline, 1990; Katz, 1998b; Bourbaki, 2008; Stillwell and Stillwell, 1989) until it was formulated in its current form, as we discuss, state, and prove in the following theorem.

Theorem 10.6 (Binomial theorem) *If* $a, b \in \mathbb{R}$*, and* $n \in \mathbb{N}$ *then:*

$$\begin{aligned}
&\sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^{k} \\
&= a^{n} + n a^{n-1} b + \frac{n(n-1)}{2!} a^{n-2} b^{2} \\
&+ \frac{n(n-1)(n-2)...(n-k+1)}{n!} a^{n-k} b^{k} + ... + b^{n} \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{k} b^{n-k}.
\end{aligned}$$

Proof We will prove the theorem using mathematical induction.

(1) If $n = 1$ then:
$(a+b)^1 = \binom{1}{0} a^1 b^0 + \binom{1}{1} a^0 b^1 = a + b.$
(2) Assume that the statement is true for $n = l$. Thus:
$(a+b)^l = \sum_{k=0}^{l} \binom{l}{k} a^{l-k} b^k.$

(3) Multiplying both sides by $(a+b)$, we get:

$$(a+b)^l(a+b) = \sum_{k=0}^{l} \binom{l}{k} a^{l-k} b^k (a+b)$$
$$(a+b)^{l+1} = (a+b) \sum_{k=0}^{l} \binom{l}{k} a^{l-k} b^k$$

Now, by adding the similar terms on the right side, we get:

$$(a+b)^{l+1} = (a+b) \sum_{k=0}^{l} \binom{l}{k} a^{l-k} b^k$$
$$= a^{l+1} + \left(\binom{l+1}{1} a^l b\right) + \left(\binom{l}{2} + \binom{l}{1}\right) a^{l-1} b^2$$
$$+ \ldots + \left(\binom{l}{k} + \binom{l}{k-1}\right) a^{l+1-k} b^k + \ldots + b^{l+1}$$
$$\because \binom{l}{1} + 1 = \binom{l+1}{1}, \binom{l}{k} + \binom{l}{k-1} = \binom{l+1}{k}$$
$$\therefore (a+b)^{l+1} = a^{l+1} + \binom{l+1}{1} a^l b + \binom{l+1}{2} a^{l+1-2} b^2$$
$$+ \ldots + \binom{l+1}{k} a^{l+1-k} b^k + b^{l+1}$$
$$= \sum_{k=0}^{l+1} \binom{l+1}{k} a^{l+1-k} b^k$$

Since the statement is true for $l+1$ hence it is true for all n. Thus, the theorem has been proved. ⧫

The binomial theorem can be generalized such that n is any number in the following theorem (Bell, 1986; Clarke, 1966; Spiegel et al., 2009; Anderson et al., 1983; Weltner et al., 2014; Knuth, 2014; Knuth, 2005).

Theorem 10.7 (General Binomial Theorem) *If $a, b \in \mathbb{R}$, and for any n then:*
$(a+b)^n = [a(1+\frac{b}{a})]^n = a^n(1+\frac{b}{n})^n.$

Proof

By putting $x = \dfrac{b}{a}$

$\therefore (1+x)^n = 1 + nx + \dfrac{n(n-1)}{2!}x^2 + \dfrac{n(n-1)(n-2)}{3!}x^3$
$+ ... + \dfrac{n(n-1)(n-2)...(n-k+1)}{k!}x^k + ...$

The above expansion is correct in the following cases:

$$\begin{cases} (i)\ n \in \mathbb{Z}^+, \forall x, \\ (ii)\ \ |x| < 1, \forall n, \\ (iii)\ n > -1, x = 1, \\ (iv)\ n > 0, x = -1. \end{cases}$$

The condition $|x| < 1$remains that the term:
$\dfrac{n(n-1)(n-2)...(n-k+1)}{k!}x^k$ approaches zero whenever k big enough.

Now, let us start the proof.

Let $f(x) = (1+x)^n = a_0 + a_1x + a_2x^2 + ... + a_kx^k... + \infty$ (1)

$f(0) = 1$

Differentiating (1) with respect to x on both sides, we get:

$f'(x) = n(1+x)^{n-1} = a_1 + 2a_2x + 3a_3x^2 + 4a^3.... + ka_kx^{k-1} + ...(2)$

Put $x = 0,$ we get $a_1 = n$

Differentiating (2) with respect to x on both sides, we get:

$n(n-1)(1+x)^{n-2} = 2a_2 + 6a_3x + 12a_4x^2 + ... + k(k-1)a_kx^{k-2} + ...(3)$

Put $x = 0,$ we get $a_2 = \dfrac{n(n-1)}{2!}$

Differentiating (3) with respect to x on both sides, we get:

$n(n-1)(n-2)(1+x)^{n-3}$
$= 6a_3 + 24a_4x + ... + k(k-1)(k-2)a_kx^{k-3} + ...$

Put $x = 0,$ we get $a_3 = \dfrac{n(n-1)(n-2)}{3!}$

By the same method, we get:

$a_4 = \dfrac{n(n-1)(n-2)(n-3)}{4!}$, and so on

$\therefore a_k = \dfrac{n(n-1)(n-2)...(n-k+1)}{k!}$.

Putting the values of $a_0, a_1, a_2, ...a_k$ obtained in (1), we get:

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + ...$$
$$+ \frac{n(n-1)(n-2)...(n-k+1)}{k!}x^k + ... \blacklozenge$$

Example 10.16 Evaluate $\binom{-2}{7}$.

Solution:

$$\because \binom{n}{k} = \frac{n!}{(n-k)!} = n(n-1)(n-2)...(n-k+1)$$
$$\therefore \binom{-2}{7} = \frac{(-2)(-3)...(-8)}{7!}$$
$$= \frac{(-8)7!}{7!}$$
$$= -8.$$

Note: If $n \in \mathbb{Z}^+$ then $\binom{-n}{k} = (-1)^k \binom{n+k-1}{k}$.

Example 10.17 Work out $(1+x)^{-3}$.

Solution: We have to know what $\binom{-3}{k}$ gives, for various of k?

$$\because \binom{-3}{k} = (-1)^k \binom{3+k-1}{k}$$
$$= (-1)^k \binom{k+2}{k}$$
$$= (-1)^k \frac{(k+2)(k+1)}{2}$$
$$= \begin{cases} (-1)^0 \cdot 2 \cdot \frac{1}{2} = 1, \text{ if } k = 0 \\ (-1)^1 \cdot 3 \cdot \frac{2}{2} = -3, \text{ if } k = 1 \\ (-1)^2 \cdot 4 \cdot \frac{3}{2} = 6, \text{ if } k = 2 \end{cases}$$

In general, we conclude that:

$$(1+x)^{-3} = 1 - 3x + 6x^2 - ... + (-1)^k \frac{(k+2)(k+1)}{2} x^n + ...$$

Example 10.18 Write down the first four terms of the binomial expansion of $\frac{1}{1+x}$.

Solution:

$$\because (1+x)^n = 1^n + n1^{n-1}x^1 + \frac{n(n-1)}{2!}1^{n-2}x^2$$
$$+ \frac{n(n-1)(n-2)}{3!}1^{n-3}x^3 + ... + \frac{n(n-1)(n-2)...(n-k)}{k!}1^{n-k}x^k + ...$$
$$= 1 + x + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3$$
$$+ ... + \frac{n(n-1)(n-2)...(n-k)}{k!}x^k + ...$$

Since $n \notin \mathbb{Z}^+$, hence this is an infinite series, valid when $|x| < 1$

By rewriting $\dfrac{1}{1+x} = (1+x)^{-1}$ and this is an expression of the term; $(1+x)^n$, with $n = -1$

$$\therefore (1+x)^{-1} = 1 + (-1)x + \frac{(-1)(-2)}{2!}x^2 + \frac{(-1)(-2)(-3)}{3!}x^3 + ...$$
$$= 1 - x + x^2 - x^3 + ...,$$ which is an infinite series, valid when $|x| < 1$

Thus, the first four terms of the expansion are: $1 - x + x^2 - x^3$.

Example 10.19 In the binomial expansion of $(1+bx)^{-5}$, the coefficient of x is -15. Find the value of the constant b, and the coefficient of x^2.

Solution:

Since $n \in \mathbb{R}$, hence we have the expansion:

$$(1+bx)^n = 1 + n(bx) + \frac{n(n-1)}{2!}(bx)^2 + \frac{n(n-1)(n-2)}{3!}(bx)^3$$
$$+ ... + \frac{n(n-1)(n-2)...(n-k)}{k!}(bx)^k + ...$$

Since $n \notin \mathbb{Z}^+,$ hence this is an infinite series valid

, when $|bx| < 1,$ or $|x| < \dfrac{1}{|b|}.$

$$\because (1+bx)^{-5} = 1 + (-5)(bx) + \frac{(-5)(-6)}{2!}(bx)^2 + ...$$

$\therefore -5b = -15,$

$\therefore b = 3.$

To find the coefficient of x^2

, we can substitute the value of b back into the expansion to get:

$$+ \frac{(-5)(-6)}{2!}(3x)^2 + ... = +135x^2 + ...$$

Thus, the coefficient of x^2 is $135,$ and $b = 3.$

Example 10.20 Write down the first four terms of the binomial expansion of $\frac{1}{(4+3x)^2}$, stating the range of the values of x for which the expansion is valid.

Solution:

Since $n \in \mathbb{R}$, hence we have the expansion:

$(a+bx)^n = [a(1+\frac{b}{a}x)]^n = a^n((1+\frac{b}{a}x)^n$

$= a^n[1+n(\frac{b}{a}x)+\frac{n(n-1)}{2!}(\frac{b}{a}x)^2+\frac{n(n-1)(n-2)}{3!}(\frac{b}{a}x)^3+...]$

Since $n \notin \mathbb{Z}^+$, hence this is an infinite series valid

, when $\left|\frac{b}{a}x\right| < 1$, or $|x| < \left|\frac{a}{b}\right|$.

$\because \frac{1}{(4+3x)^2} = 4^{-2}(1+\frac{3}{4}x)^{-2} = \frac{1}{16}(1+\frac{3}{4}x)^{-2}$

$\therefore (1+\frac{3}{4}x)^{-2} = 1+(-2)(\frac{3}{4}x)+\frac{(-2)(-3)}{2!}(\frac{3}{4}x)^2$

$+\frac{(-2)(-3)(-4)}{3!}(\frac{3}{4}x)^3+...$

$= 1-\frac{3}{2}x+3(\frac{3}{4})^2x^2-4(\frac{3}{4})^3x^3+...$

$= 1-\frac{3}{2}x+\frac{27}{16}x^2-\frac{27}{16}x^3+...$

Thus, the first four terms of the binomial expansion of

$\frac{1}{(4+3x)^2}$ are:

$\frac{1}{16}[1-\frac{3}{2}x+\frac{27}{16}x^2-\frac{27}{16}x^3+...]$

$= \frac{1}{16}-\frac{3}{32}x+\frac{27}{256}x^2-\frac{27}{256}x^3+...$

The expansion is valid for $\left|\frac{3}{4}x\right| < 1$, or, $-\frac{4}{3} < x < \frac{4}{3}$.

Example 10.21 What is the coefficient x^4y^5 in $(xy)^9$?

Solution:

The general form of term is $\binom{9}{k}x^{9-k}y^{k}$

Putting $k = 5$ we find that the coefficient of x^4y^5 is:

$\binom{9}{5} = 126$

Example 10.22 Prove that:

$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + ... + \binom{n}{n} = 2^n.$

Solution:

$$\because (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k}b^k$$

Assume that $a = b = 1$ we get that:

$$2^n = (1+1)^n$$
$$= \sum_{k=0}^{n} \binom{n}{k} 1^{n-k}1^k$$
$$= \sum_{k=0}^{n} \binom{n}{k}$$
$$= \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + ... + \binom{n}{n}.$$

Example 10.23 Evaluate $\sqrt{5}$ by using binomial expansion.

Solution:

$$\because 5 = 4 + 1,$$
$$\therefore \sqrt{5} = \sqrt{4+1}$$
$$= 2(1 + \frac{1}{4})^{\frac{1}{2}}$$
$$= 2[1 + (\frac{1}{2})(\frac{1}{4}) + \frac{(\frac{1}{2})(\frac{-1}{2})}{2!}(\frac{1}{4})^2 + \frac{(\frac{1}{2})(\frac{-1}{2})(\frac{-3}{4})}{3!}(\frac{1}{4})^3 + ...]$$
$$= 2(1 + 0.125 - 0.0078125 + 0.00048828 + ...)$$
$$= 2(1.11767578 + ...)$$
$$\therefore \sqrt{5} \approx 2.23535156.$$

10.4 Multinomial theorem

The multinomial theorem describes how to expand a power of a sum in terms of powers of the terms in that sum. It is the generalization of the binomial theorem from binomials to multinomials (Goodman, 2001). The Multinomial Theorem provides a formula to expand a power of a sum of any number of terms. It tells us how to express the power of a sum as a sum of products of the terms, where each product has a specific coefficient (Merris, 2003).

Definition 10.6 The multinomial theorem provides an easy way to expand the power of a sum of variables. As multinomial is just another word for polynomial, this could also be called the polynomial theorem (Spiegel, 1968; Knuth, 2014).

Theorem 10.8 (Multinomial theorem) *For a positive integer k, and nonnegative integer n, the multinomial form given by;* $(x_1 + x_2 + ... + x_k)^n = \sum_{b_1+b_2+...+b_k=n} \binom{n}{b_1,b_2,...,b_k} \prod_{j=1}^{k} x_j^{b_j}$, *where* $\binom{n}{b_1,b_2,...,b_k}$ *gives as* $\binom{n}{b_1,b_2,...,b_k} = \frac{n!}{b_1!b_2!...b_k!}$.

Proof We are going to prove this theorem by mathematical induction in k.

(i)

If $k = 1$

$\therefore$ L.H.S. $= (x_1 + x_2 + ... + x_n)^n$

$= (x_1)^n$

$= x_1^n ...(1).$

Similarily,

$$\sum_{b_1+b_2+...+b_k=n} (b_1, b_2, ..., b_k) \prod_{j=1}^{k} x_j^{b_j}$$

Thus, we see, $b_1 = n$

$$\therefore \text{ R.H.S.} = \sum_{b_1=n} \binom{n}{b_1} \prod_{j=1}^{1} x_j^{b_j}$$

$$= \sum_{n} \binom{n}{n} x_1^{b_1}$$

$$= \frac{n!}{n!}(x_1)^{b_1}$$

$= x_1^n ...(2)$

From (1)&(2) L.H.S. = R.H.S.

Thus, the theorem is true for $k = 1$.

(ii) Assume that the multinomial theorem is true for $K = m$, where m is a positive integer. Thus, we have:

$$(x_1 + x_2 + ... + x_m)^n = \sum_{b_1+b_2+...+b_m=n} \binom{n}{b_1, b_2, ..., b_m} \prod_{j=1}^{m} x_j^{b_j}.$$

(iii) Now, we have to prove the validity of the theorem in case of

$k = m + 1.$

We have, L. H. S. $= (x_1 + x_2 + ... + x_m + x_{m+1})^n$

Now, let us assume that $x_m + x_{m+1}$ is a single term,

$\because b_m + b_{m+1}$ is a single term too.

Let us assume $b_m + b_{m+1} = M.$

$Thereby,$ the number of terms is

$m + 1 - 1 = m$

Thus, we can write the multinomial theorem as

$(x_1 + x_2 + ... + x_{m-1} + (x_m + x_{m+1}))^n$

$$= \sum_{b_1+b_2+...+b_{m-1}+M=n} \binom{n}{b_1, b_2, ..., b_{m-1}, M}$$

$$= \prod_{j=1}^{m-1} x_{j=1}^{b_j} \cdot (x_m + x_{m+1})^M$$

Let us write the expansion of $(x_m + x_{m+1})^M$

by using the binomial theorem:

$(x_m + x_{m+1})^M$

$$= \sum_{m+M-b_m=M} \binom{M}{b_m, M - b_m} x_m^{b_m} x_{m+1} M - b_m$$

$\because b_m + b_{m+1} = M,$

$\therefore b_{m+1} = M - b_m$

$\therefore (x_m + x_{m+1})^M$

$$= \sum_{b_m+b_{m+1}=M} \binom{M}{b_m, b_{m+1}} x_m^{b_m} x_{m+1} b_{m+1}$$

Now, substitute this value in the expansion, then we get:

$(x_1 + x_2 + ... + x_{m-1} + (x_m + x_{m+1}))^n$

$$= \sum_{b_1+b_2+...+b_{m-1}+M=n} \binom{n}{b_1, b_2, ..., b_{m-1}, M}$$

$$\cdot \prod_{j=1}^{m-1} x_j^{b_j} \cdot (x_m + x_{m+1})^M$$

$$= \sum_{b_1+b_2+...+b_{m-1}+M=n} \binom{n}{b_1, b_2, ..., b_{m-1}, M}$$

$$\cdot \prod_{j=1}^{m-1} x_j^{b_j} \cdot \sum_{b_m+b_{m+1}=M} \binom{M}{b_m, b_{m+1}} x_m^{b_m} x_{m+1}^{b_{m+1}}$$

$$\therefore (x_1 + x_2 + ... + x_{m-1} + (x_m + x_{m+1}))^n$$

$$= \sum_{b_1+b_2+...+b_{m-1}+M=n} \binom{n}{b_1, b_2, ..., b_{m-1}, M} \sum_{b_m+b_{m+1}}$$

$$= M \binom{M}{b_m, b_{m+1}} \cdot \prod_{j=1}^{n-1} x_j^{b_j} x_m^{b_m} x_{m+1}^{b_{m+1}}$$

Now, we will find the value of:

$$\sum_{b_1+b_2+...+b_{m-1}+M=n} \binom{n}{b_1, b_2, ..., b_{m-1}, M} \cdot \sum_{b_m+b_{m+1}}$$

$= M \binom{M}{b_m, b_{m+1}}$ by expanding both these sigmas.

$$\sum_{b_1+b_2+...+b_{m-1}+M=n} \binom{n}{b_1, b_2, ..., b_{m-1}, M} \sum_{b_m+b_{m+1}=M} \binom{M}{b_m, b_{m+1}}$$

$$= \sum_{b_1+b_2+...+b_{m-1}+M=n} \frac{n!}{b_1! b_2! ... b_{m-1}! M!} \cdot \sum_{b_m+b_{m+1}=M} \frac{M!}{b_m! b_{m+1}!}$$

$$= \sum_{b_1+b_2+...+b_m+b_{m+1}=n} \frac{n!}{b_1! b_2! ... b_{m-1}! M!} \cdot \frac{M!}{b_m! b_{m+1}!}$$

$$= \sum_{b_1+b_2+...+b_m+b_{m+1}=n} \frac{n!}{b_1! b_2! ... b_m! b_{m+1}!}$$

$$= \sum_{b_1+b_2+...+b_m+b_{m+1}=n} \binom{n}{b_1, b_2, ..., b_m, b_{m+1}}$$

Now, we can write as:

$$\prod_{j=1}^{m-1} x_j^{b_j} x_m^{b_m} x_{m+1}^{b_{m+1}} = x_1^{b_1} \cdot x_2^{b_2} \cdot ... x_{m-1}^{b_{m-1}} \cdot x_m^{b_m} \cdot x_{m+1}^{b_{m+1}}$$

$$= \prod_{j=1}^{m+1} x_j^{b_j}$$

Now, substituting all the values in the expansion, we get:

$(x_1 + x_2 + ... + x_{m-1} + (x_m + x_{m+1}))^n$

$$= \sum_{b_1+b_2+...+b_{m-1}+M=n} \binom{n}{b_1, b_2, ..., b_{m-1}, M} \sum_{b_m+b_{m+1}=M} \binom{M}{b_m, b_{m+1}}$$

$$\cdot \prod_{j=1}^{m-1} x_j^{b_j} x_m^{b_m} x_{m+1}^{b_{m+1}}$$

$\therefore (x_1 + x_2 + ... + x_m + x_{m+1})^n$

$$= \sum_{b_1+b_2+...+b_m+b_{m+1}=n} \binom{n}{b_1, b_2, ..., b_m, b_{m+1}} \cdot \prod_{j=1}^{m+1} x_j^{b_j}$$

Now, if we substitute $k = m + 1$

in R. H S. of the multinomial theorem, then;

$$\text{R. H. S.} = \sum_{b_1+b_2+...+b_k=n} \binom{n}{b_1, b_2, ..., b_k} \prod_{j=1}^{k} x_j^{b_j}$$

$$= \sum_{b_1+b_2+...+b_m+b_{m+1}=n} \binom{n}{b_1, b_2, ..., b_m, b_{m+1}} \prod_{j=1}^{m+1} x_j^{b_j}$$

$\therefore$ L. H. S. = R. H. S.

Thereby, the theorem is true for all $k = m + 1$

Thus, the multinomial theorem is true for all natural number k.

Hence, the theorem is proved based on (i), (ii), and (iii). ♦

Example 10.24 Find expansion of $(x + y + z + w)^4$.

Solution:

$$(x+y+z+w)^4 = \sum_{b_1+b_2+b_3+b_4=4} \binom{4}{b_1, b_2, b_3, b_4} \prod_{j=1}^{4} x_j^{b_j}$$

The values of b_j shown in Table 10.2 respectively, $j = 1, 2, 3, 4$.

Table 10.2: Values of $b_j; j = 1, 2, 3, 4$

b_1	b_2	b_3	b_4
1	1	1	1
0	0	0	2
0	0	3	0
3	0	0	1
2	1	0	1
2	0	0	2
1	2	0	1
1	0	1	2
0	1	0	3
0	2	0	2
0	1	2	1

$$(x+y+z+w)^4 = \binom{4}{1,1,1,1} xyzw + \binom{4}{0,0,0,2} w^2$$
$$+ \binom{4}{0,0,3,0} z^3 + \binom{4}{3,0,0,1} x^3 w + \binom{4}{2,1,0,1} x^2 y w^3$$
$$+ \binom{4}{2,0,0,2} x^2 w^2 + \binom{4}{1,2,0,1} x y^2 w + \binom{4}{1,0,1,2} x z w^2$$
$$+ \binom{4}{0,1,0,3} y w^3 + \binom{4}{0,2,0,2} y^2 w^2 + \binom{4}{0,1,2,1} y z^2 w$$
$$\therefore (x+y+z+w)^4 = 24xyzw + 12w^2 + 4z^3 + 4x^3 w$$
$$+ 12x^2 yw + 6x^2 zw^2 + 12xy^2 w + 6xzw^2 + 4yw^3 + 6y^2 w^2 + 12yz^2 w.$$

Example 10.25 What is a coefficient of $x^2 y^3 z$ in the expansion $(x + y + l + z)^6$?

Solution:

$$\because n = 6, b_1 = 2, b_2 = 3, b_3 = 0, b_4 = 1$$
$$\therefore \text{ the coefficient of } \binom{6}{2,3,0,1} = \frac{6!}{2!3!0!1!} = 60.$$

Example 10.26 Find the total terms in the expansion of $(x_1 + x_2 + x_3 + x_4)^5$.

Solution:

$$\because n = 5, k = 4,$$
$$\therefore \text{ The total terms is } C_{k-1}^{n+k-1} = C_{4-1}^{5+4-1} = C_3^7$$
$$\therefore C_3^7 = \frac{8!}{3!5!} = 56.$$

Hence, the number of terms of the given expansion is 56

Example 10.27 Find the coefficient of $x^2y^3z^4w$ in the expansion of $(x - y - z + w)^{10}$.

Solution:

$$\because (x - y - z + w)^{10}$$
$$= \sum_{b_1+b_2+b_3+b_4=10} \binom{10}{b_1, b_2, b_3, b_4} (x)^{b_1}(-y)^{b_2}(-z)^{b_3}(w)^{b_4},$$
$$\because b_1 = 2, b_2 = 3, b_3 = 4, b_4 = 1,$$

$\therefore$ the coefficient of $x^2y^3z^4w$ in the expansion of $(x - y - z + w)^{10}$ is

$$\frac{10!}{2!3!4!1!}(-1)^3(-1)^4 = -12600.$$

Example 10.28 Determine the coefficient of x^5 in the expansion $(2 - x + 3x^2)^6$.

Solution:

The general term in the expansion of:

$$(2-x+3x^2)^6 = \frac{6!}{r!s!t!}2^r(-x)^s(3x^2)^t; r+s+t=6$$

$$= \frac{6!}{r!s!t!}2^r(-1)^s(3)^t x^{s+2t}$$

For the coefficient of x^5, we must have:

$s+2t=5$, but, we have $r+s+t=6$

$\therefore (s+2t=5) \wedge (r+s+t=6); 0 \leq r,s,t \leq 6.$

Now, if:

$t=0 \Rightarrow r=1, s=5,$
$t=1 \Rightarrow r=2, s=3,$
$t=2 \Rightarrow r=3, s=1.$

Thus, the three containing x^5 and coefficient of:

$$x^5 = \frac{6!}{1!5!0!}2^1(-1)^5 3^0 + \frac{6!}{2!3!1!}2^2(-1)^3 3^1 + \frac{6!}{3!1!2!}2^1(-1)^1 3^2$$

$= -5052$

Thus, the coefficient of x^5 of expansion $(2-x-3x^2)^6$ is -5052.

Example 10.29 Find the coefficient of x^2y^3 in the expansion $(1+x+y)^{10}$.

Solution:

Given the expansion is $(1+x+y)^{10}$.

$$\therefore (1+x+y)^{10} = \sum_{p+q+r=10} \frac{n!}{p!q!r!}(1)^p(x)^q(y)^r$$

For the coefficient of $x^2y^3 = 1^5x^2y^3$.

$\therefore p=5, q=2, r=3.$

Thus, the coefficient of: x^2y^3

$$= \frac{10!}{5!2!3!}$$

$= 2520.$

10.5 Harmonic series

Definition 10.7 For all $n \in \mathbb{Z}^+$, then $\sum_{n=1}^{\infty} \frac{1}{n}$ is called harmonic series, and denoted by H_n (Rice, 2011; Kullman, 2001; Bernoulli, 1689; Bernoulli and Bernoulli, 1713).

10.5.1 Properties of harmonic series

In this subsection we review some properties of harmonic series. Let us consider the harmonic series:

$$\begin{aligned} H_n &= 1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n} \\ &= \sum_{k=1}^{n} \frac{1}{k} \end{aligned}$$

characterized by the following main properties:

(i) $H_n \to \infty$ whenever $n \to \infty$ (Kifowit and Stamps, 2006; Rice, 2011).

(ii) For the harmonic series $H_{2^m} \geq 1 + \frac{m}{2}, m > 0$ (Kifowit and Stamps, 2006; Roy, 2007). Knuth1

(iii) For the harmonic series:

$$\begin{aligned} H_{2^{m+1}} &= H_{2^m} + \frac{1}{2^{m+1}} + \frac{1}{2^{m+1}} + ... + \frac{1}{2^{m+1}} \\ &> H_{2^m} + \frac{1}{2^{m+1}} + \frac{1}{2^{m+1}} + ... + \frac{1}{2^{m+1}} \\ &= H_{2^m} + \frac{1}{2} \end{aligned}$$

Or, whenever adding one to m, then the right hand side will be increasing at least by half (Bressoud, 2022; Kifowit and Stamps, 2006; Knuth, 2005).

(iv) Approximate value of the harmonic series: The Approximation value of H_n given as follows;

$$H_n = \ln(n) + \gamma + \frac{1}{2^n} - \epsilon_n$$

where $\gamma \simeq 0.5772$ is the Euler–Mascheroni constant, and

$0 \leq \epsilon_n \leq \dfrac{1}{8n^2}$ which approaches zero as n goes to infinity.

(Bressoud, 2022; Boas Jr and Wrench Jr, 1971). That is, the harmonic series has the nature and properties of a logarithm (Cormen et al., 2022).

(v) The harmonic series can be generalized as follows (Selin, 2013; Smith and Mikami, 1914; Kitagawa, 2022; Lovelace, 1842):

$$H_n^{(x)} = 1 + \frac{1}{2^x} + \frac{1}{3^x} + ... + \frac{1}{n^x}$$

Assume that $1 < x \in \mathbb{R}$,

$\therefore \left|H_n^{(x)}\right| < M,$ where M is a fixed amount.

Thus the amount is known:

$$H_n^{(x)} = \frac{1}{2}\left|B_x\right| \frac{(2\pi)^x}{x!}$$

$= \zeta(x)$ Zeta function

, where, B_x are Bernoulli numbers, and $x \in \mathbb{Z}_e$.

If $x = 2,$ then $H_n^{(2)} = \dfrac{\pi^2}{6}$.

If $x = 4,$ then $H_n^{(4)} = \dfrac{\pi^4}{90}$.

10.5.2 Summation by fragmentation

Suppose we want to gathering a certain amount in the form of $\sum_k a_k b_b$. In such case, we appeal to the method of fragmenting. This method is useful when the terms $\sum_k a_k, \sum_k(b_{k+1} - b_k)$ are in their simple form (Rudin et al., 1964; Browder, 2012; Malik and Arora, 1992). Or;

$$\sum_{1 \leq k < n} (a_{k+1} - a_k)b_k = a_n b_n - a_1 b_1 - \sum_{1 \leq k < n} a_{k+1}(b_{k+1} - b_k).$$

10.6 Exercises

Solve the following questions:

Q1: In how many different ways can five people sit on five chairs?

Q2: In how many ways can the positions of department head, secretary and accountant be filled in the board of directors consisting of seven people?

Q3: How many four-digit numbers can be formed from the numbers $2, 3, 4, 5, 6$?

Q4: How many three-digit numbers can be formed from the numbers $1, 2, , 4, 5, 6, 7, 8, 9$, Provided that repetition is not permitted?

Q5: How many five-digit numbers starting with odd can be formed from the numbers $1, 2, 3, 4, 5, 6$?

Q6: Evaluate the following permutations:

(i) P_0^6.

(ii) P_2^9.

(iii) $P_2^n, 2 < n \in \mathbb{N}$.

Q7: How many arrangements can be made of four types of flowers, each consisting of two flowers?

Q8: What is the relationship between P_r^n, P_{n-r}^n?

Q9: Prove that the following relationships:

(i) $\binom{n}{r} + \binom{n}{r-1} = \binom{n+1}{r}$.

(ii) $\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} ... + (-1)^n \binom{n}{n} = 0$.

(iii) $\sum_{i=1}^{n} \binom{n}{i} = \binom{n+1}{i+1} - \binom{1}{i+1}$.

(iv) $\sum_{k \leq n} \binom{r}{k} (-1)^k \binom{-r+n}{n} = (-1)^n \binom{r-1}{n}$.

(v) $\binom{n}{m} = (-1)^{n-m} \binom{-(m+1)}{n-m}$.

(vi) $r\binom{r-1}{k} = (r-k)\binom{r}{k}, k \in \mathbb{N}$.

Q10: Find the values of $a_0, a_1, a_2, ...$ in the;
$n! = a_0 + a_1 n + a_2 n(n-1) + a_3 n(n-1)(n-2) + ..., \forall n \in \mathbb{Z}^+$.
Q11: Find the value of n in the following equalities:

(i) $\binom{n}{n-2} = 10$.

(ii) $\binom{n}{15} = \binom{n}{11}$.

Q12: How many different words, regardless of their meanings, can be formed from the word "Dilan" ? by taking it;

(i) one at a time,

(ii) three of time.

Q13: There is a social event attended by 15 men and 20 women. They want to seat 7 of the attendees at a table to the right of which there are 4 chairs and to the left of it are 3 chairs.

(i) If only women are seated on the right and only men on the left, in how many ways can the attendees sit?

(ii) If a certain woman sits in the first chair on the right, and if a certain man sits in the first chair on the left, in how many ways can the attendees sit?

Q14: John owns three types of pens, and Kennedy owns nine types of pens. Find the number of ways in which John and Kennedy can permuting pens with each other, provided that each of them keeps the original number of pens that they had with them.

Q15: Find the number of ways in which selecting seven balls, provided three balls red from a box including six red balls, and five white balls.

Q16: Evaluate $(28)\frac{1}{3} + (1.01)\frac{1}{5} + (1.01)\frac{1}{2}$.

Q17: Find the coefficient of x^3y^4 in the following expansions:

(i) $(x + 2y)^7$.

(ii) $(x^3 + y^2)^3$.

(iii) $(x - y^4)^4$.

Q18: Find the expansion of:

(i) $(x - y + 2z)^5$.

(ii) $(1 + x + y + z)^3$.

Q19: Find the coefficient of $x^2y^3z^4$ in the expansion $(x - y - z)^4$.
Q20: Prove that $m^n = \sum_{n_1+n_2+...+n_m} \binom{n}{n_1,n_2,...,n_m}$.
Q21: Prove that the number of terms of the expansion $(a_1 + a_2 + ... + a_m)^n$ is $\binom{n+m-1}{n} = \frac{(n+m-1)!}{n!(n-m)!}$.
Q22: A patient's weight decreased from $80kg$ to $60kg$ within 25 days, and if the weight after t of days is $W(t) = 160(\frac{7}{8})^{\frac{t}{25}}$. Find $W(1), W(2)$, and $W(10)$.
Q23: Prove the following equalities:

(i) $(x + y)^r = \sum_k \binom{r}{k} x(x - kz)^{k-1}(y + kz)^{r-k}, r \in \mathbb{Z}^+$.

(ii) $H_{2m} \leq 1 + m$.

(iii) $\left|H_n^{(r)}\right| < M, \forall n$.

(iv) $\sum_{1\leq k\leq n} H_k = (n + 1)H_n - n$.

Q24: Consider the finite geometric series:

$$S(x) = 1 + x + x^2 + x^3 + ... + x^n$$

can be evaluated by using $(1 - x)S(x)$. Prove that:

(i) $S(x) = \frac{1-x^{n+1}}{1-x}, x \neq 1$.

(ii) $S(1) = n + 1$.

Q25: Express of the summation:

$$1 + \frac{1}{3} + \frac{1}{5} + ... + \frac{1}{2n+1}$$

in terms of harmonic numbers.
Q26: If Beta function defined as follows:

$$\beta(x, y) = \int_0^1 t^{x-1}(1 - t)^{y-1}dt, x, y \in \mathbb{R}^+$$

Prove that:

(i) $\beta(k,1) = \beta(1,k) = \frac{1}{k}$.

(ii) $\sum_k \beta(k,1) = H_k$.

(iii) $\beta(x+1,y) + \beta(x,y+1) = \beta(x,y)$.

(iv) $\beta(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$, where $\Gamma(x) = \int_0^\infty e^{-y} y^{x-1} dy, x > 0$

(v) $\beta(\frac{1}{2}, \frac{1}{2}) = \pi$.

Bibliography

Abramowitz, M. and Stegun, I. A. (1948). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Vol. 55. US Government printing office.

Abramowitz, M., Stegun, I. A. and Romer, R. H. (1988). 'Handbook of mathematical functions with formulas, graphs, and mathematical tables'.

Adams, R. A. and Essex, C. (2018). *Calculus: a complete course*. Pearson.

Ahlfors, L. V. (1979). *Complex analysis*.

Aigner, M. and Ziegler, G. M. (1999). Proofs from the book. *Berlin. Germany*. 1: 2.

Akritas, A. G., Strzebonski, A. W. and Vigklas, P. S. (2008). Improving the performance of the continued fractions method using new bounds of positive roots. *Nonlinear Analysis: Modelling and Control*. 13(3): 265–279.

Aliabadi, M. and Darafsheh, M. (2015). On maximal and minimal linear matching property. *arXiv preprint arXiv:1508.00937*. .

Alston, S. (1959). Householder. dandelin, lobatchevskii or graeffe. *American Mathematical Monthly.* .

Anderson, J. C. et al. (1983). *Data and formulae: for engineering students*. Pergamon Press.

Anderson, R. B. (1979). *Proving programs correct*. Wiley.

Andreescu, T. and Andrica, D. (2006). *Complex Numbers from A to... Z*. Springer.

Andrilli, S. and Hecker, D. (2022). *Elementary linear algebra*. Academic Press.

Anton, H. (1987). Elementary linear algebra 9th edition.. . .

Anton, H. and Rorres, C. (2013). *Elementary linear algebra: applications version*. John Wiley & Sons.

Apostol, T. (1981). *Mathematical analysis*. Addison-Wesley.

Apostol, T. M. (1967). One-variable calculus, with an introduction to linear algebra. *(No Title)*. .

Apostol, T. M. (1974). 'Mathematical analysis addison'.

Apostol, T. M. and Ablow, C. M. (1958). Mathematical analysis. *Physics Today*. 11(7): 32.

Arfken, G. (1985). 'Mathematical methods for physicists. 3rd edn academic press'.

Arruda, M. A. Z. (2007). *Trends in sample preparation*. Nova Publishers.

Atiyah, M. (2018). *Introduction to commutative algebra*. CRC Press.

Atkinson, K. (1991). *An introduction to numerical analysis*. John wiley & sons.

Avriel, M. (1976). Nonlinear programming: Analysis and methods. printice-hall. *Inc, Englewood Cliffs, New Jersey*. .

Axelsson, O. (2007). Solution of linear systems of equations: iterative methods. In *Sparse Matrix Techniques: Copenhagen 1976 Advanced Course Held at the Technical University of Denmark Copenhagen, August 9–12, 1976*. Springer. 1–51.

Axler, S. (2010). *College Algebra*. Wiley Global Education.

Axler, S. (2015). *Linear algebra done right*. Springer.

Ayres, J. F. (1962). *Theory and problems of matrices*. McGraw-Hill.

Bak, J., Newman, D. J. and Newman, D. J. (2010). *Complex analysis*. Vol. 8. Springer.

Baker, R. J. (2019). *CMOS: circuit design, layout, and simulation*. John Wiley & Sons.

Balfour, A. and Beveridge, W. (1972). *Basic Numerical Analysis with Algol*. Heinemann Educational Publishers.

Ballantine, J. P. (1959). Complex roots of real polynomials. *The American Mathematical Monthly*. 66(5): 411–414.

Barbeau, E. J. (2003). *Polynomials*. Springer Science & Business Media.

Barnett, R. A., Ziegler, M. R., Byleen, K. E. and Stocker, C. J. (2019). *Finite mathematics for business, economics, life sciences, and social sciences*. Pearson.

Basu, S. (2021). Strictly real fundamental theorem of algebra using polynomial interlacing. *Bulletin of the Australian Mathematical Society*. 104(2): 249–255.

Bather, J. A. (1994). Mathematical induction. . .

Beachy, J. A. and Blair, W. D. (2006). Abstract algebra , waveland pr. *Inc.* .

Beals, R. and Wong, R. (2010). *Special functions: a graduate text*. Vol. 126. Cambridge University Press.

Beauregard, R. A. and Fraleigh, J. B. (1973). *A first course in linear algebra: with optional introduction to groups, rings, and fields.* Houghton Mifflin.

Bell, E. T. (1934). Exponential polynomials. *Annals of Mathematics.* 258–277.

Bell, E. T. (1986). *Men of mathematics.* Simon and Schuster.

Bellman, R. (1997). *Introduction to matrix analysis.* SIAM.

Bender, N., Yamilov, A., Yılmaz, H. and Cao, H. (2020). Fluctuations and correlations of transmission eigenchannels in diffusive media. *Physical Review Letters.* 125(16): 165901.

Berggren, J. L., Berggren, J. L. et al. (1986). *Episodes in the mathematics of medieval Islam.* Vol. 2003. Springer.

Bernoulli, J. (1689). Propositiones arithmeticae de seriebus infinitis earumque summa finita [arithmetical propositions about infinite series and their finite sums]. basel. *J. Conrad.* .

Bernoulli, J. and Bernoulli, J. (1713). *Ars conjectandi: opus posthumum: accedit Tractatus de seriebus infinitis; et Epistola gallice scripta de ludo pilae reticularis.* Impensis Thurnisiorum.

Besenyei, A. (2012). A brief history of the mean value theorem. *Talk slides.* 12.

Bessis, D., Itzykson, C. and Zuber, J. B. (1980). Quantum field theory techniques in graphical enumeration. *Advances in Applied Mathematics.* 1(2): 109–157.

Best, G. C. (1949). Notes on the graeffe method of root squaring. *The American Mathematical Monthly.* 56(2): 91–94.

Bidyāraṇya, S. and Singh, A. N. (1962). History of hindu mathematics: a source book. *(No Title).* .

Biggs, N. L. (1979). The roots of combinatorics. *Historia mathematica.* 6(2): 109–136.

Billings, S. A. (2013). *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley & Sons.

Bini, D. and Pan, V. Y. (2012). *Polynomial and matrix computations: fundamental algorithms*. Springer Science & Business Media.

Blank, B. (1999). An imaginary tale book review. *Notices of the AMS*. 46(10): 1233–1236.

Boas Jr, R. P. and Wrench Jr, J. W. (1971). Partial sums of the harmonic series. *The American Mathematical Monthly*. 78(8): 864–870.

Böhm, A. (2013). *Quantum mechanics: foundations and applications*. Springer Science & Business Media.

Borwein, P. and Erdélyi, T. (1995). *Polynomials and polynomial inequalities*. Vol. 161. Springer Science & Business Media.

Bourbaki, N. (1994). Foundations of mathematics; logic; set theory. In *Elements of the History of Mathematics*. Springer. 1–44.

Bourbaki, N. (2008). Elements of the history of mathematics paperback. *J. Meldrum (Translator)*. .

Bourbaki, N. (2013). *Algebra II: Chapters 4-7*. Springer Science & Business Media.

Brand, L. (1942). The roots of a quaternion. *The American Mathematical Monthly*. 49(8): 519–520.

Branson, W. (2013). Solving the cubic with cardano. *Convergence (September 2013). http://www. maa. org/press/periodicals/convergence/solving-the-cubic-with-cardano*. .

Bressoud, D. (2022). *A radical approach to real analysis*. Vol. 10. American Mathematical Society.

Brézin, É., Kazakov, V., Serban, D., Wiegmann, P. and Zabrodin, A. (2006). *Applications of random matrices in physics*. Vol. 221. Springer Science & Business Media.

Britannica, E. et al. (1993). 'Encyclopædia britannica'.

Britton, J. R. and Snively, L. C. (1954). *Algebra*. Holt Reinhart and Winston.

Browder, A. (2012). *Mathematical analysis: an introduction*. Springer Science & Business Media.

Brown, J. W. and Churchill, R. V. (2009). *Complex variables and applications eighth edition*. McGraw-Hill Book Company.

Brown, W. (1991). *Matrices and Vector SPates*. Vol. 145. CRC Press.

Brualdi, R. A. (2004). *Introductory combinatorics*. Pearson Education India.

Bruce, M. and Shernock, E. (2002). 'Stage based population projection matrices and their biological applications'.

Burden, R. L. and Faires, J. D. (1985). 'Numerical analysis, 2.1 the bisection algorithm'.

Cajori, F. (1911). Horner's method of approximation anticipated by ruffini. . .

Calinger, R., Brown, J. E. and West, T. R. (1999). A contextual history of mathematics: To euler. *(No Title)*. .

Callegaro, L. (2012). *Electrical impedance: principles, measurement, and applications*. CRC Press.

Callier, F. M. and Desoer, C. A. (2012). *Linear system theory*. Springer Science & Business Media.

Candido, G. (1941). Le risoluzioni della equazione di quarto grado (ferrari-eulero-lagrange). *Period. Mat.* 21(4): 88–106.

Cardano, G. (2002). *The book of my life*. New York Review of Books.

Cardano, G. and Witmer, T. R. (1993). Ars magna, or, the rules of algebra. *(No Title)*. .

Caswell, H. (2000). *Matrix population models*. Vol. 1. Sinauer Sunderland, MA.

Cauchy, A. L. B. (1828). *Exercices de mathématiques*. Vol. 3. De Bure Frères.

Cayley, A. (1858a). Ii. a memoir on the theory of matrices. *Philosophical transactions of the Royal society of London*. (148): 17–37.

Cayley, A. (1858b). Ii. a memoir on the theory of matrices. *Philosophical transactions of the Royal society of London*. (148): 17–37.

Cayley, A. (1894). *The collected mathematical papers of Arthur Cayley*. Vol. 7. University of Michigan Library.

Charles, H. C. (1892). *An elementary course in theory of quations*. Wiley & sons.

Chávez-Pichardo, M., A, M. A. M.-C., Trejo-Martínez, A., Martínez-Carbajal, D. and Arenas-Resendiz, T. (2022). A complete review of the general quartic equation with real coefficients and multiple roots. *Mathematics*. 10(14): 2377.

Chávez-Pichardo, M., Martínez-Cruz, M. A., Trejo-Martínez, A., Vega-Cruz, A. B. and Arenas-Resendiz, T. (2023). On the practicality of the analytical solutions for all third-and fourth-degree algebraic equations with real coefficients. *Mathematics*. 11(6): 1447.

Clarke, L. H. (1966). *A note book in pure mathematics*. Heinemann Educational Books.

Coburn, N. (1955). Vector and tensor analysis. *(No Title)*. .

Conrey, J. B. (2007). *Ranks of elliptic curves and random matrix theory*. number 341. Cambridge University Press.

Conte, S. D. and De Boor, C. (2017). *Elementary numerical analysis: an algorithmic approach*. SIAM.

Coolidge, J. L. (1949). The story of the binomial theorem. *The American Mathematical Monthly*. 56(3): 147–157.

Cormen, T. H., Leiserson, C. E., Rivest, R. L. and Stein, C. (2022). *Introduction to algorithms*. MIT press.

Courant, R. and Hilbert, D. (2008). *Methods of mathematical physics: partial differential equations*. John Wiley & Sons.

Cramer, G. (1750). *Introduction à l'analyse des lignes courbes*. Cramer.

Cullen, C. G. (2012). *Matrices and linear transformations*. Courier Corporation.

David, M. S. (1890). *Numbers universalized: an advanced algebra*. American Book Company.

Delmonte, C. (1997). *Advances in AFM & STM Applied to the Nucleic Acids*. Clive Delmonte Publications.

Descartes, R. (1954). '1637. the geometry of rené descartes'.

Dickenstein, A. (2005). *Solving polynomial equations*. Springer.

Diekmann, O. and Heesterbeek, J. A. P. (2000). *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. Vol. 5. John Wiley & Sons.

Diekmann, O., Heesterbeek, J. A. P. and Metz, J. A. J. (1990). On the definition and the computation of the basic reproduction ratio r 0 in models for infectious diseases in heterogeneous populations. *Journal of mathematical biology*. 28: 365–382.

Dieudonné, J. A. (1978). Abrégé d'histoire des mathématiques: 1700-1900. *(No Title)*. .

Dimo, P. (1975). Nodal analysis of power systems. . .

Djukić, D., Janković, V., Matić, I. and Petrović, N. (2011). *The IMO Compendium: A Collection of Problems Suggested for the International Mathematical Olympiads: 1959-2009 Second Edition*. Springer.

Dorf, R. C. (2003). *CRC handbook of engineering tables*. CRC Press.

Dossey, J. A., Otto, A. D., Spence, L. E. and Eynden, C. V. (1987). *Discrete mathematics*. Pearson Education India.

Dybkaer, R. (2003). Units for quantities of dimension one. *Metrologia*. 41(1): 69.

Edward, H. M. (1995). *Linear algebra*. Springer.

Evans, D. J. A. and Benn, D. I. (2014). *A practical guide to the study of glacial sediments*. Routledge.

Eves, H. W. (1980). *Elementary matrix theory*. Courier Corporation.

Faires, J. D. and Burden, R. L. (2012). *Numerical methods, 4th*. Cengage Learning.

Farrell, O. J. and Ross, B. (2013). *Solved problems in analysis: as applied to gamma, beta, legendre and bessel functions*. Courier Corporation.

Fifield, F. W. and Haines, P. J. (2000). *Environmental analytical chemistry*. Vol. 378. Blackwell science London.

Finkbeiner, D. T. (2013). *Introduction to matrices and linear transformations*. Courier Corporation.

Flanigan, F. J. (1983). *Complex variables: harmonic and analytic functions*. Courier Corporation.

Fraleigh, J. B. (2003). *A first course in abstract algebra*. Pearson Education India.

Francesco, M. (1575). *Arithmeticorum libri duo*.

Frankel, T. (2011). *The geometry of physics: an introduction*. Cambridge university press.

Fraser, D. A. S. (1958). Statistics: An introduction. . .

Fridy, J. A. (2000). *Introductory analysis: the theory of calculus*. Gulf Professional Publishing.

Froberg, C. E. (1965). Introduction to numerical analysis. . .

Fudenberg, D. and Tirole, J. (1991). *Game theory*. MIT press.

Fujiwara, M. (1916). Über die obere schranke des absoluten betrages der wurzeln einer algebraischen gleichung. *Tohoku Mathematical Journal, First Series*. 10: 167–171.

Funkhouser, H. G. (1930). A short account of the history of symmetric functions of roots of equations. *The American mathematical monthly*. 37(7): 357–365.

Fusy, É. (2005). Transversal structures on triangulations, with application to straight-line drawing. In *International Symposium on Graph Drawing*. Springer. 177–188.

Fuzzy, É. (2009). Transversal structures on triangulations: A combinatorial study and straight-line drawings. *Discrete Mathematics*. 309(7): 1870–1894.

Gantmacher, F. R. (1959). The theory of matrices vol 1 (new york: Chelsea). *Transmission zeros in linear multivariable regulators*. 681.

Gantmakher, F. R. (2000). *The theory of matrices*. Vol. 131. American Mathematical Soc.

Gehman, H. M. (1941). Complex roots of a polynomial equation. *The American Mathematical Monthly*. 48(4): 237–239.

George, M. (1967). *Topology: An Introduction with Application to Topological Groups*. McGraw-Hill.

Gill, P. E., Murray, W. and Wright, M. H. (2021). *Numerical linear algebra and optimization*. SIAM.

Gloub, G. H. and Van Loan, C. F. (1996). Matrix computations. *Johns Hopkins Universtiy Press, 3rd edtion.* .

Godsil, C. and Royle, G. F. (2001). *Algebraic graph theory*. Vol. 207. Springer Science & Business Media.

Gold, D. L. (1984). Webster's ninth new collegiate dictionary. *Dictionaries: Journal of the Dictionary Society of North America*. 6(1): 200–235.

Goldenberg, E. P. and Carter, C. J. (2017). A student asks about (–5)!. *The Mathematics Teacher*. 111(2): 104–110.

Golub, G. H. and Van de Vorst, H. A. (2000). Eigenvalue computation in the 20th century. *Journal of Computational and Applied Mathematics*. 123(1-2): 35–65.

Gong, Z., Aldeen, M. and Elsner, L. (2002). A note on a generalized cramer's rule. *Linear algebra and its applications*. 340(1-3): 253–254.

Goodman, A. W. (2001). *Algebra From A To Z-Volume 2*. World Scientific Publishing Company.

Goult, R. J. (1974). *Computational methods in linear algebra*. Stanley Thornes (Publishers) Limited.

Gove, P. B. (1963). Webster's seventh new collegiate dictionary, springfield, ma: G. & c. *Meriam Company.* .

Gowers, T., Barrow-Green, J. and Leader, I. (2010). *The Princeton companion to mathematics*. Princeton University Press.

Gradshteyn, I. S. and Ryzhik, I. M. (1980). Tables of integrals, series, and products academic press. *Inc., San Diego.* .

Graham, D. J. and Midgley, N. G. (2000). Graphical representation of particle shape using triangular diagrams: an excel spreadsheet method. *Earth Surface Processes and Landforms*. 25(13): 1473–1477.

Graham, R., Knuth, D. E. and Patashnik, O. (1994). *Concrete mathematics: a foundation for computer science*. Addison-Wesley.

Grötschel, M. (2012). *Optimization stories*.

Guenther, R. D. (1990). Modern optics john wiley & sons. *Inc., New York*. .

Gunderson, D. S. (2014). *Handbook of mathematical induction: theory and applications*. CRC Press.

Hahn, L. (1994). *Complex numbers and geometry*. Vol. 1. Cambridge University Press.

Hamadameen, A. O. (2022). *Foundations of mathematics*. Koya KOY45 Kurdistan Region - F.R. Iraq: Koya University.

Hamilton, W. R. (2008). Lectures on quaternions. *(No Title)*. .

Hamkins, J. D. (2020). *Proof and the Art of Mathematics*. MIT Press.

Hartman, G. (2011). *Fundamentals of matrix algebra*. APEX Calculus.

Hartmanis, J. and Stearns, R. E. (1965). On the computational complexity of algorithms. *Transactions of the American Mathematical Society*. 117: 285–306.

Haylock, D. and Cockburn, A. D. (2008). *Understanding mathematics for young children: A guide for foundation stage and lower primary teachers*. Sage.

Healy, M. J. R. (2000). *Matrices for statistics*. Oxford University Press.

Heath, T. L. (2013). *A history of Greek mathematics*. Vol. 1. Cambridge University Press.

Hefferon, J. (2017). *LINEAR ALGEBRA, Jim Hefferon 3rd Ed, 2017, Joshua Smcvt Edu*. Joshua Smcvt Edu.

Heinbockel, J. H. (2001). Introduction to tensor calculus and continuum mechanics trafford publishing. *Victoria*. .

Henkin, L. (1960). On mathematical induction. *The American Mathematical Monthly*. 67(4): 323–338.

Herstein, I. N. (1964). Topics. *Algebra. Waltham: Blaisdell Publishing Company*. .

Herstein, I. N. (1991). *Topics in algebra*. John Wiley & Sons.

Hirst, H. P. and Macey, W. T. (1997). Bounding the roots of polynomials. *The College Mathematics Journal*. 28(4): 292–295.

Hoffman, K. and Kunze, R. (1967). *Linear Algebra*. Printice Hall Inc.

Hohn, F. E. (1972). *Introduction to Linear Algebra*. The Macmillan Company.

Hom, R. A. and Johnson, C. R. (1985). Matrix analysis. *Cambridge University Express*. 455.

Horn, R. A. (1985). Cr johnson matrix analysis. *Cambridge UP, New York*. .

Horn, R. A. and Johnson, C. R. (2012). *Matrix analysis*. Cambridge university press.

Horner, W. G. (1819). Xxi. a new method of solving numerical equations of all orders, by continuous approximation. *Philosophical Transactions of the Royal Society of London*. (109): 308–335.

Horner, W. G. (1833). A new method of solving numerical equations of all orders, by continuous approximation. In *Abstracts of the Papers Printed in the Philosophical Transactions of the Royal Society of London*. number 2. The Royal Society London. 117–117.

Householder, A. S. (2013). *The theory of matrices in numerical analysis*. Courier Corporation.

Itō, K. (1993). *Encyclopedic dictionary of mathematics*. Vol. 1. MIT press.

Itzykson, C. and Zuber, J. B. (1980). *Introduction to Quantum Field Theory*. McGraw-Hill International Book Company.

Ivanov, A. B. (2001). 'Vector'.

Jacobi, C. G. (1846). Über ein leichtes verfahren die in der theorie der säcularstörungen vorkommenden gleichungen numerisch aufzulösen*. . .

Jacobs, H. R. and Geometry, W. H. (1974). Freeman and company. *San Francisco*. .

Jakob, B. and Waerden, B. L. (1985). Bartel l. van der waerden 1903-1996:(titel der diplomarbeit: Die briefsammlung bl van der waerden). *Handschriften und Autographen der ETH-Bibliothek*. 57.

Jones, N. (2011). *Structural impact*. Cambridge university press.

Katz, V. J. (1998a). A history of mathematics: An introduction. *(No Title)*. .

Katz, V. J. (1998b). A history of mathematics: An introduction. *(No Title)*. .

Khamsi, M. A. and Kirk, W. A. (2011). *An introduction to metric spaces and fixed point theory*. John Wiley & Sons.

Khan, S. A. (2020). Trigonometric ratios using geometric methods. *Advances in Mathematics: Scientific Journal*. 9(10): 8685–8702.

Khayyam, O. (1963). A paper of omar khayyam. *Scripta Math*. 26: 323–337.

Kifowit, S. J. and Stamps, T. A. (2006). The harmonic series diverges again and again. *To appear in The AMATYC Review*. 1.

Kitagawa, T. L. (2022). The origin of the bernoulli numbers: Mathematics in basel and edo in the early eighteenth century. *The Mathematical Intelligencer*. 44(1): 46–56.

Kline, M. (1990). *Mathematical Thought from Ancient to Modern Times: Volume 2*. Vol. 2. Oxford university press.

Kline, R. R. (2019). *Steinmetz: engineer and socialist*. Plunkett Lake Press.

Knopp, K. (1952). *Elements of the Theory of Functions*. Vol. 1. Courier Corporation.

Knox-Robinson, C. M. and Gardoll, S. J. (1998). Gis-stereoplot: an interactive stereonet plotting module for arcview 3.0 geographic information system. *Computers & Geosciences*. 24(3): 243–250.

Knuth, D. E. (2005). *The art of computer programming*. Pearson Education.

Knuth, D. E. (2014). *Art of computer programming, volume 2: Seminumerical algorithms*. Addison-Wesley Professional.

Kojima, T. (1917). On the limits of the roots of an algebraic equation. *Tohoku Mathematical Journal, First Series*. 11: 119–127.

Königsberger, K. (2004). *Analysis 1*. Springer.

Körner, T. W. (2004). *A companion to analysis: a second first and first second course in analysis*. Vol. 62. American Mathematical Soc.

Kosinski, A. A. (2001). Cramer's rule is due to cramer. *Mathematics Magazine*. 74(4): 310–312.

Kosko, E. (1957). Matrix inversion by partitioning. *Aeronautical Quarterly*. 8(2): 157–184.

Krantz, S. G., Kress, S. and Kress, R. (1999). *Handbook of complex variables*. Springer.

Kriete, H. (1998). *Progress in holomorphic dynamics*. Vol. 387. CRC Press.

Krzanowski, W. J. (2000). *Principles of multivariate analysis*. Oxford University Press.

Kudryashov, N. A. and Demina, M. V. (2007). Relations between zeros of special polynomials associated with the painlevé equations. *Physics Letters A*. 368(3-4): 227–234.

Kullman, D. E. (2001). What's harmonic about the harmonic series?. *The College Mathematics Journal*. 32(3): 201–203.

Lagrange, J. L. (1879). Traité de la résolution des équations numériques, paris (1798). *Reprinted in Œuvres*. 8: 637–649.

Lancaster, P. and Tismenetsky, M. (1985). *The theory of matrices: with applications*. Elsevier.

Landau, J. A. (2007). Historia matematica mailing list archive: Re:[hm] pascal's triangle. *Archives of Historia Matematica. Retrieved*. 04–13.

Landau, S. I. (1994). The american college dictionaries. *International Journal of Lexicography*. 7(4): 311–351.

Lang, S. (1984). 'Algebra, advanced book program'.

Lang, S. (2002). 'Graduate texts in mathematics, vol. 211 algebra,(rev'.

Lang, S. (2012a). *Calculus of several variables*. Springer Science & Business Media.

Lang, S. (2012b). *Introduction to linear algebra*. Springer Science & Business Media.

Lankham, I., Nachtergaele, B. and Schilling, A. (2006). Permutations and the determinant. . .

Latouche, G. and Ramaswami, V. (1999). *Introduction to matrix analytic methods in stochastic modeling*. SIAM.

Lay, D. C. (2003). *Linear algebra and its applications*. Pearson Education India.

Lazard, D. (2009). Thirty years of polynomial system solving, and now?. *Journal of symbolic computation*. 44(3): 222–231.

Ledermann, W. (2013a). *Complex Numbers*. Springer Science & Business Media.

Ledermann, W. (2013b). *Complex Numbers*. Springer Science & Business Media.

Leung, K. T., Suen, S. N. and Mok, I. A. C. (1992). *Polynomials and Equations: A Chinese Merchant Elite in Colonial Hong Kong (with a new preface)*. Vol. 1. Hong Kong University Press.

Li, C. and Zhang, F. (2019). Eigenvalue continuity and ger\v {s} gorin's theorem. *arXiv preprint arXiv:1912.05001*. .

Lial, M. L. (2016). *College algebra and trigonometry*. Pearson Education.

Lipschutz, S. and Lipson, M. (2001a). *Schaum's outline of theory and problems of linear algebra*. Erlangga.

Lipschutz, S. and Lipson, M. L. (2001b). *Linear Algebra 4th ed.* McGraw-Hill.

Lipschutz, S. and Lipson, M. L. (2018). *Schaum's Outline of Linear Algebra*. McGraw-Hill Education.

Liu, J. and Li, X. (2015). Hermitian-adjacency matrices and hermitian energies of mixed graphs. *Linear Algebra and its Applications*. 466: 182–207.

Lovelace, A. A. (1842). Sketch of the analytical engine invented by charles babbage, by lf menabrea, officer of the military engineers, with notes upon the memoir by the translator. *Taylor's Scientific Memoirs*. 3: 666–731.

Lozada-Cruz, G. (2020). Some applications of cauchy's mean value theorem. *Asia Pac. J. Math.* 7(30): 1–9.

MacDuffee, C. C. (1954). Theory of equations.. . .

Macedo, H. D. and Oliveira, J. N. (2013). Typing linear algebra: A biproduct-oriented approach. *Science of Computer Programming*. 78(11): 2160–2191.

Malik, S. C. and Arora, S. (1992). *Mathematical analysis*. New Age International.

Marcus, M. and Minc, H. (1992). *A survey of matrix theory and matrix inequalities*. Vol. 14. Courier Corporation.

Marden, M. (1949). *Geometry of polynomials*. number 3. American Mathematical Soc.

Marecek, L. (2017). *Intermediate algebra*. OpenStax, Rice University.

Martzloff, J. C. (2007). *A history of Chinese mathematics*. Springer.

Masotti, A. (1960). *Sui" Cartelli di matematica disfida" scambiati fra Lodovico Ferrari e Niccolo'Tartaglia*. Istituto Lombardo di Scienze e Lettere.

Mathews, J. H. (1992). *Numerical methods for mathematics, science and engineering*. Vol. 10. Prentice-Hall International.

Mazur, D. R. (2022). *Combinatorics: a guided tour*. Vol. 55. American Mathematical Society.

McCoy, N. H. (1968). *INTRODUCTION TO MODERN ALGEBRA REVISED EDITION*.

McNamee, J. M. and Pan, V. (2013). *Numerical Methods for Roots of Polynomials-Part II*. Newnes.

McNaught, A. D., Wilkinson, A. et al. (1997a). *Compendium of chemical terminology*. Vol. 1669. Blackwell Science Oxford.

McNaught, A. D., Wilkinson, A. et al. (1997b). *Compendium of chemical terminology*. Vol. 1669. Blackwell Science Oxford.

Merino, O. (2006). A short history of complex numbers. *University of Rhode Island*. .

Merris, R. (2003). *Combinatorics*. John Wiley & Sons.

Meyer, C. D. and Stewart, I. (2023). *Matrix analysis and applied linear algebra*. SIAM.

Mikami, Y. (1913). *The development of mathematics in China and Japan*. Vol. 30. Chelsea Publishing Company.

Minkowski, H. (1908). Die grundgleichungen für die elektromagnetischen vorgänge in bewegten körpern. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*. 1908: 53–111.

Minkowski, H. (1988). *Raum und zeit*. Springer.

Mohr, P. J. and Phillips, W. D. (2014). Dimensionless units in the si. *Metrologia*. 52(1): 40.

Morozov, A. et al. (2007). *Introduction to non-linear algebra*. World Scientific.

Mukhopadhyay, U. (2006). Some interesting features of hyperbolic functions. *Resonance*. 11(8): 81–85.

Nearing, J. C. et al. (2003). *Mathematical tools for physics*. Dover Publications New York.

Nering, E. D. (1970). Linear algebra and matrix theory. *(No Title)*. .

Nicholson, W. K. (2020). Linear algebra with applications. . .

Nickalls, R. W. D. (2006). Viete, descartes and the cubic equation. *The Mathematical Gazette*. 90(518): 203–208.

Nilsson, J. W. and Riedel, S. A. (2001). *Introductory circuits for electrical and computer engineering*. Pearson Education.

Norman, C. W. (1986). Undergraduate algebra: a first course. *(No Title)*. .

O'Connor, J. J. and Robertson, E. F. (1999). Abu bekr ibn muhammad ibn al-husayn al-karaji. *MacTutor History of Mathematics archive, University of St Andrews*. .

O'Connor, J. J. and Robertson, E. F. (2001). The mactutor history of mathematics archive. *University of St. Andrews, Scotland, www-gap. dcs. st-and. ac. uk/˜ history/or www-history. mcs. st-and. ac. uk/˜ history/index. html*. .

Pan, V. Y. (1997). Solving a polynomial equation: some history and recent progress. *SIAM review*. 39(2): 187–220.

Papakonstantinou, J. M. and Tapia, R. A. (2013). Origin and evolution of the secant method in one dimension. *The American Mathematical Monthly*. 120(6): 500–517.

Parsonson, S. L. (1970). Pure mathematics. *Cambridge University Press*. .

Pease, M. C. (1969). Inversion of matrices by partitioning. *Journal of the ACM (JACM)*. 16(2): 302–314.

Phillips, E. G. (2020). *Functions of a complex variable*. Courier Dover Publications.

Poincaré, H. (1913). *La dynamique de l'électron*. A. Dumas.

Prasolov, V. V. (2004). *Polynomials*. Vol. 11. Springer Science & Business Media.

Press, W. H. (1992). Root finding and nonlinear sets of equations. *Numerical Recipes in C.* .

Press, W. H., Teukolsky, S. A., Vettering, W. T. and Flannery, B. P. (1992). 'Numerical recipes, the art of scientifc programming'.

Press, W. H., Teukolsky, S. A. and Vetterling, W. T. (2007). *Numerical recipes with source code CD-ROM 3rd edition: the art of scientific computing.* Cambridge university press.

Priestley, H. A. (2003). *Introduction to complex analysis.* Oxford University Press, USA.

Pshenichnyj, B. N. (1987). The linearization method. *Optimization.* 18(2): 179–196.

Punnen, A. and Kabadi, S. (2002). Domination analysis of some heuristics for the traveling salesman problem. *Discrete Applied Mathematics.* 119(1-2): 117–128.

Rabinovitch, N. L. (1970). Rabbi levi ben gershon and the origins of mathematical induction. *Archive for History of Exact Sciences.* 237–248.

Ralston, A. (1965). A first course in numerical analysis, 1965 mcgraw-hill. *Inc. New York. MR.* 191070.

Ralston, A. and Rabinowitz, P. (2001). *A first course in numerical analysis.* Courier Corporation.

Rashed, R. (2013). *The development of Arabic mathematics: between arithmetic and algebra.* Vol. 156. Springer Science & Business Media.

Reichl, L. E. (2016). *A modern course in statistical physics.* John Wiley & Sons.

Rice, A. (2011). The harmonic series: A primer. *Mathematical Time Capsules: Historical Modules for the Mathematics Classroom.* 77.

Richard, L. (1993). 'Burden and j. douglas faires. numerical analysis'.

Riley, K. F., Hobson, M. P. and Bence, S. J. (1999). 'Mathematical methods for physics and engineering'.

Roman, S., Axler, S. and Gehring, F. W. (2005). *Advanced linear algebra*. Vol. 3. Springer.

Rotter, S. (2017). Mesoscopic scattering meets wave control. . .

Rowe, D. (1994). A history of mathematics: An introduction by victor j. katz. *Isis*. 85.

Roy, R. (2007). 'A radical approach to real analysis'.

Rudin, W. et al. (1964). *Principles of mathematical analysis*. Vol. 3. McGraw-hill New York.

Rutishauser, H. (1966). The jacobi method for real symmetric matrices. *Numerische Mathematik*. 9(1): 1–10.

Ryser, H. J. (1963). *Combinatorial mathematics*. Vol. 14. American Mathematical Soc.

Saad, Y. (2023). Revisiting the (block) jacobi subspace rotation method for the symmetric eigenvalue problem. *Numerical Algorithms*. 92(1): 917–944.

Salomon, D. (2006). *Coding for data and computer communications*. Springer Science & Business Media.

Sameh, A. H. (1971). On jacobi and jacobi-like algorithms for a parallel computer. *Mathematics of computation*. 25(115): 579–590.

Sarkar, S. K. (2008). *A Textbook of Discrete Mathematics*. S. Chand Publishing.

Schaumberger, N. (1968). Another application of de moivre's theorem. *The Mathematics Teacher*. 61(5): 496–497.

Schiff, L. I. (1968). 'Quantum mechanics.'.

Schönhage, A. (1964). Zur quadratischen konvergenz des jacobi-verfahrens. *Numerische Mathematik*. 6(1): 410–412.

Schwartz, J. T. (2001). *Introduction to matrices and vectors*. Courier Corporation.

Schwerdtfeger, H. (2020). *Geometry of complex numbers*. University of Toronto Press.

Selby, P. H. and Slavin, S. (1991). *Practical algebra: A self-teaching guide*. Vol. 110. John Wiley & Sons.

Selin, H. (2013). *Encyclopaedia of the history of science, technology, and medicine in non-westen cultures*. Springer Science & Business Media.

Sharpe, D. (1987). *Rings and factorization*. CUP Archive.

Shilov, G. (1977). 'Linear algebra (dover books on mathematics)'.

Shilov, G. E. (2012). *Linear algebra*. Courier Corporation.

Shipman, J. (2007). Improving the fundamental theorem of algebra. *The Mathematical Intelligencer*. 29(4): 9–14.

Shroff, G. M. (1990). A parallel algorithm for the eigenvalues and eigenvectors of a general complex matrix. *Numerische Mathematik*. 58(1): 779–805.

Sikka, H. (2017). Complex numbers: Its operations and properties. . .

Sitton, G. A., Burrus, C. S., Fox, J. W. and Treitel, S. (2003). Factoring very-high-degree polynomials. *IEEE Signal Processing Magazine*. 20(6): 27–42.

Slurzberg, M. and Osterheld, W. (1950). *Essentials of Electricity for Radio and Television*. McGraw-Hill.

Smith, D. E. and Mikami, Y. (1914). *A history of Japanese mathematics*. Open Court publishing company.

Sneed, E. D. and Folk, R. L. (1958). Pebbles in the lower colorado river, texas a study in particle morphogenesis. *The Journal of Geology*. 66(2): 114–150.

Ŝolín, P. (2005). *Partial differential equations and the finite element method*. John Wiley & Sons.

Souza, M. R. and Tatiana, R. M. (2018). Tempo que não é linha: uma representação da história da matemática. *VIDYA*. 38(1): 1–17.

Sparks, J. C. (2008). The pythagorean theorem. *Bloomington (Indiana): AuthorHouse*. .

Spiegel, M. R. (1968). Mathematical handbook of formulas and tables. . .

Spiegel, M. R., Lipschutz, S., Schiller, J. J. and d. Spellman (2009). *Schaum's outline of Complex Variables*. McGraw Hill Professional.

Srinivasan, S. K. and Mehata, K. M. (1976). Stochastic processes. *(No Title)*. .

Srivastava, A. N. and Sahami, M. (2009). *Text mining: Classification, clustering, and applications*. CRC press.

Ştefănescu, D. (2007). Bounds for real roots and applications to orthogonal polynomials. In *International Workshop on Computer Algebra in Scientific Computing*. Springer. 377–391.

Stewart, I. (2022). *Galois theory*. CRC press.

Stillwell, J. and Stillwell, J. (1989). *Mathematics and its History*. Vol. 3. Springer.

Strang, G. (2006). *Linear algebra and its applications*. Belmont, CA: Thomson, Brooks/Cole.

Strang, G. (2011). 18.06 sc linear algebra. *Massachusetts Institute of Technology: MIT OpenCourseWare. http://ocw. mit. edu/courses/mathematics/18-06sc-linear-algebra-fall-2011/index. htm*. .

Strang, G. (2012). 'Linear algebra and its applications 4th ed.'.

Strang, G. (2022). *Introduction to linear algebra*. SIAM.

Süli, E. and Mayers, D. F. (2003). *An introduction to numerical analysis*. Cambridge university press.

Sylvester, J. J., Cayley, A., Stokes, G. G., Ferrers, N. M. and Hermite, C. (1855). Address to the reader. *Quarterly Journal of Pure and Applied Mathematics*. 1.

Szetela, W. (1979). Hand-held calculators and the learning of trigonometric ratios. *Journal for Research in Mathematics Education*. 10(2): 111–118.

Thomas, G. B., Weir, M. D., Hass, J., Giordano, F. R. and Korkmaz, R. (2010). *Thomas' calculus*. Vol. 12. Pearson Boston.

Trefethen, L. N. and Bau, D. (2022). *Numerical linear algebra*. SIAM.

Trudinger, N. S. (1983). Elliptic partial differential equations of second order. *(No Title)*. .

Ushida, J., Tokushima, M., Shirane, M., Gomyo, A. and Yamada, H. (2003). Immittance matching for multidimensional open-system photonic crystals. *Physical Review B*. 68(15): 155115.

Uspensky, J. V. (1948). *Theory of equations*. Tata McGraw-Hill Education.

Van, A. C. M., Hablanian, M., Lerner, R. G. and Trigg, G. L. (1991). Vacuum and vacuum technology. *Encyclopaedia of Physics*. 2: 1330–1333.

van, H. M. and Monagan, M. (2004). Algorithms for polynomial gcd computation over algebraic function fields. In *Proceedings of the 2004 international symposium on Symbolic and algebraic computation*. 297–304.

Varga, R. S. (1962). Matrix iterative analysis, pretice-hall. *Inc., Englewood Cliffs*. .

Varga, R. S. (2010). *Geršgorin and his circles*. Vol. 36. Springer Science & Business Media.

Vázquez-Ávila, A. (2021). On transversal numbers of intersecting straight line systems and intersecting segment systems. *Boletín de la Sociedad Matemática Mexicana*. 27(3): 64.

Vellekoop, I. M. and Mosk, A. P. (2007). Focusing coherent light through opaque strongly scattering media. *Optics letters*. 32(16): 2309–2311.

Veselić, K. (1979). On a class of jacobi-like procedures for diagonalising arbitrary real matrices. *Numerische Mathematik*. 33: 157–172.

Vetterling, W. T., Teukolsky, S. A., Press, W. H. and Flannery, B. P. (1992). 'Root finding and nonlinear sets of equations: Numerical recipes'.

Vinberg, Ė. B. (2003). *A course in algebra*. number 56. American Mathematical Soc.

Volkov, E. A. (1990). *Numerical methods*. Taylor & Francis.

W, A. W. A., Reyment, R. and Joreskog, K. G. (1997). Applied factor analysis in the natural sciences. *Palaios*. 12(6): 613.

Wang, X. (2004). A simple proof of descartes's rule of signs. *The American Mathematical Monthly*. 111(6): 525.

Weinberg, S. (1995). *The quantum theory of fields*. Vol. 2. Cambridge university press.

Weisstein, E. W. (2002). *CRC concise encyclopedia of mathematics*. Chapman and Hall/CRC.

Weisstein, E. W. (2014). Invertible matrix theorem. *https://mathworld. wolfram. com/*. .

Wells, J. (2008). *Longman pronunciation dictionary cased for pack*. Pearson Longman.

Weltner, K., John, S., j. Weber, W., Schuster, P., Grosjean, J., Weltner, K., John, S., Weber, W. J., Schuster, P. and Grosjean, J. (2014). Taylor series and power series. *Mathematics for Physicists and Engineers: Fundamentals and Interactive Study Guide.* 229–248.

Wherrett, B. S. (1986). Group theory for atoms, molecules and solids. *(No Title).* .

Whitelaw, T. A. (2019). *Introduction to Linear Algebra.* Routledge.

Wilkinson, J. H. (1971). The algebraic eigenvalue problem. In *Handbook for Automatic Computation, Volume II, Linear Algebra.* Springer-Verlag New York.

William, H., Kemmerly, J. E. and Durbin, S. M. (2007). *Engineering circuit analysis.* McGraw-Hill Higher Education.

Wilson, W. A. and Tracey, J. I. (1925). *Analytic geometry.* DC Heath and Company.

Wright, S. J. (2006). *Numerical optimization.*

Wu, C. W. (2005). Algebraic connectivity of directed graphs. *Linear and multilinear algebra.* 53(3): 203–223.

Xirouhakis, Y., Votsis, G. and Delopoulos, A. (1999). Estimation of 3d motion and structure of human faces. In *Advances in Intelligent Systems: Concepts, Tools and Applications.* Springer. 333–344.

Yadegari, M. (1980). The binomial theorem: a widespread concept in medieval islamic mathematics. *Historia Mathematica.* 7(4): 401–406.

Ypma, T. J. (1995). Historical development of the newton–raphson method. *SIAM review.* 37(4): 531–551.

Zake Sheet, L. M. (2021). Efficiency comparison of algorithms for finding determinants via permutations.. *AL-Muthanna Journal of Pure Science.* 8(1).

Zhang, Z. Y., Li, T., Ding, C., Ren, X. W. and Zhang, X. S. (2010). Binary matrix factorization for analyzing gene expression data. *Data Mining and Knowledge Discovery*. 20: 28–52.

Zielke, G. (1985). 'Horn, ra; johnson, cr, matrix analysis. cambridge etc'.

Index

T

W

Z

Appendices

Appendix A

Proof of the Fundamental Theorem of Algebra

Before proceeding to the proof, we need the following axiom to help us with the proof.

Axiom: *In the following $p(z)$ will denote nth degree polynomial $p(z) = p_0 + p_1 z + p_2 z^2 + ... p_n z^n$, where the coefficients p_i are any complex numbers with neither p_0 nor p_n equal to zero (otherwise the polynomial equivalent to one of lesser degree). We will utilize a fundamental completeness property of real and complex numbers, namely that a continuous function on a closed set achieves its minimum at some point in the domain. This can be taken as an axiom, or can be easily proved by applying other well-known completeness axioms, such as the Cauchy sequence axiom or the nested interval axiom (Körner, 2004).*

Proof Suppose that $p(z)$ has no roots in the complex plane. First note that for large z, say $|z| > 2 \underset{i}{max} \left| \frac{p_i}{p_n} \right|$, the z^n term of $p(z)$ is greater in absolute value than the sum of all the other terms. This gives some $B > 0$, then for any sufficiently large δ, we have $|f(z)| > B, \forall z$ with $|z| \geq \delta$. We will take $B = 2\,|p(0)| = 2\,|p_0|$. Since $|p(z)|$ s continuous on the interior and boundary of the circle with radius δ, it follows by the completeness axiom mentioned above that $|p(z)|$ achieves its minimum value at some point t in this circle (possibly on the boundary). But since $|p(0)| < \frac{1}{2}.p(z), \forall z$ on the circumference of the circle, it follows that $|p(z)|$ achieves its minimum at some point t in the interior of the circle.

Now rewrite the polynomial $p(z)$, translating the argument z by t, thus producing a new polynomial:

$q(z) = p(0z + t) = q_0 + q_1 z + q_2 z^2 + q_n z^n$, and similarly translate the circle described above. Presumably the polynomial $q(z)$, defined on some circle centered at the origin (which circle is contained within the circle above), has a minimum absolute value $M > 0$ at $z = 0$. Note that $M = |p(0)| = |q_0|$.

Our proof strategy is to construct some point x, close to the origin, such that $|q(x)| < |q(0)|$, thus contradicting the presumption that $|q(z)|$ has a minimum nonzero value at $z = 0$. If our method gives us merely a direction in the complex plane for which the function value decreases in magnitude (a descent direction), then by moving a small distance in that direction, we hope to achieve our goal of constructing a complex x such that $|q(x)| < |q(0)|$. This is the strategy we will pursue.

Construction of x such that $|q(x)| < |q(0)|$:

Let the first nonzero coefficient of $q(z)$, following q_0, be q_m, so that $q(z) = q_0 + q_m z^m + q_{m+1} z^{m+1} + ... + q_n z^n$.

We will choose x to be the complex number $x = r(\frac{-q_0}{q_m})^{\frac{1}{m}}$, where r is a small positive real value we will specify below, and where $(\frac{-q_0}{q_m})^{\frac{1}{m}}$ denotes any of the mth roots of $(\frac{-q_0}{q_m})$.

Comment: As an aside, note that unlike the real numbers, in the complex number system the mth roots of a real or complex number are always guaranteed to exist: if $z = z_1 + iz_2$, with z_1 and z_2 real, then the mth roots of z are given explicitly by

$\left\{R^{\frac{1}{2}} cos(\frac{\phi+2k\pi}{m}) + iR^{\frac{1}{2}} sin(\frac{\phi+2k\pi}{m}), k = 0, 1, m-1, ..., n\right\}$, where $R = \sqrt{z_1^2 + z_2^2}, \phi = arctan(\frac{z_2}{z_1})$. The guaranteed existence of mth roots, a feature of the complex number system, is the key fact behind the fundamental theorem of algebra.

Proof that $|q(x)| < |q(0)|$: With the definition of x given above, we can write

$q(x) = q_0 - q_0 r^m + q_{m+1} r^{m+1} (\frac{-q_0}{q_m})^{\frac{m+1}{m}} + ... + q_n r^n (\frac{-q_0}{q_m})^{\frac{n}{m}}$
$= q_0 - q_0 r^m + E,$

where the extra terms E can be bounded as follows. Assume that $q_0 \leq q_m$ (a very similar expression is obtained for $|E|$ in the case $q_0 \geq q_m$), and define $\delta = (\left|\frac{q_0}{q_m}\right|)^{\frac{1}{m}}$. Then, by applying the well-known formula for the sum of a geometric series, we can write

$|E| \leq r^{m+1} \underset{i}{max} |q_i| \left|\frac{q_0}{q_m}\right|^{\frac{m+1}{m}} (1 + \delta + \delta^2 + ... + \delta^{n-m-1}) \leq \frac{r^{m+1} \underset{i}{max} |q_i|}{1-\delta} \left|\frac{q_0}{q_m}\right|^{\frac{m+1}{m}}.$

Thus $|E|$ can be made arbitrarily smaller than $|q_0 r^m| = |q_0| r^m$ by choosing r small enough. For instance, select r so that $|E| < \left|\frac{q_0 r^m}{2}\right|$. Then for such an r, we have

$|q(x)| = |q_0 - q_0 r^m + E| < \left|q_0 - \frac{q_r^m}{2}\right| = |q_0| (1 - \frac{r^m}{2}) < |q_0| = |q(0)|$, which contradicts the original assumption that $|q(z)|$ has a minimum nonzero value at $z = 0$. ⧫

Appendix B

Proof of Taylor's Theorem

Some reflection on the proof(s) of Taylor's theorem. In the beginning, we recall the mathematical form of the theorem:

Taylor's theorem: *Suppose* $f : (a,b) \to \mathbb{R}$ *is a function on* (a,b)*, where* $a, b \in \mathbb{R}$ *with* $a < b$*. Assume that for some positive integer* n*,* f *is* n*-times differentiable on the open interval* (a,b)*, and that* $f, f^{'}, f^{''}, ..., f^{(n-1)}$ *all extended continuously to the closed interval* $[a,b]$ *(the extended functions will still be called* $f, f^{'}, f^{''}, ..., f^{(n-1)}$ *respectively). Then, there exists* $c \in (a,b)$ *such that*

$$f(b) = \sum_{k=0}^{n-1} \frac{f^k(a)}{k!}(b-a)^k + \frac{f^n(a)}{n!}(b-a)^n$$

It should be noted that, when $n = 1$, this reduces to the ordinary mean value theorem. This suggests that we may modify the proof of the mean value theorem, to give a proof of Taylor's theorem (Besenyei, 2012).

The proof of the mean value theorem comes in two parts:

(i) By subtracting a linear polynomial (degree one), we we reduce the case where $f(a) = f(b) = 0$.

(ii) The special case where $f(a) = f(b) = 0$ from Rolle's theorem.

In the proof of the therm, we follow the strategy to generalization of Roll's therm by state the following proposition.

Proposition: *Suppose $f : (a, b) \to \mathbb{R}$ is a function on (a, b), where $a, b \in \mathbb{R}$ with $a < b$. Assume that for some positive integer n, F is n-times differentiable on the open interval (a, b), and that $F, F^{'}, F^{''}, ..., F^{(n-1)}$ all extended continuously to the closed interval $[a, b]$(the extended functions will still be called $F, F^{'}, F^{''}, ..., F^{(n-1)}$ respectively). In addition*

$$F(a) = F^{'}(a) = ... = F^{(n-1)} = 0, \; F(b) = 0$$

then there exists $c \in (a, b)$ such that,

$$F^{(n)}(c) = 0$$

Proof The proof of this proposition follows readily from an n- fold application of Rolle's theorem:

Since $F(a) = F(b) = 0$, by Rolle's theorem applied to F on $[a, b]$, there exists $c_1 \in (a, b)$ such that

$$F^{'}(c_1) = 0.$$

Next, since $F^{'}(a) = F^{'}(c_1) = 0$, by Roll's theorem applied to F on $[a, c_1]$, there exists $c_2 \in (a, c_1)$ such that

$$F^{'}(c_2) = 0.$$

Repeat, then we get $c_1, ..., c_n$ such that

$$a < c_n < c_{n-1} < ... < c_1 < b,$$

with

$$F^{(k)}(c_k) = 0, \forall k = 1, 2, ..., n$$

By, setting $c = c_n$, we have $c \in (a, b)$, and

$$F^{(n)}(c) = 0. \blacklozenge$$

Proof Here we can prove this theorem in two different ways, and we call the first method the first proof, while we call the second the alternative proof.

First proof of Taylor's Theorem.

We apply the proposition to prove the theorem. The key is to construct a degree n polynomial, that allows us to reduce the case in Proposition. The fact that such polynomial exists follows by a dimension counting argument in linear algebra. But we will need the explicit expression of the polynomial. So let us construct the polynomial explicitly:

Let, f be as in Taylor's theorem. Let

$$P(x) = \sum_{k=0}^{n} a_k (x-a)^k.$$

This is convenient from of expressing a polynomial of degree k, since we will need to compute higher order derivatives of this polynomial at the point a. We will find coefficients $a_0, a_1, ..., a_n$, such that $F(x) := f(x) - P(x)$ satisfies the conditions of the Proposition. Indeed, for $k = 0, 1, ..., n-1$, we have

$$F^{(k)}(a) = f^{(k)}(a) - k! a_k,$$

so in order for $F(a) = F^{'}(a) = ... = F^{(n-1)}(a) = 0$, it suffices to set

$$a_k = \frac{f^{(k)}}{k!}, \forall k = 0, 1, ..., n-1.$$

It remains then to determine a_n. But this is determined by the equation $F(b) = 0$. Indeed

$$\begin{aligned} F(b) &= f(b) - \sum_{k=0}^{n} a_k (b-a)^k \\ &= f(b) - \sum_{k}^{n-1} \frac{f^{(k)}(a)}{k!}(b-a)^k - a_n(b-a)^n, \end{aligned}$$

Setting $F(b) = 0$, we get

$$a_n = \frac{1}{(b-a)^n}\left(f(b) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(b-a)^k \right).$$

Now, we have found a polynomial P such that $F(x) := f(x) - P(x)$ satisfies the conditions of the Proposition. Hence there exists $c \in (a,b)$ such that $F^{(n)}(c) = 0$. But

$$\begin{aligned} F^{(n)}(c) &= f^{(n)}(c) - P^{(n)}(c) \\ &= f^{(n)}(c) - n! a_n \\ &= f^{(n)}(c) - \frac{n!}{(b-a)^n} \sum_{k}^{n-1} \frac{f^{(k)}(a)}{k!}(b-a)^k \end{aligned}$$

Since $F^{(n)}(c) = 0$, it follows that

$$0 = f^{(n)}(c) - \frac{n!}{(b-a)^n}\left(f(b) - \sum_{k}^{n-1} \frac{f^{(k)}(a)}{k!}(b-a)^k \right)$$

It means

$$f(b) = \sum_{k}^{n-1} \frac{f^{(k)}(a)}{k!}(b-a)^k + \frac{f^{(n)}(c)}{n!}(b-a)^n. \blacklozenge$$

Second proof of Taylor's Theorem: The alternative proof
Let f be as in Taylor's theorem, and

$$F(x) = f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

Also,

$$G(x) = (x-a)^n.$$

Then both F and G vanishes to order $(n-1)$ at a, in the sense that $F, F^{'}, ..., F^{(n-1)}$ and $G, G^{'}, ..., G^{(n-1)}$ are extends continuously to $[a, b]$, and the extends functions satisfy

$$F(a) = F'(a) = ... = F^{(n-1)}(a) = 0$$
$$G(a) = G'(a) = ... = G^{(n-1)}(a) = 0$$

Note also that $G^{''}, G^{'''}, ..., G^{(n)}$ all never vanishes on (a, b). Hence we may apply Cauchy's mean value theorem (Lozada-Cruz, 2020) n times, the first time we obtain

$$\frac{F(b)}{G(b)} = \frac{F^{'}(b) - F^{'}(a)}{G^{'}(b) - G^{'}(a)} = \frac{F^{'}(c_1)}{G^{'}(c_1)}$$

for some $c_1 \in (a, b)$.

Next, we can repeat this argument, on the interval $[a, c_1]$ instead of $[a, b]$, we then obtain

$$\frac{F^{'}(c_1)}{G^{'}(c_1)} = \frac{F^{'}(c_1) - F^{'}(a)}{G^{'}(c_1) - G^{'}(a)} = \frac{F^{''}(c_1)}{G^{''}(c_1)}$$

for some $c_2 \in (a, c_1)$.

Repeating this procedure, we obtain $c_1, c_2, ..., c_n$ such that

$$a < c_n < c_{n-1} < ... < c_1 < b$$

with

$$\frac{F(b)}{G(b)} = \frac{F^{'}(c_1)}{G^{'}(c_1)} = \frac{F^{''}(c_2)}{G^{''}(c_2)} = ... = \frac{F^{(n)}(c_n)}{G^{(n)}(c_n)}$$

Setting $c = c_n$, we have $c \in (a, b)$, and

$$\frac{F(b)}{G(b)} = \frac{F^{(n)}(c)}{G^{(n)}(c)}$$

This is equivalent to saying that

$$\frac{f(b) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(b-a)^k}{(b-a)^n} = \frac{f^{(n)}(c)}{n!}$$

Now, rearranging the term yields

$$f(b) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(b-a)^k + \frac{f^{(n)}(c)}{n!}(b-a)^n. \blacklozenge$$