# MODULE 2 SOLUTION OF LINEAR ALGEBRAIC EQUATIONS

## UNIT 1 DIRECT METHOD

### CONTENTS

**Notations and Symbols**

| | |
|---|---|
| $A = [a_{ik}]$ | Matrix with the elements $a_{ik}$ |
| det A = \|A\| | Determinant of a square matrix A |
| ¥ | infinity |
| r | Rho |
| u | Nu |
| m | Mu |
| l | Lambda |
| \|\|A\|\| | Norm of a matrix A |
| i | Imaginary unit, $i^2 = -1$. |

Also see the list given in Block 1.

## 1.0 INTRODUCTION

One of the commonly occurring problems in applied mathematics is finding one or more roots of an equation f(x) = 0. In most cases explicit solutions are not available and we are satisfied with being able to find one or more roots to a specified degree of accuracy. In Block 1, we have discussed various numerical methods for finding the roots of an equation f(x) = 0. there we have also discussed the convergence of these methods. Another important problem of applied mathematics is to find the solution of systems of linear equations arise in a large number of areas, both directly in modeling physical situations and indirectly in the numerical solution of other mathematical models. These applications occur in all areas of the physical, biological and engineering sciences. For instance, in physics, the problem of steady state temperature in a plate is reduced to solving linear equations.

Engineering problems such as determining the potential in certain electrical networks, stresses in a building frame, flow rates in a hydraulic system etc. are all reduced to solving a set of algebraic equations simultaneously. Linear algebraic systems are also involved in the optimization theory, least squares fitting of data, numerical solution of boundary value problems for ordinary and partial differential equations, statistical inference etc. Hence, the numerical solution of systems linear algebraic equations plays a very important role.

Numerical methods for solving linear algebraic systems may be divided into two types, direct and iterative. Direct methods are those which, in the absence of round-off or other errors, yield the exact solution in a finite number of elementary arithmetic operations. Iterative methods start with an initial approximation.

To understand the numerical methods for solving linear system of equations it is necessary to have some knowledge of the properties of matrices. You might have already studied matrices, determinants and their properties in your linear algebra courses. However, we begin with a quick recall of few definitions here. In this unit, we have also discussed some direct methods for finding the solution of system of linear algebraic equations.

## 2.0 OBJECTIVES

At the end of this unit, you should be able to:

- state the difference between the direct and iterative methods of solving the system of linear algebraic equations
- obtain the solution of system of linear algebraic equations by using the direct method
- use the pivoting technique while transforming the coefficient matrix to upper or lower triangular matrix.

## 3.0 MAIN CONTENTS

### 3.1 Preliminaries

As we have mentioned earlier, you might be already familiar with vectors, matrices, determinants and their properties (Ref. Linear algebra MTE-02). A rectangular array of (real or complex) numbers of the from

$$\begin{bmatrix} \\ a_{11} & a_{12}\ldots\ldots & a_{2n} \\ a_{21} & a_{22}\ldots\ldots & a_{2n} \\ \\ a_{n1} & a_{n2}\ldots\ldots\ldots & a_{nn} \end{bmatrix}$$

is called a matrix. The numbers $a_{11}$, $a_{12}$, ..., $a_{nn}$ are the elements of the matrix. The horizontal lines are called rows and the vertical lines called columns of the matrix. A matrix with m rows and n columns is called an m´ n matrix (read as m by n matrix). We usually denote matrices by capital letters A, b etc., or by ($a_{jk}$), ($b_{ik}$) etc.

If the matrix has the same number of rows and columns, we call it a square matrix and the number of rows or columns is called its order. If a matrix has only one column it is a column matrix or column vector and if it has only one row it is a row matrix or row vector.

The matrices $A = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} = [a_{11}, a_{21}, \ldots a_{n1}]^T$ and

B = [$a_{11}$, $a_{12}$, ..., $a_{1n}$] are respectively the column and row matrices. We give below some special square matrices A = ($a_{ij}$) of order n.

1 A matrix A = ($a_{ij}$) in which $a_{ij}$ = 0 (i, j = 1, 2 ....., n) is called a null matrix and is denoted by 0.
e.g.,
$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ is a 2 ´ 2 null matrix.

2. A matrix A in which all the non-diagonal elements vanish i.e., $a_{ij}$ = 0 for i $^1$ j is called a diagonal matrix.

$$\text{E.g., } A = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}$$

is a 3 ′ 3 diagonal matrix.

3 The identity matrix I is a diagonal matrix in which all the diagonal elements are equal to one. The identity matrix of order 4 is

$$I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

4 A square matrix is lower triangular if all the elements above the main diagonal vanish i.e., $a_{ij} = 0$ for $j > i$. A lower triangular matrix of order 3 has the form

$$A = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Similarly upper triangular matrices are matrices in which,
$a_{ij} = 0$ for $i > j$.

$$\text{e.g., } A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$$

Two matrices $A = (a_{ij})$ and $B = (b_{ij})$ are equal iff they have the same number of rows and columns and their corresponding elements are equal, that is $a_{ij} = b_{ij}$ for all i, j.

You must also be familiar with the addition and multiplication of matrices.

Addition of matrices is defined only for matrices of same order. The sum C = A + B of two matrices A and B, is obtained by adding the corresponding elements of A and B, i.e., $c_{ij} = a_{ij} + b_{ij}$.

For example, if $A = \begin{bmatrix} 4 & 6 & 3 \\ 0 & 1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & -1 & 0 \\ 3 & 1 & 0 \end{bmatrix}$ then

$$A + B = \begin{bmatrix} 1 & 5 & 3 \\ 3 & 2 & 2 \end{bmatrix}$$

Product of an m ´ n matrix A = ($a_{ij}$) and an n ´ p matrix B = ($b_{ij}$) is an m ´ p matrix C. C = AB, whose (i, k)th entry is

$$c_{ij} = \sum_{j=1}^{0} a_{ij}b_{ij} = a_{ij}\, b_{ij} + a_{i2}b_{i2} + ... + a_{in}\, b_{nk}$$

That is, to obtain the (i, k)th element of AB, take the ith row of A and kth column of B, multiply their corresponding elements and add up all these products. For example, if

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 1 & 0 & 2 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \text{ then (1, 2)}$$

the elementof AB is

$$[2\ \ 3\ \ -1] \begin{bmatrix} 1 \\ 4 \\ 2 \\ \ \end{bmatrix} = 2 * 1 + 3 * 4 + (-1) *2 = 12$$

Note that two matrices A and B can be multiplied only if the number of columns of A equals the number of rows of B. In the above example the product BA is not defined.

The matrix obtained by interchanging the rows and columns of A is called the transpose of A and is denoted by $A^T$

$$\text{If } A = \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} \text{ then } A^T = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$$

Determinant is a number associated with square matrices.

$$\text{For a 2 ´ 2 matrix } A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\det (A) = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\text{For a 3 ´ 3 matrix } A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\det(A) = a_{11}\det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12}\det\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13}\det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

A determinant can be expanded about any row or column. The determinant of an n ´ n matrix A = ($a_{ij}$) is given by det(A) = $(-1)^{i+1}a_{ij}\det(A_{ij}) + (-1)^{i+2}a_{i2}\det(A_{i2}) + \ldots + (-1)^{i+n}a_{in}\det(A_{in})$, where the determinant is expanded about the ith row and $A_{ij}$ is the (n – 1) ´ (n – 1) matrix obtained from A by deleting the ith row and jth column and i £ i£ n. Obviously, computation is simple if det(A) is expanded along a row or column that has maximum number of zeros. This reduces the number of terms to be computed.

The following example will help you to get used to calculating determinants.

**Example 1:**

If A = $\begin{bmatrix} 1 & 2 & 6 \\ 5 & 4 & 1 \\ 7 & 3 & 2 \end{bmatrix}$ calculated det (A).

**Solution:** Let us expand by the first row. We have

$|A_{11}| = \begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix} = 4 * 2 – 1 * 3 = 5$, $|A_{12}| = \begin{bmatrix} 5 & 1 \\ 7 & 2 \end{bmatrix} = 5 * 2 = 7 * 1 = 3$,

$|A_{13}| = \begin{bmatrix} 5 & 4 \\ 7 & 3 \end{bmatrix} = 5 * 3 – 4 * 7 = -13$.

Thus,

$|A| = (-1)^{1+1}*1*|A_{11}|+(-1)^{1+2}*2*|A_{12}|+(-1)^{1+3}*6*|A_{13}|=5–6–78 = -79$

If the determinant of a square matrix A has the value zero, then the matrix A is called a singular matrix, otherwise, A is called a nonsingular matrix.

We shall now give some more definitions.

**Definition**: The inverse of an n ´ nnonsingular matrix A is an n ´ n matrix B having the property

A B = B A = i
where I is an identity matrix of order n ´ n.

the inverse matrix B if it exists, is denoted by $A^{-1}$ and is unique.

**Definition**: For a matrix A = ($a_{ij}$), the cofactor $A_{ij}$ of the element $a_{ij}$ is given by

$A_{ij} = (-1)^{i+j}M_{ij}$

where$M_{ij}$ (minor) is the determinant of the matrix of order (n – 1) ´ (n – 1) obtained from A after deleting its ith row and the jth column.

**Definition**: The matrix of cofactors associated with the n ´ n matrix A is an n ´ n matrix $A^c$ obtained from A by replacing each element of A by its cofactor.

**Definition**: The transpose of the cofactor matrix $A^c$ of A is called the adjoint of A and is written as adj(A). Thus

$\text{adj}(A) = (A^c)^T$

Let us now consider a system of n linear algebraic equations in n unknowns

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + .... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + .... + a_{2n}x_n = b_2 \\ \quad . \qquad\quad . \qquad . \\ \quad . \qquad\quad . \qquad . \\ \quad . \qquad\quad . \qquad . \\ a_{n1}x_1 + a_{n2}x_2 + .... + a_{nn}x_n = b_n \end{array} \qquad (1)$$

where the coefficients $a_{ij}$ and the constant $b_i$ (i = 1, ...., n) are real and known. This system of equations in matrix from may be written as

$$A\,x = b \qquad (2)$$

Where

$$A = \begin{bmatrix} a_{11} & a_{12} & & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ & & & \\ & & & \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \\ \\ x_n \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \\ \\ b_n \end{bmatrix}$$

A is called the coefficient matrix and has real elements.

Our problem is to find the values $x_i$, i = 1, 2 ...., n if they exist, satisfying Eqn. (2). Before we discuss some methods of solving the system (2), we give the following definitions.

**Definition**: A system of linear Eqns. (2) is said to be consistent if it has at least one solution. If no solution exists, then the system is said to be inconsistent.

**Definition**: The system of Eqns. (2) is said to be homogeneous if b = 0, that is, all the elements $b_i$, $b_2$, ...., $b_n$ are zero, otherwise the system is called non-homogeneous.

In this unit, we shall consider only non-homogeneous systems.

You also know from you linear algebra that the non-homogeneous system of Eqns. (2) has a unique solution, if the matrix A is nonsingular. You may recall the following basic theorem on the solvability of linear systems (Ref. Theorem 4, Sec. 5.0, Unit 1, Block 3, Module 1).

**Theorem 1**: A non-homogeneous system of n linear equations in n known has a unique solution if and only if the coefficient matrix A is nonsingular.

If A is nonsingular, then$A^{-1}$ exists, and the solution of system (2) can be expressed as

$x = A^{-1}b$.

In case the matrix A is singular, then the system (2) has no solution if b ¹ 0 or has an infinite number of solutions if b = 0. here we assume that A is a nonsingular matrix.

As we have already mentioned in the introduction, the methods of solution of the system (2) may be classified into two types:

i Direct Methods: which in the absence of round-off errors give the exact solution in a finite number of steps.

ii. Iterative Methods: Starting with an approximate solution vector $x^{(0)}$, these methods generates a sequence of approximate solution vectors $\{x^{(k)}\}$ which converge to the exact solution vector x as the number of iterations k ® ¥ . Thus iterative methods are infinite processes. Since we perform only a finite number of iterations, these methods can only find some approximation to the solution vector x. We shall discuss iterative methods later in Units 4 and 5.

In this unit we shall discuss only the direct methods. You are familiar with one such method due to the mathematician Cramer and known as Cramer's Rule. Let us briefly review it.

## 3.2 Cramer's Rule

In the system (2), let d = det(A) ¹ 0 and b ¹ 0. Then the solution of the system is obtained as

$x_i = d_i/d$, i = 1, 2, ...., n (3)

where $d_i$ is the determinant of the matrix obtained from A by replacing the ith column of A by the column vector b. let us illustrate the method through an example.

**Example 2**: Solve the system of equations.

$3x_1 + x_2 + 2x_3 = 3$
$2x_1 - 3x_2 - x_3 = -3$
$x_1 - 2x_2 - x_3 = 4$
using Cramer's rule.

**Solution**: We have,

$$d = |A| = \begin{vmatrix} 3 & 1 & 2 \\ 1 & -3 & -1 \\ 1 & 2 & 1 \end{vmatrix} = 8$$

$$d_1 = \begin{vmatrix} 3 & 1 & 2 \\ -3 & -3 & -1 \\ 4 & 2 & 1 \end{vmatrix}$$

= 8 (first column in A is replaced by the column vector b)

$$d_2 = \begin{vmatrix} 3 & 3 & 2 \\ 2 & -3 & -1 \\ 1 & 4 & 1 \end{vmatrix}$$

= 16 (second column in A is replaced by the column vector b

$$d_3 = \begin{vmatrix} 3 & 1 & 3 \\ 2 & -3 & -3 \\ 1 & 2 & 4 \end{vmatrix}$$

= -8 (third column in A is replaced by the column vector b)

Using (3), we get the solution
$x_1 = d_1/d = 1$; $x_2 = d_2/d = 2$; $x_3 = d_3/d = -1$

While going through the example and attempting the self assessment exercises you must have observed that in Cramer's methods we need to evaluate n + 1 determinants each of order n, where n is the number of equations. If the number of operations required to evaluate a determinant is measured in terms of multiplications only, then to evaluate a determinant of second order, i.e.,

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}\, a_{22} - a_{12}\, a_{21}$$

we need two multiplications or (2 – 1) 2! multiplications. To evaluate a determinant of third order

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} =(a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31})$$

we need 12 multiplication or (3 – 1)3! multiplications. In general, to evaluate a determinant of nth order we need (n – 1)n! multiplications.

Also for a system of n equations, Cramer's rule requires n + 1determinants each of order n and performs n divisions to obtain $x_i$, i = 1, 2, ...., n. Thus the total number of multiplications and divisions needed to solve a system of n equations, using Cramer's rule becomes

M = total number of multiplications + total number of divisions
= (n + 1) (n - 1)n! + n

In Table 1, we have given the values of M for different values of n.

**Table 1**

| Number of equations<br>N | Number of operations<br>n |
|---|---|
| 2 | 8 |
| 3 | 51 |
| 4 | 364 |
| 5 | 2885 |
| 6 | 25206 |
| 7 | 241927 |
| 8 | 2540168 |
| 9 | 29030409 |
| 10 | 359251210 |

From the table, you will observe that as n increases, the number of operations required for Cramer's rule increases very rapidly. For this reason, Cramer's rule is not generally used for n > 4. hence for solving large systems, we need more efficient methods. In the next section we describe some direct methods which depend on the form of the coefficient matrix.

## 3.3 Direct Methods for Special Matrices

We now discuss three special forms of matrix A in Eqn. (2) for which the solution vector x can be obtained directly.

**Case 1**: A = D, where D is diagonal matrix. In this case the systems of Eqns. (2) are of the form

$$\begin{array}{llllll}
a_{11}x_1 & \ldots\ldots\ldots\ldots\ldots\ldots\ldots & & = b_1 \\
. & a_{22}x_2 & & . = b_2 \\
. & . & & . = & . \\
. & & . & . & . \\
. & & & . & . & . \\
. & & & a_{nn}x_n = b_n
\end{array}$$

and det (A) – $a_{11}$ $a_{22}$ .... $a_{nn}$

Since the matrix A is nonsingular, $a_{11}{}^{1}$ 0 for 1, 2, ....., n and we obtain the solution as

$x_i = b_i/a_{ii}$, i = 1, 2, ...., n.

Note that in this case we need only n divisions to obtain the solution vector.

**Case 2** : A = L, where L is a lower triangular matrix ($a_{ij}$ = 0, j > i). The system of Eqns. (2) is now of the form

$$\begin{array}{ll}
a_{11}x_1 & = b_1 \\
a_{21}x_1 + a_{22}x_2 & = b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 & = b_3 \\
. & \\
. & \\
. & \\
a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \ldots + a_{nn}x_n & = b_n
\end{array} \qquad (4)$$

and det (A) = $a_{11}a_{22}...a_{nn}$.

You may notice here that the first equation of the system (4) contains only $x_1$, the second equation contains only $x_1$ and $x_2$ and so on. Hence, we find $x_1$ from the first equation, $x_2$ from the second equation and proceed in that order till we get $x_n$ from the last equation.

Since the coefficient matrix A is nonsingular, $a_{11}{}^{1}$ 0, i = 1, 2, ..., n. we thus obtain

$$\begin{aligned}
x_1 &= b_1/a_{11} \\
x_2 &= (b_2 - a_{21}x_1)/a_{22} \\
x_3 &= (b_3 - a_{31}x_1 - a_{32}x_2)/a_{33} \\
&. \\
&. \\
&. \\
x_n &= (b_n - \sum_{j=1}^{n-1} a_{ij}x_j)/a_{nn}
\end{aligned}$$

In general, we have for any i

$$x_i = (b_i - \sum_{j=1}^{n-1}(a_{ij}x_j)/a_{ii} \quad i = 1, 2, ...., n. \qquad (5)$$

For example, consider the system of equations

$5x_1 = 5$
$-x_1 - 2x_2 = -7$
$-x_1 + 3x_2 + 2x_3 = 5$

From the first equation we have,
$x_1 = 1$

From the second equation we get,
$$x_2 = \frac{-7 + x_1}{-2} = 3$$
and from the third equation we have,
$$x_3 = \frac{5 + x_1 - 3x_2}{2} = -\frac{3}{2}.$$

Since the unknowns in this methods are obtained in the order $x_1, x_2, ...., x_n$, this method is called the forward substitution method.

The total number of multiplications and divisions needed to obtain the complete solution vector x, using this method is

$M = 1 + 2 + ..... + n = n(n + 1)/2$.

**Case 3**: A = U, where U is an upper triangular matrix ($a_{ij} = 0, j < 1$). The system (2) is now of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1n}x_n &= b_1 \\ a_{22}x_2 + a_{23}x_3 + ... + a_{2n}x_n &= b_2 \\ a_{33}x_3 + ... + a_{3n}x_n &= b_3 \end{aligned} \qquad (6)$$

$a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n = b_{n-1}$
$a_{nn}x_n = b_n$
and det (A) = $a_{11}a_{22}...a_{nn}$.

You may notice here that the nth (last) equation contains only $x_n$, the (n – 1)th equation contains $x_n$ and $x_{n-1}$ and so on. We can obtain $x_n$ from the nth equation, $x_{n-1}$ from the (n – 1)th equation and proceed in that order till we get $x_1$ from the first equation. Since the coefficient matrix A is nonsingular, $a_{ii}{}^1$ 0, i = 1, 2, ...., n and we obtain

$x_n = b_n/a_{nn}$

$x_{n-1} = (b_{n-1} - a_{n-1,n}x_n)/a_{n-1,n-1}$

$x_1 = (b_i - \sum_{j=2}^{n} a_{ij}x_j)/a_{11}$

or in general

$$x_i = (b_i - \sum_{j=i+1}^{n} a_{ij}x_j)/a_{ii} \quad i = 1, 2, ..., n \qquad (7)$$

Since the unknowns in this method are determined in the order $x_n$, $x_{n-1}$, ..., $x_1$, this method is called the back substitution method. The total number pf multiplications and divisions needed to obtain the complete solution vector x using this method is again n(n + 1)/2.

Let us consider the following example.

**Example 3**: Solve the linear system of equations

$$\begin{aligned} 2x_1 + 3x_2 - x_3 &= 5 \\ -2x_2 - x_3 &= -7 \\ -5x_3 &= -15 \end{aligned}$$

**Solution**: From the last equation, we have

$x_3 = 3.$

From the second equation, we have

$$x_2 = \frac{b_2 - a_{23}x_3}{a_{22}} = \frac{(-7 + 3)}{(-2)} = 2.$$

Hence from the first equation, we get

$$x_1 = \frac{b_1 - a_{12}x_2 - a_{13}x_3}{a_{11}} = \frac{(5 - 3.2 + 3)}{2} = 1$$

In the above discussion you have observed that the system of Eqns. (2) can be easily solved if the coefficient matrix A in Eqns. (2) has one of the three forms D, L or U or if it can be transformed to one of these forms. Now, you would like to know how to reduce the given matrix A into one of these three forms? One such method which transforms the matrix A to the form U is the Gauss elimination method which we shall describe in the next section.

## 3.4 Gauss Elimination Method

Gauss elimination is one of the oldest and most frequently used methods for solving systems of algebraic equations. It is attributed to the famous German mathematician, Carl Fredrick Gauss (1777 – 1855). This method is the generalization of the familiar method of eliminating one unknown between a pair of simultaneous linear equations. You must have learnt this method in your linear algebra course (MTH 122). In this method the matrix A is reduced to the form U by using the elementary row operations which include:

i) interchanging any two rows
ii) multiplying (or dividing) any row by a non-zero constant
iii) adding (or subtracting) a constant multiple of one row to another row.

The operation $R_i + mR_j$ is an elementary row operation, that means, add to the elements of the ith row m times the corresponding elements of the jth row. The elements in the jth row remain unchanged.
If any matrix A is transformed into another matrix B by a series of elementary row operations, we say that A and B are equivalent matrices. Consequently, we have the following definition.

To understand the Gauss elimination method let us consider a system of three equations:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned} \qquad (8)$$

Let $a_{11}{}^{1}$ 0. In the first stage of elimination we multiply the first equation in Eqns. (8) by $m_{21} = (-a_{21}/a_{11})$ and add to the second equation. Then multiply the first equation by $m_{31} = (-a_{31}/a_{11})$ and addto the third equation. This eliminates $x_1$ from the second and third equations. The new system called the first derived system then becomes

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 &= b_2^{(1)} \\ a_{32}^{(1)}x_2 + a_{33}^{(1)}x_3 &= b_3^{(1)} \end{aligned} \qquad (9)$$

where,

$$a_{22}^{(1)} = a_{22} - \frac{a_{21}}{a_{11}}a_{12}$$

$$a_{23}^{(1)} = a_{23} - \frac{a_{21}}{a_{11}}a_{13}$$

$$b_2^{(1)} = b_2 - \frac{a_{21}}{a_{11}}b_1$$

$$a_{32}^{(1)} = a_{32} - \frac{a_{31}}{a_{11}}a_{12}$$

$$a^{(1)}_{33} = a_{33} - \frac{a_{31}}{a_{11}} a_{13}$$

$$b^{(1)}_{3} = b_3 - \frac{a_{31}}{a_{11}} b_1$$

In the second stage of elimination we multiply the second equation in (9) by $m_{32} = (-a^{(1)}_{32}/a^{(1)}_{22})$, $a^{(1)}_{22}\,{}^{1}$ 0 and add to the third equation. This eliminates $x_2$ from the third equation. The new system called the second derived system becomes

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$
$$a^{(1)}_{22}x_2 + a^{(1)}_{23}x_3 = b^{(1)}_{2} \qquad (11)$$
$$a^{(2)}_{33}x_3 = b^{(2)}_{3}$$

where

$$a^{(2)}_{33} = a^{(1)}_{33} - \frac{a^{(1)}_{32}}{a^{(1)}_{22}} a^{(1)}_{23}$$

$$b^{(2)}_{3} = b^{(1)}_{3} - \frac{a^{(1)}_{32}}{a^{(1)}_{22}} b^{(1)}_{2} \qquad (12)$$

You may note here that the system of Eqns. (11) is an upper triangular system of the form (6) and can be solved using the back substitution provided method $a^{(2)}_{33}\,{}^{1}$ 0.

Let us illustrate the method through an example.

**Example 4**: Solve the following linear system

$$2x_1 + 3x_2 – x_3 = 5$$
$$4x_1 + 4x_2 – 3x_3 = 3 \qquad (13)$$
$$-2x_1 + 3x_2 – x_3 = 1$$

using Gauss elimination method.

**Solution**: to eliminate $x_1$ from the second and third equations of the system (13) add $\frac{-4}{2}$ = -2 times the first equation to the second equation and add -(-2)/2 = 1 times the first equation to the third equation. We obtain the new system as

$$2x_1 + 3x_2 – x_3 = 5$$
$$-2x_2 – x_3 = -7 \qquad (14)$$
$$6x_2 – 2x_3 = 6$$

In the second stage, we eliminate $x_2$ from the third equation of system (14). Adding -6/(-2) = 3 times the second equation to the third equation, we get

$$2x_1 + 3x_2 - x_3 = 5$$
$$-2x_2 - x_3 = -7 \qquad (15)$$

$-5x_3 = -15$

System (15) is in upper triangular form and its solution is

$x_3 = 3, x_2 = 2, x_1 = 1.$

You may observe that we can write the above procedure more conveniently in matrix form. Since the arithmetic operations we have performed here affect only the elements of the matrix A and the vector b, we consider the augmented matrix i., [A|b] (matrix A augmented by the vector b) and perform the elementary now operations on the augmented matrix.

$$[A|b] = \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right] R_2 - \frac{a_{21}}{a_{11}} R_1, R_3 - \frac{a_{31}}{a_{11}} R_1$$

$$» \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ & a_{22}^{(1)} & a_{32}^{(1)} & b_2^{(1)} \\ & a_{32}^{(1)} & a_{33} & b_3 \end{array}\right] R_3 - \frac{a_{32}^{(1)}}{a_{22}^{(1)}} R_2$$

$$» \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ & a_{22}^{(1)} & a_{32}^{(1)} & b_2^{(1)} \\ & & a_{33} & b_3 \end{array}\right]$$

which is in the desired from where, $a_{22}^{(1)}$, $a_{23}^{(1)}$ ,$a_{32}^{(1)}$, $a_{33}^{(1)}$, $b_2^{(1)}$, $b_3^{(1)}$, $a_{33}^{(2)}$, $a_3^{(2)}$ are given by Eqns. (10) and (12).

**Definition**: The diagonal elements $a_{11}$, $a_{22}^{(1)}$ and $a_{33}^{(2)}$ which are used as divisors are called pivots.

You might have observed here that for a linear system of order 3, the elimination was performed in 3 – 1 = 2 stages. In general for a system of n equations given by Eqns. (2) the elimination is performed in (n – 1) stages. At the ith stage of elimination, we eliminate $x_i$, starting from (i + 1)th row upto the nth row. Sometimes, it may happen that the elimination process stops in less than (n – 1) stages. But this is possible only when no equations containing the unknowns are left or when the coefficients of all the unknowns in remaining equations become zero. Thus if the process stops at the rth stage of elimination then we get a derived system of the form

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1\\
&a_{22}^{(1)}x_2 + \ldots + a_{2n}^{(1)}x_n = b_2^{(1)}\\
&.\\
&.\\
&. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (16)\\
&a_{rr}^{(r-1)}x_r + \ldots + a_{rn}^{(r-1)}x_n = b_r^{(r-1)}\\
&\qquad\qquad\qquad 0 = b_{r+1}^{(r-1)}\\
&\qquad\qquad\qquad .\quad\ .\\
&\qquad\qquad\qquad .\quad\ .\\
&\qquad\qquad\qquad .\quad\ .\\
&\qquad\qquad\qquad 0 = b_n^{(r-1)}
\end{aligned}$$

Where $r \leq n$ and $a_{11}{}^1\ 0, a_{22}^{(1)}{}^1\ 0, \ldots., a_{rr}^{(r-1)}{}^1\ 0$.

In the solution of system of linear equations we can thus expect two different situations

1) $r = n$
2) $r < n$.

Let us now illustrate these situations through examples.

**Example 5**: Solve the system of equations

$$\begin{aligned}
4x_1 + x_2 + x_3 &= 4\\
x_1 + 4x_2 - 2x_3 &= 4\\
-x_1 + 2x_2 - 4x_3 &= 2
\end{aligned}$$

using Gauss elimination method

**Solution**: Here we have

$$[A|b] = \left[\begin{array}{ccc|c} 4 & 1 & 1 & 4\\ 1 & 4 & -2 & 4\\ 1 & 2 & -4 & 2 \end{array}\right] R_2 - \frac{1}{4}R_1, R_3 + \frac{1}{4}R_1$$

$$= \left[\begin{array}{ccc|c} 4 & 1 & 1 & 4\\ 0 & 15/4 & 9/4 & 3\\ 0 & 9/4 & 15/4 & 3 \end{array}\right] R_3 - \frac{3}{5}R_2$$

$$= \left[\begin{array}{ccc|c} 4 & 1 & 1 & 4 \\ 0 & 15/4 & -9/4 & 3 \\ 0 & 0 & -12/5 & 6/5 \end{array}\right]$$

using back substitution method, we get

$x_3 = -1/2$; $x_2 = 1/2$; $x_1 = 1$

Also, $\det(A) = 4 * \frac{15}{4} * \frac{(-\ 12)}{5} = -36$

Thus in this case we observe that r = n = 3 and the given system of equations has a unique solution. Also the coefficient matrix A in this case is nonsingular. Let us look at another example.

**Example 6**: Solve the system of equations

$$\begin{aligned} 3x_1 + 2x_2 + x_3 &= 3 \\ 2x_1 + x_2 + x_3 &= 0 \\ 6x_1 + 2x_2 + 4x_3 &= 6 \end{aligned}$$

using Gauss elimination method. Does the solution exist?

**Solution**: We have

$$[A|b] = \left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 2 & 1 & 1 & 0 \\ 6 & 2 & 4 & 6 \end{array}\right] \quad R_2 - \frac{2}{3} R_1,\ R_3 - 2R_1$$

$$= \left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 0 & -1/3 & 1/3 & -2 \\ 0 & -2 & 2 & 0 \end{array}\right] \quad R_3 - 6R_2$$

$$= \left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 0 & -1/3 & 1/3 & -2 \\ 0 & 0 & 0 & 12 \end{array}\right]$$

In this case you can see that r < n and elements $b_1$, $b_2^{(1)}$ and $b_3^{(2)}$ are all non-zero.

Since we cannot determine $x_3$ from the last equation, the system has no solution. In such a situation we say that the equations are inconsistent. Also note that det (A) = 0 i.e., the coefficient matrix is singular.

We now consider a situation in which not all b's are non-zero.

**Example 7**: Solve the system of equations

$$16x_1 + 22x_2 + 4x_3 = -2$$
$$4x_1 - 3x_2 + 2x_3 = 9$$
$$12x_1 + 25x_2 + 2x_3 = -11$$

using gauss elimination method.

**Solution**: In this case we have

$$[A|b] = \left[\begin{array}{ccc|c} 6 & 22 & 4 & -2 \\ 4 & -3 & 2 & 9 \\ 12 & 25 & 2 & -11 \end{array}\right] R_2 - \frac{1}{4} R_1, R_3 - \frac{3}{4} R_1$$

$$= \left[\begin{array}{ccc|c} 6 & 22 & 4 & -2 \\ 0 & -17/2 & 1 & 19/2 \\ 0 & 17/2 & -1 & -19/2 \end{array}\right] R_3 + R_2$$

$$= \left[\begin{array}{ccc|c} 6 & 22 & 4 & -2 \\ 0 & -17/2 & 1 & 19/2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Now in this case r < n and elements $b_1$, $b_2^{(1)}$ are non-zero, but $b_3^{(2)}$ is zero. Also the last equation is satisfied for any value of $x_3$. Thus, we get

$x_3$ = any value

$$x_2 = -\frac{2}{17}\left(\frac{19}{2} - x_3\right)$$

$$x_1 = \frac{1}{16}(-2 - 22x_2 - 4x_3)$$

Hence the system of equations has infinitely many solutions.

Note that in this case also det(A) = 0.

The conclusions derived from Examples 4, 5 and 6 are true for any system of linear equations. We now summarize these conclusions as follows:

i) If r = n, then the system of Eqns. (2) has a unique solution which can be obtained using the back substitution method. Moreover, the coefficient matrix A in this case is nonsingular.

ii) If r< n and all the elements $b_{r+1}^{(r-1)}$, $b_{r+2}^{(r-1)}$, ...., $b_{n}^{(r-1)}$ are zero then the system has no solution. In this case we say that the system of equations inconsistent.

iii) If r < n and all the elements $b_{r+1}^{(r-1)}$, $b_{r+2}^{(r-1)}$, ....., $b_{n}^{(r-1)}$, if present, are zero, then the system has infinite number of solutions. In this case the system has only r linearly independent rows.

In both the cases (ii) and (iii), the matrix A is singular.

Now we estimate the number of operations (multiplication and division) in the Gauss elimination method for a system of n linear equations in n unknowns as follows:

No. of divisions
1st step of elimination (n – 1) divisions
2nd step of elimination (n – 2) divisions

(n – 1)th step of elimination 1 divisions
\ Total number of divisions = ( n – 1) + (n – 2) + ..... + 1

$$= \text{å}\ (n-1) = \frac{n(n-1)}{2}$$

No. of multiplications
1st step of elimination n(n – 1) multiplications
2nd stepof elimination(n – 1) (n – 2) multiplications
(n – 1)th step of elimination 2.1 multiplications
\ Total number of multiplications = n(n – 1) + (n – 1) (n – 1) + .... + 2.1

$$= \text{å}\ n(n-1)$$
$$= \text{å}\ n^2 - \text{å}\ n$$
$$= \frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2}$$
$$= \frac{1}{3}n(n+1)(n-1)$$

Also the back substitution adds n divisions (one division at each step) and the numbers of multiplications added are
(n – 1)th equation 1 multiplication
(n – 2)th equation 2 multiplication

1st equation ( n – 1) multiplication

\ Total multiplications = å $(n - 1) = \frac{n(n - 1)}{2}$

Total operation added by back substitution = $\frac{n(n - 1)}{2} + n = \frac{n(n + 1)}{2}$

You can verify these results for n = 3 from Eqns. (9) and (11).

Thus to find the solution vector x using the Gauss elimination method, we need

$$
\begin{aligned}
M &= \frac{n(n - 1)}{2} + \frac{1}{3}n(n^2 - 1) + \frac{n}{2}(n + 1) \\
&= \frac{n}{6}[2n^2 + 6n - 2] \\
&= \frac{n^3}{6} + n^2 - \frac{n}{3}
\end{aligned}
$$

operations. For large n, we may say the total number of operations needed is $\frac{1}{3}n^3$ (approximately). Thus, we find that Gauss elimination method needs much lesser number of operations compared to the Cramer’s rule.

It is clear from above that you can apply Gauss elimination method to a system of equations of any order. However, what happens if one of the diagonal elements i.e., the pivots in the triangularization process vanishes? Then the method will fail. In such situations we modify the Gauss elimination method and this procedure is called pivoting.

**Pivoting**

In the elimination procedure the pivots $a_{11}$, $a_{22}^{(1)}$, ..., $a_{nn}^{(n-1)}$ are used as divisors. If at any stage of the elimination one of these pivots say $a_{ii}^{(i-1)}$, ($a_{11}^{(0)} = a_{11}$), vanishes then the elimination procedure cannot be continued further (see Example 8). Also, it may happen that the pivot $a_{ii}^{(i-1)}$, though not zero, may be very small in magnitude compared to the remaining elements in the ith column. Using a small number as a divisor may lead to the growth of the round-off error. In such cases the multipliers (e.g. $\frac{-a_{i-1,i}^{(i-2)}}{a_{ii}^{(i-1)}}$, $\frac{-a_{i-2,i}^{(i-3)}}{a_{ii}^{(i-1)}}$) will be larger than one in magnitude. The use of large multiplier will lead to magnification of error both during the elimination phase and during the back substitution phase of the solution.

To avoid this we rearrange the remaining rows (ith row upto nth row) so as to obtain a non-vanishing pivot or to make it the largest element in magnitude in that column. The strategy is called pivoting (see Example 9). The pivoting is of the two types; partial pivoting and complete pivoting.

## Partial Pivoting

In the first stage of elimination, the first column is searched for the largest element in magnitude and this largest element is then brought at the position of the pivot by interchanging the first row with the row having the largest element in magnitude in the first column. In the second stage of elimination, the second column is searched for the largest element in magnitude among the (n – 1) elements leaving the first element and then this largest element in magnitude is brought at the position of the second pivot by interchanging the second row with the row having the largest element in the second column. This searching and interchanging of rows is repeated in all the n – 1 stages of the elimination. Thus we have the following algorithm to find the pivot.

For i = 1, 2, ....., n, find j such that

$$\left|a_{ji}^{(i-1)}\right| = \max_{k} \left|a_{ki}^{(i-1)}\right|, \; i \leq k \leq n,$$

and interchange rows i and j.

## Complete Pivoting

In the first stage of elimination, we search the entire matrix A for the largest element in magnitude and bring it at the position of the pivot. In the second stage of elimination we search the square matrix of order n – 1 (leaving the first row and the first column) for the largest element in magnitude and bring it to the position of second pivot and so on. This requires at every stage of elimination not only the interchanging of rows but also interchanging of columns. Complete pivoting is much more complicated and is not often used.

In this unit, by pivoting we shall mean only partial pivoting.

Let us now understand the pivoting procedure through examples.

**Example 8**: Solve the system of equations

$x_1 + x_2 + x_3 = 6$
$3x_1 + 3x_2 + 4x_3 = 20$
$2x_1 + x_2 + 3x_3 = 13$
using Gauss elimination method with partial pivoting.

**Solution**: let us first attempt to solve the system without pivoting. We have

$$[A|b] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 3 & 3 & 4 & 20 \\ 2 & 1 & 3 & 13 \end{array}\right] \; R_2 - 3R_1, \; R_3 - 2R_1$$

$$= \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 0 & 1 & 2 \\ 0 & -1 & 1 & 1 \end{array}\right]$$

Note that in the above matrix the second pivot has the value zero and the elimination procedure cannot be continued further unless, pivoting is used.

Let us now use the partial pivoting. In the first column 3 is the largest element. Interchanging the rows 1 and 2, we have

$$[A|b] = \left[\begin{array}{ccc|c} 3 & 3 & 4 & 20 \\ 1 & 1 & 1 & 6 \\ 2 & 1 & 3 & 13 \end{array}\right] R_2 - \frac{1}{3} R_1, R_3 - \frac{2}{3} R_1$$

$$= \left[\begin{array}{ccc|c} 3 & 3 & 4 & 20 \\ 0 & 0 & -1/3 & -2/3 \\ 0 & -1 & 1/3 & -1/3 \end{array}\right]$$

In the second column, 1 is the largest element in magnitude leaving the first element. Interchanging the second and third rows we have

$$[A|b] = \left[\begin{array}{ccc|c} 3 & 3 & 4 & 20 \\ 0 & -1 & 1/3 & -1/3 \\ 0 & 0 & -1/3 & -2/3 \end{array}\right]$$

You may observe here that the resultant matrix is in triangular form and no further elimination is required. Using back substitution method, we obtain the solution

$x_3 = 2, x_2 = 1, x_1 = 3.$

Let us consider another example.

**Example 9**: Solve the system of equations

$$\begin{aligned} &0.0003\, x_1 + 1.566\, x_2 = 1.569 \\ &0.3454\, x_1 - 0.436\, x_2 = 3.018 \qquad (17) \end{aligned}$$

using Gauss elimination method with and pivoting. Assume that the numbers in arithmetic calculations are rounded to four significant digits. The solution of the system (17) is $x_1 = 10$, $x_2 = 1$.

**Solution**: Without Pivoting

$m_{21} = -\frac{a_{21}}{a_{11}} = -\frac{0.3454}{0.0003} = -1151.0$ (rounded to four places)

$a_{22}^{(1)} = -0.436 - 1.566\ '\ \ 1151$

$= -0.436 - 1802.0 - 1802.436$

$= -1802.0$

$b_2^{(1)} = 3.018 - 1.569\ '\ \ 1151.0$

$= 3.018 - 1806.0$

$= -1803.0$

Thus, we get the system of equations

$$0.0003\, x_1 + 1.566\, x_2 = 1.569$$
$$-\, 1802.0\, x_2 = -1803.0$$

which gives

$x_2 = \frac{1803.0}{1802.0} = 1.001$

$x_1 = \frac{1.569 -\ \ 1.566\ '\ \ 1.001}{0.0003} = \frac{1.569 -\ \ 1.568}{0.0003}$

$= 3.333$

which is highly inaccurate compared to the exact solution.

We interchange the first and second equations in (17) and get

$0.3454\, x_1 - 0.436\, x_2 = 3.018$
$0.0003\, x_1 + 1.566\, x_2 = 1.569$

we obtain

$m_{21} = -\frac{a_{21}}{a_{11}} = -0.0009$

$a_{22}^{(1)} = 1.566 - 0.0009\ '\ \ (0.436)$

$1.566 - 0.0004$

$= 1.566$

$b_2^{(1)} = 1.569 - 3.018\ '\ \ (0.0009)$

$= 1.569 - -.0027$

$= 1.566$

Thus, we get the system of equations

$0.3454\, x_1 - 0.436\, x_2 = 3.018$

$$1.566\, x_2 = 1.566$$

which gives

$x_2 = i$

$$x_1 = \frac{3.018 + 0.436}{0.3454} = \frac{3.454}{0.3454} = 10$$

which is the exact solution.

We now make the following two remarks about pivoting.

**Remark**: If the matrix A is diagonally dominant i.e.,
$|a_{ii}|^3 \sum_{\substack{i=1 \\ j=1}}^{n} |a_{ii}|$, then no pivoting is needed. See Example 5 in which A is diagonally dominant.

**Remark**: If exact arithmetic is used throughout the computation, pivoting is not necessary unless the pivot vanishes. However, if computation is carried upto a fixed number of digits, we get accurate results if pivoting is used.

There is another convenient way of carrying out the pivoting procedure. Instead of physically interchanging the equations all the time, the n original equations and the various changes made in them can be recorded in a systematic way. Here we use an n ′ (n + 1) working array or matrix which we call W and is same as our augmented matrix [A|b]. Whenever some unknown is eliminated from an equation, the changed coefficients and right side for this equation are calculated and stored in the working array W in place of the previous coefficients and right side. Also, we use an n-vector which we call p = ($p_i$) to keep track of which equations have already been used as pivotal equation (and therefore should not be changed any further) and which equations are still to be modified. Initially, the ith entry $p_i$pf p contains the integer i, i = 1, ........, n and working array W is of the form

$$W = (w_{ij}) = \left[\begin{array}{cccc|c} a_{11} & a_{12} & & a_{1n} & b_1 \\ a_{21} & a_{22} & & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & \\ a_{n1} & a_{n2} & & a_{nn} & b_n \end{array}\right]$$

Further, one has to be careful in the selection of the pivotal equation for each step. For each step the pivotal equation must be selected on the basis of the current state of the system under consideration i.e. without foreknowledge of the effect of the i = 1, ......, n, where $d_i$ is the number

$d_i = \max |a_{ij}|$

$1 \leq j \leq n$

At the beginning of say kth step of elimination, e pick as pivotal equation that one from the available n – k, which has the absolutely largest coefficient of $x_k$ relative to the size of the equation. This means that the integer j is selected between k and n for which

$\frac{|w_{p_{jk}}|}{d_{p_j}}$ 3 $\frac{|w_{ik}|}{d_i}$, " $i = p_k, .....,p_n$

We can also store the multipliers in the working array W instead of storing zeros. That is, if $p_i$ is the first pivotal equation and we use the multipliers $m_{pi,1}$, i = 2, ....., n to eliminate $x_1$ from the remaining (n – 1) positions of the first column then in the first column we can store the multipliers $m_{pi,1}$, i = 2, ....., n, instead of storing zeros.

Let us now solve the following system of linear equations by scaled partial pivoting by storing the multipliers and maintaining pivotal vector.

**Example 10**: Solve the following system of linear equations with pivoting

$x_1 – x_2 + 3x_3 = 3$
$2x_1 + x_2 + 4x_3 = 7$
$3x_1 + 5x_2 – 2x_3 = 6$

**Solution**: Here the working matrix is

$$W = \begin{bmatrix} 1 & -1 & 3 & 3 \\ 2 & 1 & 4 & 7 \\ 3 & 5 & -2 & 6 \end{bmatrix} \quad p = [p_1, p_2, p_3]^T = [1, 2, 3]^T$$

and $d_1 = 3$, $d_2 = 4$and $d_3 = 5$.

Note that d's will not change in the successive steps.

Step 1: Now $\frac{|w_{p1,1}|}{d_1} = \frac{1}{3} \frac{|w_{p2,1}|}{d_2} = \frac{2}{4} = \frac{1}{2}, \frac{|w_{p3,1}|}{d_3} = \frac{3}{5}.$

Since $\frac{3}{5} > \frac{1}{2}, \frac{1}{3},$

Hence, $p_1 = 3$, $p_2 = 2$ and $p_3 = 1$.

We use the third equation to eliminate $x_1$ from first and second equations and store corresponding multipliers instead of storing zeros in the working matrix.

The multipliers are $m_{pi,1} = \dfrac{w_{p_{i,1}}}{w_{p_{i,1}}}$, i = 2, 3

Therefore, $m_{2,1} = \dfrac{w_{p_{2,1}}}{w_{p_{1,1}}} = \dfrac{w_{2,1}}{w_{3,1}} = \dfrac{2}{3}$

and $m_{1,1} = \dfrac{w_{p_{3,1}}}{w_{p_{1,1}}} = \dfrac{w_{1,1}}{w_{3,1}} = \dfrac{1}{3}$

After the first step the working matrix is transformed to

$$W^{(1)} = \begin{bmatrix} (1/3) & -8/3 & 11/3 & 1 \\ (2/3) & -7/3 & 16/3 & 3 \\ \lceil 3 \rceil & 5 & -2 & 6 \end{bmatrix} \quad p = (p_1, p_2, p_3)^T = (3, 2, 1)^T$$

$$\text{Step 2: } \frac{|w_{p_{2,2}}|}{dp_2} = \frac{|w_{2,2}|}{d_2} = \frac{7/3}{4} = \frac{7}{12}$$

$$\frac{|w_{p_{3,2}}|}{dp_3} = \frac{|w_{1,2}|}{d_1} = \frac{8/3}{3} = \frac{8}{9}$$

Now $\dfrac{8}{9} > \dfrac{7}{12}$ so that we have p = $(p_1, p_2, p_3)^T = (3, 2, 1)^T$.

Multiplier is $m_{p_{i,2}} = \dfrac{w_{p_{i,2}}}{w_{p_{2,2}}}$, i = 3

$$Þ \quad m_{p_{3,2}} = \frac{w_{p_{i,2}}}{w_{p_{2,2}}} = \frac{-7/3}{-8/3} = \frac{7}{8}.$$

That is, we use the first equation as pivotal equation to eliminate $x_2$ from second equation and also we store the multiplier. After the second step, we have the following working matrix.

$$W^{(2)} = \begin{bmatrix} \frac{1}{3} & -\frac{8}{3} & \frac{11}{3} & 1 \\ \frac{2}{3} & \frac{7}{8} & \frac{51}{24} & \frac{17}{8} \\ 3 & 5 & -2 & 6 \end{bmatrix} \quad p = [3, 1, 2]^T$$

In the working matrix the circled numbers denote multipliers and squared ones denote pivotal elements. Rearranging the equations (i.e., 3rd equation becomes the first equation, 1st becomes the 2nd and 2nd becomes the third) we get the reduced upper triangular system which can be solved by back substitution.

$$3x_1 + 5x_2 - 2x_3 = 6$$
$$-\frac{8}{3}x_2 + \frac{11}{3}x_3 = 1$$
$$\frac{51}{24}x_3 = \frac{17}{8}$$

By back substitution, we get $x_1 = 1$, $x_2 = 1$ and $x_3 = 1$.

We now make the following two remarks.

**Remark**: We do not interchange rows in Step 1 and 2, instead we maintain a pivotal vector and use it at the end to get upper triangular system.

**Remark**: We store multipliers in the working matrix so that we can easily solve Ax = c, once we have solved Ax = b. This will be explained to you in detail in Unit 2 when we discuss the method of obtaining inverse of a matrix A.

We shall now describe the triangularization method which is also a direct method for the solution of system of equations.
In this method the matrix of coefficients of the linear system being solved is factored into the product of two triangular matrices. This method is frequently used to solve a large system of equations. We shall discuss the method in the next section.

## 3.5 LU Decomposition Method

Let us consider the system of Eqns. (2), where A is a non-singular matrix. We first write the matrix A as the product of a lower triangular matrix L and an upper triangular matrix U in the form

A = LU
or in matrix form we write (18)

$$\begin{bmatrix} a_{11} & a_{12} & & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ & & & \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & & 0 \\ l_{21} & l_{22} & & 0 \\ & & & \\ l_{n1} & l_{n2} & & l_{nn} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & & u_{1n} \\ 0 & u_{22} & & u_{2n} \\ & & & \\ 0 & 0 & & u_{nn} \end{bmatrix} \quad (19)$$

The left side matrix A has $n^2$ elements, whereas L and U have 1 + 2 + ... + n = n(n + 1)/2 elements each. Thus, we have $n^2 + n$ unknowns in L and U which are to be determined. On comparing the corresponding elements on two sides in Eqn. (19), we get $n^2$ equations in $n^2 + n$ unknowns and hence n unknowns are determined. Thus, we get a solution in terms of these n unknowns i.e., we get a n parameter family of solutions. In order to obtain a unique solution we either take all the diagonal elements of L as 1, or all the diagonal elements of U as 1.

For $u_{ij} = 1$, i = 1, 2, ...., n, the method is called the Crout LU decomposition method. For $1_{ii} = 1$, i = 1, 2, ...., n we have Doolittle LU decomposition method. Usually Crout's LU decomposition method is used unless it is specifically mentioned. We shall now explain the method for n = 3 with $u_{ii} = 1$, i = 1, 2, 3. We have

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & o \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix}$$

or

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} l_{11} & l_{11}u_{12} & l_{11}u_{13} \\ l_{21} & l_{21}u_{22} + l_{22} & l_{21}u_{23} + l_{22}u_{23} \\ l_{31} & l_{31}u_{12} + l_{32} & l_{31}u_{13} + l_{32}u_{23} + l_{33} \end{bmatrix}$$

On comparing the elements of the first column, we obtain

$1_{11} = a_{11}$, $1_{21} = a_{21}$, $1_{31} = a_{31}$ (20)
i.e., the first column of L is determined.

On comparing the remaining elements of the first row, we get

$1_{11}u_{12} = a_{12}$; $1_{11}u_{13} = a_{13}$

which gives

$u_{12} = a_{12}/1_{11}$; $u_{13} = a_{13}/1_{11}$ (21)

Hence the first row of U is determined

On comparing the elements of the second column, we get

$1_{21}u_{12} + 1_{22} = a_{22}$
$1_{31}u_{12} + 1_{32} = a_{32}$

which gives

$$\begin{bmatrix} l_{22} & = a_{22} - l_{21}u_{12} \\ l_{32} & = a_{32} - l_{31}u_{12} \end{bmatrix} \qquad (22)$$

Now the second column of L is determined.

On comparing the elements of the second row, we get

$l_{21}u_{13} + l_{22}u_{23} = a_{23}$

which gives $u_{23} = (a_{23} - l_{21}\ u_{13})/l_{22}$ (23)

and the second row of U is determined.

On comparing the elements of the third column, we get

$l_{31}u_{13} + l_{32}u_{23} + l_{33} = a_{33}$

which gives $l_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23}$ (24)

You must have observed that in this method, we alternate between getting a column of L and a row of U in that order. If instead of $u_{ii}$ = 1 1, 2, ...., n, we take $l_{ii}$ = 1, i = 1, 2, ...., n, then we alternative between getting a row of U and a column of L in that order.

Thus, it is clear from Eqns. (20) – (24) that we can determine all the elements of L and U provided the nonsingular matrix A is such that

$$a_{11}{}^{1}\ 0, \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{1}\ 0.$$

Similarly, for the general system of Eqns. (2), we obtain the elements of L and U using the relations

$$l_{ij} = a_{ij} - \sum_{i=1}^{j=1} l_{ik}u_{kj},\ i^{3}\ j$$

$$u_{ij} = (a_{ij} - \sum_{i=1}^{j=1} l_{ik}u_{kj})/l_{ii},\ i^{3}\ j$$

$u_{ii} = 1$

Also, det (A) = $l_{11}l_{22}$ ....., $l_{nn}$.

Thus w can say that every nonsingular matrix A can be written as the product of a lower triangular matrix and an upper triangular matrix if all principal minors of A are nonsingular, i.e., if

$$a_{11}{}^{1}\ 0, \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{1}\ 0, \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix}{}^{1}\ 0, .....|A|\ {}^{1}\ 0.$$

Once we have obtained the elements of the matrices L and U, we write the system of equations

$$A\,x = b \qquad (25)$$

in the form

$$L\,U\,x = b \qquad (26)$$

The system (26) may be further written as the following two systems

$$U\,x = y \qquad (27)$$
$$L\,y = b \qquad (28)$$

Now, we first solve the system (28), i.e.,

$L\,y = b$,

using the forward substitution method to obtain the solution vector y. Then using this y, we solve the system (27), i.e.,

$U\,x = y$,
using the backward substitution method to obtain the solution vector x.

The number of operations for this method remains the same as that in the Gauss-elimination method.

We now illustrate this method through an example.

**Example 11**: Use the LU decomposition method to solve the system of equations
$x_1 + x_2 + x_3 = 1$
$4x_1 + 3x_2 - x_3 = 6$
$3x_1 + 5x_2 + 3x_3 = 4$
**Solution**: Using $1_{ii} = 1$, i = 1, 2, 3, we have

$$\begin{bmatrix} 1 & 1 & 1 \\ 4 & 3 & -1 \\ 3 & 5 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{31} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

$$=\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ l_{21}u_{11} & l_{21}u_{12}+u_{22} & l_{21}u_{13}+u_{23} \\ l_{31}u_{11} & l_{31}u_{12}+l_{32}u_{22} & l_{31}u_{13}+l_{32}u_{23}+u_{33} \end{bmatrix}$$

On comparing the elements of row and column alternatively, on both sides, we obtain

| | |
|---|---|
| first row | : $u_{11} = 1$, $u_{12} = 1$, $u_{13} = 1$ |
| first column | : $l_{21} = 4$, $l_{31} = 3$ |
| second row | : $u_{22}$ - -1, $u_{23} = -5$ |
| second column | : $l_{32} = -2$ |
| third row | : $u_{33} = -10$ |

Thus, we have

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 3 & -2 & 1 \end{bmatrix} \quad U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -5 \\ 0 & 0 & -10 \end{bmatrix}$$

Now from the system

L y = b

or

$$\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 3 & -2 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ 4 \end{bmatrix}$$

we get

$y_1 = 1$, $y_2 = 2$, $y_3 = 5$

and from the system

U x = y

or

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -5 \\ 0 & 0 & -10 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$$

we get

$x_3 = -1/2$, $x_2 = 1/2$, $x_1 = 1$.

## 4.0 CONCLUSION

As in the summary

## 5.0 SUMMARY

In this unit we have covered the following:

a) For a system of n equations
Ax = b (see Eqn. (2))
in n unknowns, where A is n ´ n non-singular matrix, the methods of finding the solution vector x may be broadly classified into two types: (1) direct methods and (ii) iterative methods

b) Direct methods produce the exact solution in a finite number of steps provided there are no round-off errors. Cramer's rule is one such method. This method gives the solution vector as

$$x_i = \frac{d_i}{d} \quad i = 1, 2, ..., n$$

where d = |A| and $d_i$ is the determinant pf the matrix obtained from A by replacing the ith column of A by the column vector b. Total number of operations required for Cramer's rule in solving a system of n equations are
M = (n + 1) (n – 1)n! + n
Since the number M increases very rapidly, Cramer's rule is not used for n > 4.

c) For larger systems, direct methods becomes more efficient if the coefficient matrix A is in one of the forms D (diagonal), L (lower triangular) or U (upper triangular).

d) Gauss elimination method is another direct method for solving large systems (n > 4). In this method the coefficient matrix A is reduced to the form U by using the elementary row operations. The solution vector x is then obtained by using the back substitution method. For large n, the total numbers of operations required in Gauss elimination method are $\frac{1}{3}n^3$ (approximately).

e) In Gauss elimination method if at any stage of the elimination any of the pivots vanishes or become small in magnitude, elimination procedure cannot be continued further. In such cases pivoting is used to obtain the solution vector x.

f) Every non-singular matrix A can be written as the product of a lower triangular matrix and an upper triangular matrix, by the LU decomposition method, if all the principal minors of A are non-singular. Thus, LU decomposition method, which is a modification of the Gauss elimination method can be used to obtain the solution vector x.

## 6.0 TUTOR-MARKED ASSIGNMENT (TMA)

i If $A = \begin{bmatrix} 3 & -2 & 0 & 2 \\ 2 & 1 & 0 & -1 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & -3 & 1 \end{bmatrix}$ calculate det (A).

ii Solve the system of equations

$3x_1 + 5x_2 \quad = 8$

$-x_1 + 2x_2 - x_3 = 0$

$3x_1 - 6x_2 + 4x_3 = 1$

using Cramer's rule.

iii Solve the system of equations

$x_1 + 2x_2 - 3x_3 + x_4 = -5$

$x_2 + 3x_3 + x_4 = 6$

$2x_1 + 3x_2 + x_3 + x_4 = 4$

$x_1 \quad + x_3 + x_4 = 1$

using Cramer's rule.

iv Solve the system of equations

$x_1 = 1$

$2x_1 = x_2 = 1$

$3x_1 - x_2 - 2x_3 = 0$

$4x_1 + x_2 - 3x_3 + x_4 = 3$

$5x_1 - 2x_2 - x_3 - 2x_4 + x_5 = 1$

using forward substitution method.

v Solve the system of equations

$x_1 - 2x_2 + 3x_3 - 4x_4 + 5x_5 = 3$

$x_2 - 2x_3 + 3x_4 - 4x_5 = -2$

$x_3 - 2x_4 + 3x_5 = 2$

$x_4 - 2x_5 = -1$

$x_5 = 1$

using backward substitution method.

vi Use Gauss elimination method to solve the system of equations

$x_1 + 2x_2 + x_3 = 3$

$3x_1 - 2x_2 - 4x_3 = -2$

$2x_1 + 3x_2 - x_3 = -6$

vii Solve the system of equations

$$\begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & 1 & 3 & 1 \\ 2 & 3 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ 4 \\ 1 \end{bmatrix}$$

viii Use Gauss elimination method to solve the system of equations

$$\begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ 1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

ix Solve the system of equations

$0.729x + 0.81y + 0.9z = 0.6867$

$x + y + z = 0.8338$

$1.331x + 1.21y + 1.1z = 1.000$

using gauss eliminating method with and without pivoting. Round off the numbers in arithmetic calculations to four significant digits. The exact solution of the system rounded to four significant digit is

$x = 0.2245, y = 0.2814 \ \ z = 0.3279$

x. Use the LU decomposition method with $u_{ii} = 1$, i = 1, 2, 3 to solve the system of equations given in Example 11.

xi Use the LU decomposition method with $1_{ii} = 1$, i = 1, 2, 3 to solve the system of equations given in TMA Question 4 no. 1.

xii Use L U decomposition method to solve the system of equations given in TMA Question 4 no. 3.

## 7.0 REFERENCES/FURTHER READINGS.

Wrede, R.C. and Spegel M. (2002). Schaum's and Problems of Advanced Calculus. McGraw – Hill N.Y.

Keisler, H.J. (2005). Elementary Calculus. An Infinitesimal Approach. 559 Nathan Abbott, Stanford, California, USA

# UNIT 2 INVERSE OF A SQUARE MATRIX

## CONTENTS

## 1.0 INTRODUCTION

In the previous unit, you have studied the Gauss elimination and LU decomposition methods for solving systems of algebraic equations A x = , when A is a n ′ nnonsingular matrix. Matrix inversion is another problem associated with the problem of finding solutions of a linear system. If the inverse matrix $A^{-1}$ of the coefficient matrix A is known then the solution vector x can be obtained from $x = A^{-1}b$. In genral, inversion of matrices for solving system of equations should be avoided whenever possible. This is because, it involves greater amount of work and also it is difficult to obtain the inverse accurately in many problems. However, there are two cases in which the explicit computation of the inverse is desirable. Firstly, when several systems equations, having the same coefficient matrix A but different right hand side b, have to be solved. Then computations are reduced if we first find the inverse matrix and then find the solution. Secondly, when the elements of $A^{-1}$ themselves have some special physical significance. For instance, in the statistical treatment of the fitting of a function to observational data by the method of least squares, the elements of $A^{-1}$ give information about the kind and magnitude of errors in the data.

In this unit, we shall study a few important methods for finding the inverse of a nonsingular square matrix.

## 2.0 OBJECTIVES

At the end of this unit, you should be able to:

- obtain the inverse by adjoint method for $n < 4$
- obtain the inverse by the Gauss-Jordan and LU decomposition methods
- obtain the solution of a system of linear equations using the inverse method.

## 3.0 MAIN CONTENTS

### 3.1 The Method of Adjoints

You already know that the transpose of the matrix of the cofactors of elements of A is called the adjoint matrix and is denoted by adj(A).

Formally, we have the following definition.

**Definition**: The transpose of the cofactor matrix $A^c$ of A is called the adjoint of A and is written a adj(A).

$$\text{adj}(A) = (A^c)^T$$

The inverse of a matrix can be calculated using the adjoint of a matrix.

E obtain the inverse matrix $A^{-1}$ of A from

$$A^{-1} = \frac{1}{\det(A)}\,\text{adj}(A) \qquad (1)$$

This method of finding the inverse of a matrix is called the method of adjoints.

Note that det(A) in Eqn. (1) must not be zero and therefore the matrix A must be nonsingular.

We shall not be going into the details of the method here. We shall only illustrate it through examples.

**Example 1**: Find $A^{-1}$ for the matrix

$$A = \begin{bmatrix} 5 & 8 & 1 \\ 0 & 2 & 1 \\ 4 & 3 & -1 \end{bmatrix}$$

and solve the system of equations

$$A\,x = b \qquad (2)$$

for

i) $b = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$ ii) $b = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ iii) $b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

**Solution**: Since det (A) = -1 $^{1}$ 0, the inverse of A exists. We obtain the cofactor matrix $A^c$ from A by replacing each element of A by its cofactor as follows:

$$A^c = \begin{bmatrix} 5 & 4 & -8 \\ 11 & -9 & 17 \\ 6 & -5 & 10 \end{bmatrix}$$

$$\backslash\ \text{adj}(A) = (A^c)^T = \begin{bmatrix} 5 & 11 & 6 \\ 4 & -9 & -5 \\ 8 & 17 & 10 \end{bmatrix}$$

Now $A^{-1} = \dfrac{1}{\det(A)}\text{adj}(A)$

$$\backslash\ A^{-1} = - = \begin{bmatrix} 5 & 11 & 6 \\ 4 & -9 & -5 \\ 8 & 17 & 10 \end{bmatrix} = \begin{bmatrix} 5 & -11 & -6 \\ 4 & 9 & 5 \\ 8 & -17 & -10 \end{bmatrix}$$

Also the solution of the given system of equations are

i) $x = A^{-1}b = \begin{bmatrix} 5 & -11 & -6 \\ 4 & 9 & 5 \\ 8 & -17 & -10 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix}$

ii) $x = A^{-1}b = \begin{bmatrix} 5 & -11 & -6 \\ 4 & 9 & 5 \\ 8 & -17 & -10 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \\ 8 \end{bmatrix}$

iii) $x = A^{-1}b = \begin{bmatrix} 5 & -11 & -6 \\ 4 & 9 & 5 \\ 8 & -17 & -10 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 7 \\ 12 \end{bmatrix}$

We now take up an example in which the given matrix A is lower triangular and we shall show that its inverse is also a lower triangular matrix.

**Example 2**: Find $A^{-1}$ for the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{bmatrix}$$

**Solution**: We have
det(A) = 18 $^{1}$ 0. Thus $A^{-1}$ exists.

Now

$$A^c = \begin{bmatrix} 18 & -12 & -2 \\ 0 & 6 & -5 \\ 0 & 0 & 3 \end{bmatrix}$$

$$\backslash \quad A^{-1} = \frac{(A^c)^T}{\text{adj}(A)} = \frac{1}{18}\begin{bmatrix} 18 & 0 & 0 \\ 12 & 6 & 0 \\ -2 & -5 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1/3 & 0 \\ 1/9 & -5/18 & 1/6 \end{bmatrix}$$

Thus, $A^{-1}$ is again a lower triangular matrix. Similarly, we can illustrate that the inverse of an upper triangular matrix is again upper triangular.

**Example 3**: Find $A^{-1}$ for the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$$

**Solution**: Since, det(A) = 24 $^{1}$ 0, $A^{-1}$ exists.

We obtain

$$A^c = \begin{bmatrix} 24 & 0 & 0 \\ 12 & 6 & 0 \\ -2 & -5 & 4 \end{bmatrix}$$

$$\backslash \quad A^{-1} = \frac{1}{24}\begin{bmatrix} 24 & -12 & -2 \\ 0 & 6 & -5 \\ 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & -1/2 & -1/12 \\ 0 & 1/4 & -5/24 \\ 0 & 0 & 1/6 \end{bmatrix}$$

which is again an upper triangular matrix.

The method of adjoints provides a systematic procedure to obtain the inverse of a given matrix and for solving systems of linear equations. To obtain the inverse of an n ′ n matrix, using this method, we need to evaluate one determinant of order n, n determinants each of order n – 1 and perform $n^2$ divisions. In addition, if this method is used for solving a linear system we also need matrix multiplication. The number of operations (multiplications and divisions) needed, for using this method, increases very rapidly as n increases. For this reason, this method is not used when n > 4.

For large n, there are methods which are efficient and are frequently used for finding the inverse of a matrix and solving linear systems. We shall now discuss these methods.

## 3.2 The Gauss-Jordan Reduction Method

This method is a variation of the Gauss elimination method. In the Gauss elimination method, using elementary row operations, we transform the matrix A to an upper triangular matrix U and obtain the solution by using back substitution method. In Gauss-Jordan reduction not only the elements below the diagonal but also the elements above the diagonal of A are made zero at the same time. In other words, we transform the matrix A to a diagonal matrix D. This diagonal matrix may then be reduced to an identity matrix by dividing each row by its pivot element.

Alternately, the diagonal elements can also be made unity at the same time when the reduction is performed. This transforms thecoefficient matrix into an identity matrix. Thus, on completion of the Gauss-Jordan method, we have

$$[A|b]\ [I|d] \Longrightarrow \qquad (3)$$

The solution is then given by

$$x_i = d_i,\ i = 1, 2, ......, n \qquad (4)$$

In this method also, we use elementary row operations that are used in the Gauss elimination method. We apply these operations both below and above the diagonal in order to reduce all the off-diagonal elements of the matrix to zero. Pivoting can be used to make the pivot non-zero or make it the largest element in magnitude in that column as discussed. We illustrate the method through an example.

**Example 4**: Solve the system of equations

$x_1 + x_2 + x_3 = 1$
$4x_1 + 3x_2 - x_3 = 6$
$3x_1 + 5x_2 + 3x_3 = 4$

using Gauss-Jordan method with pivoting.

**Solution**: We have

$$[A|b] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 4 & 3 & -1 & 6 \\ 3 & 5 & 3 & 4 \end{array}\right] \text{(interchanging first and second row)}$$

$$» \left[\begin{array}{ccc|c} 4 & 3 & -1 & 6 \\ 1 & 1 & 1 & 1 \\ 3 & 5 & 3 & 4 \end{array}\right] R_2 - \frac{1}{4} R_1, R_3 - \frac{3}{4} R_1$$

$$» \left[\begin{array}{ccc|c} 4 & 3 & -1 & 6 \\ 0 & 1/4 & 5/4 & -1/2 \\ 0 & 11/4 & 15/4 & -1/2 \end{array}\right] \text{(interchanging second and third row)}$$

$$» \left[\begin{array}{ccc|c} 4 & 3 & -1 & 6 \\ 0 & 11/4 & `15/4 & -1/2 \\ 0 & 1/4 & 5/4 & -1/2 \end{array}\right] R_3 – 1/11\ R_2, R_1 - \frac{12}{11} R_2$$

$$» \left[\begin{array}{ccc|c} 4 & 0 & -56/11 & 72/11 \\ 0 & 11/4 & 15/4 & -1/2 \\ 0 & 0 & 10/11 & -5/11 \end{array}\right] R_1 + \frac{56}{10} R_3, R_2 - \frac{33}{8} R_3$$

$$» \left[\begin{array}{ccc|c} 4 & 0 & 0 & 4 \\ 0 & 11/4 & 0 & 11/8 \\ 0 & 0 & 10/11 & 5/11 \end{array}\right]$$

R1/4 (divide first row by 4),

$\frac{4}{11}R_2$ (divide second row by 11/ 4),

$\frac{11}{10}R_3$ (divide third row by 10/ 11).

$$» \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 1 & -1/2 \end{array}\right]$$

which is the desired form.
Thus, we obtain
$x_1 = 1, x_2 = \frac{1}{2}, x_3 = -\frac{1}{2}$.

The method can be easily extended to a general system of n equations. Just as we calculated the number of operations needed for Gauss elimination method in the same way you can verify that the total number of operations needed for this method is M = $\frac{1}{2} n^3 + \frac{n^2}{2} + n$.

Clearly this method requires more number of operations compared to the Gauss elimination method. We therefore, do not use this method generally for solving system of equations but is very commonly used for finding the inverse matrix. This is done by augmenting the matrix A by the identity matrix I of the order same as that of A. Using elementary row operations on the augmented matrix [A|I] we reduce the matrix A to the form I and in the process the matrix I is transformed to $A^{-1}$

That is

$$[A|I] \Longrightarrow [I|A^{-1}] \qquad (5)$$

We now illustrate the method through examples.

**Example 5**: Find the inverse of the matrix

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 2 & -3 & -1 \\ 1 & -2 & 1 \end{bmatrix}$$

using the Gauss-Jordan method.

**Solution**: We have

$$[A|I] = \left[\begin{array}{ccc|ccc} 3 & 1 & 2 & 1 & 0 & 0 \\ 2 & -3 & -1 & 0 & 1 & 0 \\ 1 & -2 & 1 & 0 & 0 & 1 \end{array}\right]_{R/3}$$

$$\text{»} \left[\begin{array}{ccc|ccc} 1 & 1/3 & 2/3 & 1/3 & 0 & 0 \\ 2 & -3 & -1 & 0 & 1 & 0 \\ 1 & -2 & 1 & 0 & 0 & 1 \end{array}\right]_{R - 2R,\ R - R}$$

$$\text{»} \left[\begin{array}{ccc|ccc} 1 & 1/3 & 2/3 & 1/3 & 0 & 0 \\ 0 & -11/3 & -7/3 & -2/3 & 1 & 0 \\ 0 & -7.3 & 1/3 & -1/3 & 0 & 1 \end{array}\right]\ 3R/11$$

$$\text{»} \left[\begin{array}{ccc|ccc} 1 & 1/3 & 2/3 & 1/3 & 0 & 0 \\ 0 & 1 & 7/11 & 2/11 & -3/11 & 0 \\ 0 & -7/3 & 1/3 & -1/3 & 0 & 1 \end{array}\right] R_1 - \frac{1}{3}R_2,\ R_3 + \frac{7}{3}R_2$$

$$\text{»} \left[\begin{array}{ccc|ccc} 1 & 0 & 5/11 & 3/11 & 1/11 & 0 \\ 0 & 1 & 7/11 & 2/11 & -3/11 & 0 \\ 0 & 0 & 20/11 & 1/11 & -7/11 & 1 \end{array}\right] \frac{11}{20}R_3$$

$$\approx \left[\begin{array}{ccc|ccc} 1 & 0 & 5/11 & 3/11 & 1/11 & 0 \\ 0 & 1 & 7/11 & 2/11 & -3/11 & 0 \\ 0 & 0 & 0 & 1/20 & -7/20 & 11/20 \end{array}\right] R_1 - \frac{5}{11}R_3, R_2 - \frac{7}{11}R_3$$

$$\approx \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/4 & 1/4 & -1/4 \\ 0 & 1 & 0 & 3/20 & -1/20 & -7/20 \\ 0 & 0 & 1 & 1/20 & -7/20 & 11/20 \end{array}\right]$$

Thus, we obtain

$$A^{-1} = \begin{bmatrix} 1/4 & 1/4 & -1/4 \\ 3/20 & -1/20 & -7/20 \\ 1/20 & -7/20 & 11/20 \end{bmatrix}$$

**Example 6**: Find the inverse of the matrix

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 1/2 & 0 & 0 \\ 1 & 0 & -3 & 0 \\ 1 & -7/2 & -17 & 55/3 \end{bmatrix}$$

using the Gauss-Jordan method

**Solution**: Here we have

$$[A|I] = \left[\begin{array}{cccc|cccc} 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1/2 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & -3 & 0 & 0 & 0 & 1 & 0 \\ 1 & -7/2 & -17 & 55/3 & 0 & 0 & 0 & 1 \end{array}\right] \frac{1}{2}R_1$$

$$\approx \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 1 & 1/2 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & -3 & 0 & 0 & 0 & 1 & 0 \\ 1 & -7/2 & -17 & 55/3 & 0 & 0 & 0 & 1 \end{array}\right]$$

$R_2 – R_1, R_3 – 2R_1, R_4 – R_1$

$$\gg \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & -1/2 & 1 & 0 & 0 \\ 0 & 0 & -3 & 0 & -1 & 0 & 1 & 0 \\ 0 & -7/2 & -17 & 55/3 & -1/2 & 0 & 0 & 1 \end{array}\right] 2R_2$$

$$\gg \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 2 & 0 & 0 \\ 0 & 0 & -3 & 0 & -1 & 0 & 1 & 0 \\ 0 & -7/2 & -17 & 55/3 & -1/2 & 0 & 0 & 1 \end{array}\right] R_4 + \frac{7}{2}R_2$$

$$\gg \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 2 & 0 & 0 \\ 0 & 0 & -3 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -17 & 55/3 & -4 & 7 & 0 & 1 \end{array}\right] \left(- \frac{1}{3}R_3\right)$$

$$\gg \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1/3 & 0 & -1/3 & 0 \\ 0 & 0 & -17 & 55/3 & -4 & 7 & 0 & 1 \end{array}\right] \left(- \frac{1}{17}R_4\right)$$

$$\gg \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1/3 & 0 & -1/3 & 0 \\ 0 & 0 & -17 & 55/3 & 4/17 & -7/17 & 0 & -1/17 \end{array}\right] R_4–R_3$$

$$\gg \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1/3 & 0 & -1/3 & 0 \\ 0 & 0 & 0 & -55/51 & -5/51 & -7/17 & 1/3 & -1/17 \end{array}\right] \left(- \frac{51}{55}R_4\right)$$

$$\gg \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1/3 & 0 & -1/3 & 0 \\ 0 & 0 & 0 & 1 & 1/11 & 21/55 & -17/55 & 3/55 \end{array}\right]$$

Hence

$$A^{-1} = \begin{bmatrix} 1/2 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 1/3 & 0 & -1/3 & 0 \\ 1/11 & 21/55 & -17/55 & 3/55 \end{bmatrix}$$

is the inverse of the given lower triangular matrix.

Let us now consider the problem of finding the inverse of an upper triangular matrix.

**Example 7**: Find the inverse of the matrix

$$A = \begin{bmatrix} 1 & 3/2 & 2 & 1/2 \\ 0 & 1 & -4 & 1 \\ 0 & 0 & 1 & 2/3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

using the Gauss-Jordan method.

$$[A|I] = \left[\begin{array}{cccc|cccc} 1 & 3/2 & 2 & 1/2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -4 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2/3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] R_1 - \frac{3}{2}R_2$$

$$» \left[\begin{array}{cccc|cccc} 1 & 0 & 8 & -1 & 1 & -3/2 & 0 & 0 \\ 0 & 1 & -4 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2/3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] R_1 – 8R_3,\ R_2 + 4R_3$$

$$» \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & -19/3 & 1 & -3/2 & -8 & 0 \\ 0 & 1 & 0 & 11/3 & 0 & 1 & 4 & 0 \\ 0 & 0 & 1 & 2/3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$R_1 + \frac{19}{3}R_4,\ R_2 - \frac{11}{3}R_4,\ R_3 - \frac{2}{3}R_4$$

$$\gg \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -3/2 & -8 & 19/3 \\ 0 & 1 & 0 & 0 & 0 & 1 & 4 & -11/3 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & -2/3 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

Hence

$$A^{-1} = \begin{bmatrix} 1 & -3/2 & -8 & 19/3 \\ 0 & 1 & 4 & -11/3 \\ 0 & 0 & 1 & -2/3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

which is the inverse of the given upper triangular matrix.

Note that in Example 2, 3, 6 and 7, the inverse of a lower/upper triangular matrix is again a lower/upper triangular matrix. There is another method of finding the inverse of a matrix A which uses the pivoting strategy. Recall that in Sec. 3.4 of Unit 1, for the solution of system of linear algebraic equation Ax = b, we showed you how the multipliers $m_{p,i,k}$'s can be stored in working array W during the process of elimination. The main advantage of storing these multipliers is that if we have already solved the linear system of equations Ax = b or order n, by the elimination method and we want to solve the system Ax = c with the same coefficient matrix A, only the right side being different, then we do not have to go through the entire elimination process again. Since we have saved in the working matrix W all the multipliers used and also have saved the p vector, we have only to repeat the operations on the right hand side to obtain β, such that Ux = β is equaivalent to Ax = c.

In order to understand the calculations necessary to derive β , from c consider the changes made in the right side b during the elimination process. Let k be an integer between 1 and n, and assume that the ith equation was used as pivotal equation during step k of the elimination process. Then $i = p_k$. initially, the right side of equation i is just $b_i$.

If k > 1, then after Step 1, the right side is
$b_i^{(1)} = b_i - m_{i1}\, b_{p_1}$

If k > 2, then after Step 2, the right side is
$b_i^{(2)} = b_i^{(1)} = m_{i2}\, b_{p_2}^{(1)}$
$= b_i - m_{i1}\, b_{p1} - m_{i2}\, b_{p_2}^{(1)}$

In the same manner, we have the right side of equation $i = p_k$ as
$$b_i^{(k-1)} = b_i - m_{i1}\, b_{p1} - m_{i2}\, b_{p_2}^{(1)} - \ldots.. - m_{i,k-1}\, b_{p_{k-1}}^{(k-2)} \qquad (6)$$

Replacing i by $p_k$ in Eqn. (6), we get

$$b_{p_k}^{(k-1)} = b_{p_k} - m_{p_{k'1}} b_{p1} - m_{p_{k'2}} b_{p_2}^{(1)} - .... - m_{p_{k'k-1}} b_{p_{k-1}}^{(k-2)} \qquad (7)$$

$$k = 1, 2, ....., n.$$

Also, since $\tilde{b}_j = b_{p_j}^{(j-1)}$, j = 1, 2, ...., n, we can rewrite Eqn. (7) as

$$\tilde{b}_k = \tilde{b}_{p_k} - m_{p_{k,1}} \tilde{b}_1 - m_{p_{k,2}} \tilde{b}_2 - ..... - m_{p_{k,k-1}} \tilde{b}_{k-1} \qquad (8)$$

$$k = 1, ...., n.$$

Eqn. (8) can then be used to calculate the entries of $\tilde{b}$. But since the multipliers $m_{ij}$'s are stored in entries $w_{ij}$'s of the working matrix W, we can also write Eqn. (8) in the form

$$\tilde{b}_k = \tilde{b}_{p_k} - \sum_{j=1}^{k-1} W_{pkj} \tilde{b}_j, \; k = 1, ...., n \qquad (9)$$

Hence, if we just know the final content of the first n columns of W and the pivoting strategy p then we can calculate the solution x of Ax = b by using the back substitution method and writing

$$x_k = \frac{\tilde{b}_k - \sum_{j=k+1}^{n} W_{p_k j} x_j}{W_{p_k k}}, \; k = n, n-1, ....., 1 \qquad (10)$$

The vector x = $[x_1 \; x_2 \; ....... \; x_n]^T$ will then be the solution of Ax = b.

For finding the inverse of an $n \times n$ matrix A, we use the above algorithm. We first calculate the final contents of the n columns of the working matrix W and the pivoting vector p and then solve each of the n systems

$$Ax = e_j, \; j = 1, ......., n \qquad (11)$$

where $e_1 = [1 \;\; 0 \; ...... \; 0]^T$, $e_2 = [0 \;\; 1 \;\; 0 \; ..... \; 0]^T$, ......., $e_n = [0 \;\; 0 \; ...... \; 1]^T$, with the help of Eqn. (9) and (10). Then for each j = 1, ......, n the solution of system of system (11) will be the corresponding column of the inverse matrix $A^{-1}$. The following example will help you to understand the above procedure.

**Example 8**: Find the inverse of the matrix

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix}$$

using partial pivoting.

**Solution**: Initially $p = [p_1, p_2, p_3]^T = [1, 2, 3]^T$ and the working matrix is

$$W^{(0)} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix}$$

Now $d_1 = 2$, $d_2 = 2$, $d_3 = 2$.

Step 1: $\dfrac{|\,W_{p_{1,1}}\,|}{d_1} = \dfrac{1}{2}$, $\dfrac{|\,W_{p_{2,1}}\,|}{d_2} = \dfrac{2}{2} = 1$, $\dfrac{|\,W_{p_{3,1}}\,|}{d_3} = \dfrac{1}{2}$

$1 > \dfrac{1}{2}, \dfrac{1}{2}$ $\backslash$ $p_1 = 2, p_2 = 1, p_3 = 3$

We use the second equation to eliminate $x_1$ from first and third equations and store corresponding multipliers instead of storing zeros in the working matrix. The multipliers are

$$m_{p_{i,1}} = \frac{w_{p_{i,1}}}{w_{p_{i,1}}},\ i = 2, 3$$

$$\backslash \quad m_{p_{2,1}} = m_{11} = \frac{w_{p_{2,1}}}{w_{p_{1,1}}} = \frac{1}{2}$$

$$m_{p_{3,1}} = m_{31} = \frac{w_{p_{3,1}}}{w_{p_{1,1}}} = -\frac{1}{2}$$

we get the following working matrix

$$W^{(1)} = \begin{bmatrix} 1/2 & 3/2 & -1 \\ \boxed{2} & 1 & 0 \\ 1/2 & 3/2 & 2 \end{bmatrix},\ p = (2, 1, 3)^T$$

Step 2: $\dfrac{|w_{p_{2,2}}|}{dp_2} = \dfrac{|w_{p_{1,2}}|}{d_1} = \dfrac{3/2}{2} = \dfrac{3}{4}$

$\dfrac{|w_{p_{3,2}}|}{dp_3} = \dfrac{|w_{p_{3,2}}|}{d_3} = \dfrac{3/2}{2} = \dfrac{3}{4}$

Since $\dfrac{3}{4} = \dfrac{3}{4}$ so we take $p = (2, 1, 3)^T$

Now $m_{p_{i,2}} = \dfrac{w_{p_{i,2}}}{w_{p_{2,2}}}$, $i = 3$

$$\backslash \quad m_{p_{3,2}} = m_{32} = \frac{w_{p_{3,2}}}{w_{p_{1,2}}} = \frac{3/2}{3/2} = 1$$

We use the first equation as pivotal equation to eliminate $x_2$ from the third equation and also store the multipliers. After the second step we have the following working matrix

$$W^{(2)} = \begin{bmatrix} 1/2 & \boxed{3/2} & -1 \\ \boxed{2} & 1 & 0 \\ \enclose{circle}{1/2} & \enclose{circle}{1} & 3 \end{bmatrix}, \ p = (2, 1, 3)^T$$

Now in this case, $w^{(2)}$ is our final working matrix with pivoting strategy $p = (2, 1, 3)^T$

Note that circled ones denote multipliers and squared ones denote pivot elements in the working matrices.

To find the inverse of the given matrix A, we have to solve

$Ax = e_1 = [b_1\ b_2\ b_3]^T$
$Ax = e_2 = [b_1\ b_2\ b_3]^T$
$Ax = e_3 = [b_1\ b_2\ b_3]^T$
where $e_1 = [1\ 0\ 0]^T$, $e_2 = [0\ 1\ 0]^T$, $e_3 = [0\ 0\ 1]^T$

First we solve the system $Ax = e_1$ and consider

$$\begin{bmatrix} \enclose{circle}{1/2} & \boxed{3/2} & -1 \\ \enclose{circle}{2} & 1 & 0 \\ \enclose{circle}{1/2} & \enclose{circle}{1} & 3 \end{bmatrix} \begin{bmatrix} x \\ x \\ x \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ p = (2, 1, 3)^T \qquad (12)$$

Using Eqn. (9), we get
with $p_1 = 2, \tilde{b}_1 = b_2 = 0$
with $p_2 = 1$, $\tilde{b}_2 = b_1 - w_{11}\tilde{b}_1$
$= 1-\left[\frac{1}{2}\right]0$
$= 1$
with $p_3 = 3$, $\tilde{b}_3 = b_3 - w_{31}\tilde{b}_1 - w_{32}\tilde{b}_2$
$= 0-\left[\frac{1}{2}\right].0-1.1=-1$

Using Eqn. (10), we then get the following system of equations
$3x_3 = -1$
$\frac{3}{2}x_2 - x_3 = 1$
$2x_1 + x_2 = 0$

which gives $x_3 = -\frac{1}{3}$, $x_2 = \frac{4}{9}$ and $x_1 = -\frac{2}{9}$

i.e., vector x = $\left[\frac{1}{2}.\frac{4}{9}\ \frac{1}{3}\right]^T$ is the solution of system (12).

Remember that the solution of system (12) constitutes the first column of the inverse matrix $A^{-1}$.

In the same way we solve the system of equations Ax = $e_2$ and Ax = $e_3$, or

$$\begin{bmatrix} \enclose{circle}{1/2} & \boxed{3/2} & -1 \\ \boxed{2} & \enclose{circle}{1} & 0 \\ \enclose{circle}{1/2} & 1 & 3 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \text{p} = (2, 1, 3)^T \qquad (13)$$

and

$$\begin{bmatrix} 1/2 & 3/2 & -1 \\ 2 & 1 & 0 \\ 1/2 & 1 & 3 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \text{p} = (2, 1, 3)^T \qquad (14)$$

Using Eqns (9) and (10), we obtain the solution of system (13) as
x = $\left[\frac{5}{9}.\frac{1}{9}\frac{1}{3}\right]^T$ which is the second column of $A^{-1}$ and the solution of system (14), i.e.,
x = $\left[\frac{5}{9}.\frac{1}{9}\frac{1}{3}\right]^T$ as the third column of $A^{-1}$

Hence $A^{-1} = \begin{bmatrix} 2/9 & 5/9 & -1/9 \\ 4/9 & -1/9 & 2/9 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$

You may recall that in Sec. 3.5 of Unit 1 we discussed the LU decomposition method. Using this method we can factorise any non-singular square matrix A into the product of a lower triangular matrix L and upper triangular matrixU. That is, we can write

A = LU. ... (15)

In the next section we shall discuss how form (15) can be used to find the inverse of non-singular square matrices.

## 3.3 L U Decomposition Method

Let us consider Eqn. (15) and take the inverse on both the sides. If we use the fact that the inverse of the product of matrices is the product of their inverses takes in reverse order, then we obtain

$$A^{-1} = (L\,U)^{-1} = U^{-1}\,L^{-1} \qquad (16)$$

We can now find the inverse of U and L separately and obtain the inverse matrix $A^{-1}$ from Eqn. (16).

Remark: It may appear to you that finding an inverse of a matrix by this method is a lengthy process. But, in practice, this method is very useful because of the fact that here we deal with triangular matrices and triangular matrices are easily invertible. It involves only forward and backward substitutions.

Let us now consider an example to understand how the method works.

**Example 9**: Find the inverse of the matrix

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 2 & -3 & -1 \\ 1 & -2 & 1 \end{bmatrix}$$

using LU decomposition method.

**Solution**: We write,

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 2 & -3 & -1 \\ 1 & -2 & 1 \end{bmatrix} = LU = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & u & u \\ 0 & 1 & u \\ 0 & 0 & 1 \end{bmatrix} \qquad (17)$$

Comparing the coefficients on both sides of Eqn. (17), we obtain

$1_{11} = 3,\ 1_{21} = 2,\ 1_{31} = 1$ (multiplying the rows of L by the first column of U)

$1_{11}u_{12} = 1,\ u_{12} = \frac{1}{3}$ (multiplying the rows of L by the

$1_{11}u_{13} = 2,\ u_{13} = 2/3$ second and third column of U)

The second column of L is obtained from

$1_{21}u_{12} + 1_{22} = a_{22},\ 1_{22} = -3 - \frac{2}{3} = -\frac{11}{3}$

$1_{31}u_{12} + 1_{32} = a_{32},\ 1_{32} = -2 - \frac{1}{3} = -\frac{7}{3}$

$u_{23}$ is obtained from

$1_{21}u_{13} + 1_{22}u_{23} = a_{23},\ u_{23} = \dfrac{-\ 1 -\ 2(2/\ 3)}{-\ 11/\ 3} = \dfrac{7}{11}$

$1_{33}$ is obtained from

$1_{31}u_{13} + 1_{32}u_{23} + 1_{33} = 1,\ 1_{33} = \dfrac{20}{11}$

Thus we have

$$L = \begin{bmatrix} 3 & 0 & 0 \\ 2 & -11/3 & 0 \\ 1 & -7/3 & 20/11 \end{bmatrix} \text{and } U = \begin{bmatrix} 1 & 1/3 & 2/3 \\ 0 & 1 & 7/11 \\ 0 & 0 & 1 \end{bmatrix}$$

Now since L is a lower triangular matrix $L^{-1}$ is also a lower triangular matrix. Let us assume that

$$L^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Using the identity $LL^{-1}$, we have

$$LL^{-1} = \begin{bmatrix} 3 & 0 & 0 \\ 2 & -11/3 & 0 \\ 1 & -7/3 & 20/11 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 2 & -11/3 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

comparing the coefficients, we get

$$1'_{11} = \frac{1}{3},\ 1'_{22} = -\frac{3}{11},\ 1'_{33} = \frac{11}{20}$$

Also,

$$2\ 1'_{11} - \frac{11}{3} 1'_{21} = 0,\ 1'_{21} = \frac{6}{33} = \frac{2}{11}$$

$$1'_{11} - \frac{7}{3} 1'_{21} + \frac{20}{11} 1'_{31} = \frac{1}{20}$$

$$-\frac{7}{3} 1'_{22} + \frac{20}{11} 1'_{32} = 0,\ 1'_{32} = -\frac{7}{20}$$

$$\backslash \quad L^{-1} = \begin{bmatrix} 1/3 & 0 & 0 \\ 2/11 & -3/11 & 0 \\ 1/20 & -7/20 & 11/20 \end{bmatrix}$$

Similarly, since U is an upper triangular matrix, $U^{-1}$ is also upper triangular matrix. Using $UU^{-1} = I$, we obtain by backward substitution.

$$U = \begin{bmatrix} 1 & 1/3 & 2/3 \\ 0 & 1 & 7/11 \\ 0 & 0 & 1 \end{bmatrix} \text{and } U^{-1} = \begin{bmatrix} 1 & -1/3 & -5/11 \\ 0 & 1 & -7/11 \\ 0 & 0 & 1 \end{bmatrix}$$

Therefore, we have from Eqn. (16)

$$A^{-1} = U^{-1} L^{-1} = \begin{bmatrix} 1 & -1/3 & -5/11 \\ 0 & 1 & -7/11 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/3 & 0 & 0 \\ 2/11 & -3/11 & 0 \\ 1/20 & -7/20 & 11/20 \end{bmatrix}$$

$$= \begin{bmatrix} 1/4 & 1/4 & -1/4 \\ 3/20 & -1/20 & -7/20 \\ 1/20 & -7/20 & 11/20 \end{bmatrix}$$

## 4.0 CONCLUSION

We now end this unit by giving a summary of what we have covered init.

## 5.0 SUMMARY

In this unit we have covered the following:

a) Using the method of adjoints, the inverse of a given non-singular matrix A can be obtained from

$$A^{-1} = \frac{1}{\det(A)} \text{adj}(A) \qquad \text{(see Eqn. (1))}$$

Since the number of operations in the adjoint method to find the inverse of n ´ nnon-singular matrix A increases rapidly as n increases, the method is not generally used for n > 4.

b) For large n, the Gauss-Jordan reduction method, which is an extension of the Gauss elimination method can be used for finding the inverse matrix and solve the linear systems.
$Ax = b$ (see Eqn. (2))
using the Gauss-Jordan method.

a) the solution of system of Eqns (2) can be obtained by using elementary now operations
$[A|b] \overset{\text{reduced to}}{\text{¾¾¾¾¾®}} [I|d]$

b) the inverse matrix $A^{-1}$ can be obtained by using elementary row operations $[A|I] \overset{\text{reduced to}}{\text{¾¾¾¾¾®}} [I|A^{-1}]$

c) For large n, another useful method of finding the inverse matrix $A^{-1}$ is LU decomposition method. Using this method any non-singular matrix A is first decomposed into the product of a lower triangular matrix L and an upper triangular matrixU. That is

$A = LU$
$U^{-1}$ and $L^{-1}$ can be obtained by backward and forward substitutions. Then the inverse can be found from
$A^{-1} = U^{-1} L^{-1}$

## 6.0 TUTOR-MARKED ASSIGNMENT

i Solve the system of equations
$3x_1 + x_2 + 2x_3 = 3$
$2x_1 – x_2 – x_3 = 1$
$x_1 – 2x_2 + x_3 = -4$
using the method of adjoints.

ii Solve the system of equations

$$\begin{bmatrix} 2 & 3 & 4 & 1 \\ 1 & 2 & 0 & 1 \\ 2 & 3 & 1 & -1 \\ 1 & -2 & -1 & 4 \end{bmatrix} \begin{bmatrix} X1 \\ X2 \\ X3 \\ X4 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 5 \end{bmatrix}$$

using the method of adjoints.

iii. Verify that the total number of operations needed for Gauss-Jordan reduction methods is $\frac{1}{2}n^3 + \frac{n^2}{2} + n$.

iv In example 6 and 7 verify that
$A A^{-1} = A^{-1} A = I$.

v Solve the system of equation
$x_1 + 2x_2 + x_3 = 0$
$2x_1 + 2x_2 + 3x_3 = 3$

$-x_1 - 3x_2 = 2$
using the Gauss-Jordan method with pivoting.

vi Find the inverse of the matrix

$$A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

using the Gauss-Jordan method.

vii Find the inverse of the matrix

$$A = \begin{bmatrix} 5 & 8 & 1 \\ 0 & 2 & 1 \\ 4 & 3 & -1 \end{bmatrix}$$

using the LU decomposition method.

viii Find the inverse of the matrix

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 2 & -1 & -1 \\ 1 & -2 & 1 \end{bmatrix}$$

Using the LU decomposition method.

## 7.0 REFERENCES/FURTHER READINGS

Wrede, R.C. and Spegel M. (2002). Schaum's and Problems of Advanced Calculus. McGraw – Hill N.Y.

Keisler, H.J. (2005). Elementary Calculus. An Infinitesimal Approach. 559 Nathan Abbott, Stanford, California, USA

# UNIT 3 ITERATIVE METHODS

## CONTENTS

## 1.0 INTRODUCTION

In the previous two units, you have studied direct methods for solving linear system of equations Ax = b, A being n ´ n non-singular matrix. Direct methods provide the exact solution in a finite number of steps provided exact arithmetic is used and there is no round-off error. Also, direct methods are generally used when the matrix A is dense or filled, that is, there are few zero elements, and the order of the matrix is not very large say $n < 50$.

Iterative methods, on the other hand, start with an initial approximation and by applying a suitably chosen algorithm, lead to successively better approximations. Even if the process converges, it would give only an approximate solution. These methods are generally used when the matrix A is sparse and the order of the matrix A is very large say $n > 50$. Sparse matrices have very few non-zero elements. In most cases these non-zero elements lie on or near the main diagonal giving rise to tri-diagonal, five diagonal or band matrix systems. It may be noted that there are no fixed rules to decide when to use direct methods and when to use iterative methods. However, when the coefficient matrix is sparse or large, the use of iterative methods is ideally suited to find the solution which take advantage of the sparse nature of the matrix involved.

In this we shall discuss two iterative methods, namely, Jacobi iteration and Gauss-Seidel iteration methods which are frequently used for solving linear system of equations.

## 2.0 OBJECTIVES

At the end of this unit, you should be able to:

- obtain the solution of system of linear equations, Ax = b, when the matrix A is large or sparse, by using the iterative method viz; Jacobi method or the Gauss-Seidel method
- tell whether these iterative methods converges or not
- obtain the rate of convergence and the approximate number of iterations needed for the required accuracy of these iterative methods.

## 3.0 MAIN CONTENT

### 3.1 The General Iteration Method

In iteration methods as we have already mentioned, we start with some initial approximate solution vector $x^{(0)}$ an generate a sequence of approximation $\{x^{(k)}\}$ which converge to the exact solution vector x as k® ¥ . If the method is convergent, each iteration produces a better approximation to the exact solution. We repeat the iterations till the required accuracy is obtained. Therefore, in an iterative method the amount of computation depends on the desired accuracy whereas in direct methods the amount of computation is fixed. The number of iterations needed to obtain the desired accuracy also depends on the initial approximation, closer the initial approximation to the exact solution, faster will be the convergence.

Consider the system of equations

$$Ax = b \qquad \text{... (1)}$$

where A is an n´ n non-singular matrix.

Writing the system in expanded form, we get

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots\ldots\ a_{1n}x_n = b_1\\
&a_{21}x_1 + a_{22}x_2 + \ldots\ldots\ a_{2n}x_n = b_2\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{n1}x_1 + a_{n2}x_2 + \ldots\ldots + a_{nn}x_n = b_n
\end{aligned} \qquad (2)$$

We assume that the diagonal coefficients $a_{ii}{}^{1}$  0, (i = 1, ....., n). If some of $a_{ii}$ = 0, then we arrange the equations so that this condition holds. We then rewrite system (2) as

$$\begin{aligned}
x_1 &= -\frac{1}{a_{11}}(a_{12}x_2 + a_{13}x_3 + \ldots. + a_{1n}x_n) + \frac{b_1}{a_{11}}\\
x_2 &= -\frac{1}{a_{22}}(a_{21}x_1 + a_{23}x_3 + \ldots. + a_{2n}x_n) + \frac{b_2}{a_{22}}\\
\\
x_n &= -\frac{1}{a_{nn}}(a_{n1}x_1 + a_{n2}x_2 + \ldots. + a_{nn-1}x_{n-1}) + \frac{b_n}{a_{nn}}
\end{aligned} \qquad (3)$$

In matrix form, system (3) can be written as
$x = Hx + c$
where

$$H = \begin{bmatrix} 0 & \frac{-a12}{a11} & \frac{-a13}{a11} & \frac{-a1n}{a11} \\ \frac{a21}{a22} & 0 & \frac{-a23}{a22} & \frac{-a2n}{a22} \\ \hdashline \frac{an1}{ann} & \frac{-an2}{ann} & \frac{-an,n-1}{ann} & 0 \end{bmatrix} \qquad (4)$$

and the elements of c are $c_i = \dfrac{b_i}{a_{ii}}$ (i = 1, 2, ..., n)

To solve system (3) we make an initial guess $x^{(0)}$ of the solution vector and substitute into the r.h.s. of Eqn. (3). The solution of Eqn. (3) will then yield a vector $x^{(1)}$, which hopefully is a better approximation to the solution than $x^{(0)}$. We then substitute $x^{(1)}$ into the r.h.s. of Eqn. (3) and get another approximation, $x^{(2)}$. We continue in this manner until the successive iterations $x^{(k)}$ have converged to the required number of significant figures.

In general we can write the iteration method for solving the linear system of Eqns. (1) in the form

$$x^{(k+1)} = Hx^{(k)} + c, \ k = 0, 1...... \qquad (5)$$

where $x^{(k)}$ and $x^{(k+1)}$ are the approximations to the solution vector x at the kth and the (k + 1)th iterations respectively. H is called the iteration matrix and depends on A. c is a column vector and depends on both A and b. The matrix H is generally a constant matrix.

When the method (5) is convergent, then
$\lim\limits_{\text{k® ¥}} x^{(k)} = \lim\limits_{\text{k® ¥}} x^{(k+1)} = x$
and we obtain from Eqn. (5)
$x = Hx + c \qquad (6)$
If we define the error vector at the kth iteration as
$\hat{\mathrm{I}}^{(k)} = x^{(k)} - x \qquad (7)$
then subtracting Eqn. (6) from Eqn. (5), we obtain
$\hat{\mathrm{I}}^{(k+1)} = H\,\hat{\mathrm{I}}^{(k)} \qquad (8)$
Thus, we get from Eqn. (8)

$$\hat{\mathrm{I}}^{(k)} = H\,\hat{\mathrm{I}}^{(k-1)} = H^2\hat{\mathrm{I}}^{(k-2)} = ... = H^k\hat{\mathrm{I}}^{(0)} \qquad (9)$$

where $\hat{\mathrm{I}}^{(0)}$ is the error in the initial approximate vector. Thus, for the convergence of the iterative method, we must have

$\lim\limits_{\text{k® ¥}} \hat{\mathrm{I}}^{(k)} = 0$
independent of $\hat{\mathrm{I}}^{(0)}$.

Before we discuss the above convergence criteria, let us recall the following definitions from linear algebra.

**Definition**: For a square matrix A of order n, and a number l the value of l for which the vector equation Ax = l x has non-trivial solution x $^{1}$ 0, is called an eigenvalue or characteristic value of the matrix A.

**Definition**: The largest eigenvalue in magnitude of A is called the spectral radius of A ad is denoted by p(A).

The eigenvalues of the matrix A are obtained from the characteristic equation
det(A - l I) = 0
which is an nth degree polynomial in l . The roots of this polynomial l $_1$, l $_2$, .......,l $_n$ are the eigenvalues of A. Therefore, we have

$$r\,(A) = \max_i |l_{\ i}| \qquad (10)$$

We now state a theorem on the convergence of the iterative methods.

**Theorem 1**: An iteration method of the form (5) is convergent for arbitrary initial approximate vector $x^{(0)}$ if and only if r (H) < 1.

We shall not be proving this theorem here as its proof makes use of advanced concepts from linear algebra and is beyond the scope of this course.

We define the rate of convergence as follows:

**Definition**: The number n = -log$_{10}$r (H) is called the rate of convergence of an iteration method.

Obviously, smaller the value of r (H), larger is the value of n .

**Definition**: The method is said to have converged to m significant digits if $\max_i |Î_{\ i}^{(k)}|$ , $10^{-m}$, that is, largest element in magnitude, of the error vector Î $^{(k)}$, $10^{-m}$. Also the number of iterations k that will be needed to make $\max_i |Î_{\ i}^{(k)}|$ , $10^{-m}$is given by

$$k = \frac{m}{n} \qquad (11)$$

Therefore, the number of iterations that are required to achieved the desired accuracy depends onn . For a method having higher rate of convergence, lesser number of iterations will be needed for a fixed accuracy and fixed initial approximation.

There is another convergence criterion for iterative methods which is based on the norm of a matrix.

The norm of a square matrix A of order n can be defined in the same way as we define the norm of an n-vector by comparing the size of Ax with the size of x (an n-vector) as follows:

i) $\|A\|_2 = \max \frac{\|Ax\|_2}{\|x\|_2}$

based on the Euclidean vector norm, $\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + ... + |x_n|^2}$

and

ii) $\|A\|_{¥} = \max \frac{\|Ax\|_{¥}}{\|x\|_{¥}}$, based on the maximum vector norm, $\|x\|_{¥} = \max\limits_{1£ i£ n} |x_i|$.

In (i) and (ii) above the maximum is taken over all (non zero) n-vector. The most commonly used norms is the maximum norm $\|A\|_{¥}$ , as it is easier to calculate. It can be calculated in any of the following two ways:

$\|A\|_{¥} = \max\limits_{x} \underset{i}{å} |a_{ik}|$ (maximum absolute column-sum)

or

$\|A\|_{¥} = \max\limits_{i} \underset{k}{å} |a_{ik}|$ (maximum absolute row sum)

The norm of a matrix is a non-negative number which in addition to the property $\|AB\| , \|A\| \|B\|$

satisfies all the properties of a vector norm, viz.,

a) $\|A\| ƒ 0$ and $\|A\| = 0$ if $A = 0$

b) $\|a A\| = |a| \|A\|$, for all numbers a .

c) $\|A + B\| , \|A\| + \|B\|$

where A and B are square matrices of order n.

We no state a theorem which gives the convergence criterion for iterative methods in terms of the norm of a matrix.

**Theorem 2**: The iteration method of the form (5) for the solution of system (1) converges to the exact solution for any initial vector, if $\|H\|<1$.

Also note that

$\|H\| ƒ r (H)$.

This ca be easily proved by considering the eignevalue problem $Ax = 1 x$.

Then $\|A\| = \|1 x\| = |1| \|x\|$

or $|l|\ \|x\| = \|Ax\|\ ,\ \ \|A\|\ \|x\|$
i.e., $|l|\ ,\ \ \|A\|$ since $\|x\|\ {}^{1}\ \ 0$

Since this results is true for all eignevalue, we have
$r\,(A)\ ,\ \ \|A\|$.

The criterion given in Theorem 2 is only a sufficient condition, it is not necessary. Therefore, for a system of equations for which the matrix H is such that either $\underset{i}{\max}\ \overset{n}{\underset{k=1}{å}}\ |h_{ik}| < 1$, the iteration always converges, but if the condition is violated it is not necessary that the iteration diverges.

There is another sufficient condition for convergence as follows:

**Theorem 3**: If the matrix A is strictly diagonally dominant that is,

$$|a_{ii}| > \overset{n}{\underset{\substack{j=1\\ j^{1}\,i}}{å}}\ |a_{ij}|,\ i = 1, 2, \ldots\ldots, n.$$

Then the iteration method (5) converges for nay initial approximation $x_{10}$. If no better initial approximation is known, we generally take
$x^{(0)} = 0$.
We shall mostly use the criterion given in Theorem 1, which is both necessary and sufficient.

For using the iteration method (5), we need the matrix H and the vector c which depend on the matrix A and the vector b. the well-known iteration methods are based on the splitting of the matrix A in the form

$$A = D + L + U \qquad (12)$$

where D is the diagonal matrix, L and U are respectively the lower and upper triangular matrices with zero diagonal elements. Based on the splitting (12), we now discuss two iteration methods of the form (5).

## 3.2 The Jacobi's Iteration Method

We write the system of Eqn. (1) in the form (2), viz.,

$$\begin{array}{l}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1\\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2\\
.\qquad .\qquad\quad .\qquad .\\
.\qquad .\qquad\quad .\qquad .\\
.\qquad .\qquad\quad .\qquad .\\
a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n
\end{array}$$

We assume that $a_{11}$, $a_{22}$, ..... $a_{nn}$ are pivot elements and $a_{ii}{}^{1}$ 0, i = 1, 2, ...., n. if any of the pivots is zero, we can interchange the equations to obtain non-zero pivots (partial pivoting).

Note that, A being a non-singular matrix, it is possible for us to make all the pivots non-zero. It is only when the matrix A is singular that even complete pivoting may not lead to all the non-zero pivots.

We rewrite system (2) in the form (3) and define the Jacobi iteration method as

$$x_1^{(k+1)} = -\frac{1}{a_{11}}(a_{12}x_2^{(k)} + a_{13}x_3^{(k)} + \ldots + a_{1n}x_n^{(k)} - b_1)$$

$$x_2^{(k+1)} = -\frac{1}{a_{22}}(a_{21}x_2^{(k)} + a_{23}x_3^{(k)} + \ldots + a_{2n}x_n^{(k)} - b_2)$$

.

.

.

$$x_n^{(k+1)} = -\frac{1}{a_{ii}}(a_{n1}x_i^{(k)} + a_{n2}x_2^{(k)} + \ldots + a_{n,n-1}x_{n-1}^{(k)} - b_n)$$

$$\text{or } x_i^{(k+1)} = -\frac{1}{a_{ii}},\ (a_{n1}x_i^{(k)} + a_{n2}x_2^{(k)} + \ldots + a_{n,n-1}x_{n-1}^{(k)} - bn) \qquad (13)$$

The method (13) can be put in the matrix form as

$$\begin{bmatrix} x_1^{(k\ 1)} \\ x_1^{(k\ 1)} \\ \\ x_1^{(k\ 1)} \end{bmatrix} = - \begin{bmatrix} \frac{1}{a11} & & & \\ & \frac{1}{a22} & & \\ & & \cdots & \\ & & & \frac{1}{ann} \end{bmatrix} \begin{matrix} 0 & a_{12} & \ldots & a_{1n} \\ a_{12} & 0 & \ldots. & a_{2n} \\ & & \ldots. & \\ a_{n1} & a_{n2} & \ldots & 0 \end{matrix} \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \\ x_n^{(k)} \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ \\ b_n \end{bmatrix} r$$

$$x^{(k+1)} = -D^{-1}(L + U)\, x^{(k)} + D^{-1}b, \ k = 0, 1, \ldots. \qquad (14)$$

where

$$D = \begin{bmatrix} a_{11} & 0 & \text{..........} & 0 \\ 0 & a_{22} & \text{..........} & 0 \\ 0 & \text{..........} & & a_{nn} \end{bmatrix}, L = \begin{bmatrix} 0 & 0 & \text{..........} & 0 \\ a_{21} & 0 & \text{..........} & 0 \\ a_{31} & a_{32} & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n,n-1} & 0 \end{bmatrix}$$

$$\text{and } U = \begin{bmatrix} 0 & a_{12} & a_{13} & \text{..........} & a_{1n} \\ 0 & 0 & a_{23} & & a_{2n} \\ & \text{..........} & & & \\ & & & & a_{n-1,n} \\ 0 & 0 & 0 & \text{..........} & 0 \end{bmatrix}$$

The method (14) is for the form (5), where

$H = -D^{-1} (L + U)$ and $c = D^{-1}b$

For computation purpose, we obtain the solution vector $x^{(k+1)}$ at the (k + 1)th iteration, element by element using Eqn. (13). For large n, we rarely use the method in its matrix form as given by Eqn. (14).

In this method in the (k + 1)th iteration we use the values, obtained at the kth iteration viz., $x_1^{(k)}, x_2^{(k)}, ...., x_n^{(k)}$ on the right hand side of Eqn. (13) and obtain the solution vector $x^{(k+1)}$. We then replace the entire vector $x^{(k)}$ on the right side of Eqn. (13) by $x^{(k+1)}$ to obtain the solution at the next iteration. In other words each of the equations is simultaneously changed by using the most recent set of x-values. It is for this reason this method is also known as the method of simultaneous displacements.

Let us now solve a few examples for better understanding of the method and its convergence.

**Example 1**: Perform four iterations of the Jacobi method for solving the system of equations

$$\begin{bmatrix} 8 & 1 & 1 \\ 1 & -5 & 1 \\ 1 & 1 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 16 \\ 7 \end{bmatrix} \qquad (15)$$

with $x^{(0)} = 0$, the exact solution is $x = [-1\ -4\ -3]^T$.

**Solution**: The Jacobi method when applied to the system of Eqns. (15) becomes

$$x_1^{(k+1)} = \frac{1}{8}[x_2^{(k)} + x_3^{(k)} - 1]$$

$$x_2^{(k+1)} = \frac{1}{5}[x_1^{(k)} + x_3^{(k)} - 16] \qquad (16)$$

$$x_3^{(k+1)} = \frac{1}{4}[x_1^{(k)} + x_2^{(k)} - 7], k = 0, 1, ....$$

Starting with $x^{(0)} = [0\ 0\ 0]^T$, we obtain form Eqns. (16), the following results:

k = 0

$$x_1^{(1)} = \frac{1}{8}[0 + 0 - 1] = -0.125$$

$$x_2^{(1)} = \frac{1}{5}[0 + 0 - 16] = -3.2$$

$$x_3^{(1)} = \frac{1}{4}[0 + 0 - 7] = -1.75$$

k = 1

$$x_1^{(2)} = \frac{1}{8}[-3.2 - 1.75 - 1] = -0.7438$$

$$x_2^{(2)} = \frac{1}{5}[-0.125 - 1.75 - 16] = 3.5750$$

$$x_3^{(2)} = \frac{1}{4}[-0.125 - 3.2 - 7] = -2.5813$$

k = 2

$$x_1^{(3)} = \frac{1}{8}[-3.5750 - 2.5813 - 1] = -0.8945$$

$$x_2^{(3)} = \frac{1}{5}[-0.7438 - 2.5813 - 16] = -3.8650$$

$$x_3^{(3)} = \frac{1}{4}[-0.7438 - 3.5750 - 7] = 2.8297$$

k = 3

$$x_1^{(4)} = \frac{1}{8}[-3.8650 - 2.8297 - 1] = 0.9618$$

$$x_2^{(4)} = \frac{1}{5}[-0.8945 - 2.8297 - 16] = -3.9448 \qquad (17)$$

$$x_3^{(4)} = \frac{1}{4}[-0.8945 - 3.8650 - 7] = -2.9399$$

Thus, after four iterations we get the solution as given in Eqns (17). We find that after iteration, we get better approximation to the exact solution.

**Example 2**: Jacobi method is used to solve the system of equations

$$\begin{bmatrix} 4 & -1 & 1 \\ 4 & -8 & 1 \\ -2 & 1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 21 \\ 15 \end{bmatrix} \qquad (18)$$

Determine the rate of convergence of the method and the number of iterations needed to make $\max_{i} |Î_{i}^{(k)}|$ , $10^{-2}$

Perform these number of iteration starting with initial approximation $x^{(0)} = [1\ 2\ 2]^T$ and compare the result with the exact solution $[2, 4\ 3]^T$

**Solution**: The Jacobi method when applied to the system of Eqns. (18), gives the iteration matrix

$$H = -\begin{bmatrix} \begin{matrix} \frac{1}{a_{11}} & 0 & 0 \\ 0 & \frac{1}{a_{22}} & 0 \\ 0 & 0 & \frac{1}{a_{33}} \end{matrix} \left| \begin{matrix} 1 & a_{12} & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & 0 \end{matrix} \right. \end{bmatrix}$$

$$= -\begin{bmatrix} \begin{matrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{8} & 0 \\ 0 & 0 & \frac{1}{5} \end{matrix} \left| \begin{matrix} 0 & -1 & 1 \\ 4 & 0 & 1 \\ 2 & 1 & 0 \end{matrix} \right. \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1/4 & -1/4 \\ 1/2 & 0 & 1/8 \\ 2/5 & -1/5 & 0 \end{bmatrix}$$

The eignevalues of the matrix H are the roots of the characteristic equation.

det (H - 1 I) = 0

Now

$$\det (H - 1\ I) = \begin{bmatrix} 1 & 1/4 & -1/4 \\ 1/2 & -1 & 1/8 \\ 2/5 & -1/5 & -1 \end{bmatrix} = 1^{3} - \frac{3}{80} = 0$$

All the three eigenvalues of the matrix H are equal and they are equal to

$l = 0.3347$

The spectral radius is

$$r\,(H) = 0.3347 \qquad (19)$$

We obtain the rate of convergence as

$$n = -\log_{10}(0.3347) = 0.4753$$

The number of iterations needed for the required accuracy is given by

$$k = \frac{2}{n} \text{ » } 5 \qquad (20)$$

The Jacobi method when applied to the system of Eqns. (18) becomes

$$x^{(k+1)} = \begin{bmatrix} 0 & 1/4 & -1/4 \\ 1/2 & 0 & 1/8 \\ 2/5 & -1/5 & 0 \end{bmatrix} x^{(k)} + \begin{bmatrix} 7/4 \\ 21/8 \\ 3 \end{bmatrix}, k = 0, 1, ... \qquad (21)$$

starting with the initial approximation $x^{(0)} = [1\ 2\ 2]^T$, we get from Eqn. (21)

$x^{(1)} = [1.75 \quad 3.375 \quad 3.0]^T$
$x^{(2)} = [1.8437 \quad 3.875 \quad 3.025]^T$
$x^{(3)} = [1.9625 \quad 3.925 \quad 2.9625]^T$
$x^{(4)} = [1.9906 \quad 3.9766 \quad 3.0000]^T$
$x^{(5)} = [1.9941 \quad 3.9953 \quad 3.0009]^T$

which is the result after five iterations. Thus, you can see that result obtained after five iterations is quite close to the exact solution $[2\ 4\ 3]^T$

**Example 3**: Perform four iterations of the Jacobi method for solving the system of equations

$$\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \qquad (22)$$

With $x^{(0)} = [0.5 \quad 0.5 \quad 0.5 \quad 0.5]^T$. What can you say about the solution obtained if the exact solution is $x = [1 \quad 1 \quad 1 \quad 1]^T$?

**Solution**: The Jacobi method when applied to the system of Eqns. (22) becomes

$$\begin{aligned}
x_1^{(k+1)} &= \frac{1}{2}[1 + x_2^{(k)}] \\
x_2^{(k+1)} &= \frac{1}{2}[x_1^{(k)} + x_3^{(k)}] \\
x_3^{(k+1)} &= \frac{1}{2}[x_2^{(k)} + x_4^{(k)}] \qquad (23) \\
x_4^{(k+1)} &= \frac{1}{2}[1 + x_3^{(k)}],\ k = 0, 1, ....
\end{aligned}$$

Using $x^{(0)} = [0.5\ \ 0.5\ \ 0.5\ \ 0.5]^T$, we obtain

$x^{(1)} = [0.75\ \ 0.5\ \ 0.5\ \ 0.75]^T$
$x^{(2)} = [0.75\ \ 0.625\ \ 0.625\ \ 0.75]^T$
$x^{(3)} = [0.8125\ \ 0.6875\ \ 0.6875\ \ 0.8125]^T$
$x^{(4)} = [0.8438\ \ 0.75\ \ 0.75\ \ 0.8438]^T$

You may notice here that the solution is improving after each iteration. Also the solution obtained after four iterations is not a good approximation to the exact solution $x = [1\ \ 1\ \ 1\ \ 1]^T$. this shows that we require a few more iterations to get a good approximation.

**Example 4**: Find the spectral radius of the iteration matrix when the Jacobi method, is applied to the system of equations

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}$$

Verify that the iterations do not converge to the exact solution
$x = [1\ \ 3\ \ -1]^T$.

**Solution**: The iteration matrix H in this case becomes

$$H = -\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & -2 \\ 1 & -1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & -2 \\ 0 & 0 & 2 \\ 1 & 1 & 0 \end{bmatrix}$$

and $c = [-1\ 5\ -3]^T$

The eigenvalue of H are roots of the characteristic equation

det (H - l I) = 0. This gives us
-l (l $^2$ – 4) = 0
i.e., l = 0, ± 2
\ r (H) = 2 > 1.

Thus, the condition in Theorem 1 is violated. The iteration method does not converge.

We now perform few iteration and see whet happens actually. Taking $x^{(0)} = 0$ and using the Jacobi method

$$x^{(k+1)} = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 0 & 2 \\ 1 & 1 & 0 \end{bmatrix} x^{(k)} + \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}$$

we obtain

$x^{(1)} = (-1\ 5\ -3)^T$
$x^{(2)} = (5\ -1\ 3)^T$
$x^{(3)} = (-7\ 11\ -9)^T$
$x^{(4)} = (17\ -13\ 15)^T$
$x^{(5)} = (-31\ 35\ -33)^T$

and so on, which shows that the iterations are diverging fast. You may also try to obtain the solution with other initial approximations.

Let us now consider an example to show that the convergence criterion given in Theorem 3 is only a sufficient condition. That is, there are systems of equation which are not diagonally dominant but, the Jacobi iteration method converges.

**Example 5**: Perform iterations of the Jacobi method for solving the system of equations

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

With $x^{(0)} = [0 \quad 1 \quad 1]^T$. What can you say about the solution obtained if the exact solution is $x = [0\ 1\ 2]^T$?

**Solution**: The Jacobi method when applied to the given system of equations becomes

$$x_1^{(k+1)} = [3 - x_2^{(k)} - x_3^{(k)}]$$
$$x_2^{(k+1)} = 1$$

$$x_3^{(k+1)} = [-1 + 3x_2^{(k)}], k = 0, 1, ....$$

Using $x^{(0)} = [0\ 1\ 1]^T$, we obtain
$x^{(1)} = [1\ 1\ 2]^T$
$x^{(2)} = [0\ 1\ 2]^T$
$x^{(3)} = [0\ 1\ 2]^T$

You may notice here that the coefficient matrix is not diagonally dominant but the iterations converge to the exact solution after only two iterations.

We have already mentioned that iterative methods are usually applied to large linear system with a sparse coefficient matrix. For sparse matrices, the number of non-zero entries is small, and hence the number of arithmetic operations to be performed per step is small.

However, iterative methods may not always converge, and even when they converge, they may require a large number of iterations.

We shall now discuss the Gauss-Seidel method which is a simple modification of the method of simultaneous displacements and has improved rate of convergence.

## 3.3 The Gauss-Seidel Iteration Method

Consider the system of Eqns. (2) written in form (3). For this system of equations, we define the Gauss-Seidel method as:

$$\begin{aligned}
x_1^{(k+1)} &= -\frac{1}{a_{11}}(a_{12}x_2^{(k)} + a_{13}x_3^{(k)} + ... + a_{1n}x_n^{(k)} - b_1)\\
x_2^{(k+1)} &= -\frac{1}{a_{22}}(a_{21}x_1^{(k+1)} + a_{23}x_3^{(k)} + ... + a_{2n}x_n^{(k)} - b_2)\\
&.\\
&. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (24)\\
&.\\
x_n^{(k+1)} &= -\frac{1}{a_{nn}}(a_{n1}x_1^{(k+1)} + a_{n2}x_2^{(k+1)} + ... + a_{n,n-1}x_{n-1}^{(k+1)} - b_n)
\end{aligned}$$

$$\text{or } x_i^{(k+1)} = -\frac{1}{a_{ii}} \sum_{j1}^{n} a_{ij}x_j{}^{(k\ \ 1)} + \sum_{ji1}^{n} a_{ij}x_j{}^{(k)}\ b_i\ , i = 1, 2, .... n$$

You may notice here that in the first equation of system (24), we substitute the initial approximation ($x_2^{(0)}$, $x_3^{(0)}$, ...., $x_n^{(0)}$) on the right hand side. In the second equation w substitute ($x_1^{(1)}$, $x_3^{(0)}$, ...., $x_n^{(0)}$) on the right hand side. In the third equation, we substitute ($x_1^{(1)}$, $x_2^{(1)}$, $x_4^{(0)}$, ...., $x_n^{(0)}$) on the right hand side. We continue in this manner until all the components have been improved. At the end of this first iteration, we will have an improved vector ($x_1^{(1)}$, $x_2^{(1)}$, ...., $x_n^{(1)}$). The entire process is then repeated. In other words, the method uses an improved component as soon as it becomes available. It is for this reason the method is also called the method of successive displacements.

We can also write the system of Eqns. (24) as follows:

$$a_{11}x_1^{(k+1)} = -a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \dots a_{1n}x_n^{(k)} + b_1$$

$$a_{21}x_2^{(k+1)} + a_{21}x_2^{(k+1)} = - a_{23}x_3^{(k)} - \dots - a_{2n}x_n^{(k)} + b_2$$

.

.

.

$$a_{n1}x_1^{(k+1)} + a_{n2}x_2^{(k+1)} + \dots + a_{nn}x_n^{(k+1)}b_n$$

In matrix form, this system can be written as

$$(D + L)\, x^{(k+1)} = -U\, x^{(k)} + b \qquad (25)$$

where D is the diagonal matrix

$$D = \begin{bmatrix} a_{11} & & & 0 \\ 0 & a_{22} & & .. \\ & & a_{33} & .. \\ & & & .. \\ & & & \\ 0 & & & a_{nn} \end{bmatrix}$$

and L and U are respectively the lower and upper triangular matrices with the zeros along the diagonal and are of the form

$$L = \begin{bmatrix} 0 & 0 & 0 & \dots\dots & 0 & 0 \\ a_{21} & 0 & 0 & \dots\dots & & 0 \\ a_{31} & a_{32} & 0 & 0 & \dots\dots. & 0 \\ \dots\dots & & & & & .. \\ \dots\dots. & & & & & .. \\ a_{n1} & a_{n2} & & & \dots\dots & a_{nn} \end{bmatrix} \quad U = \begin{bmatrix} 0 & a_{12} & a_{13} & \dots.. & a_{1n} \\ 0 & 0 & a_{23} & \dots.. & a_{2n} \\ 0 & 0 & 0 & \dots.. & a_{3n} - \\ & & & & \dots.. \\ & & & & a_{n-1,n} \\ 0 & & & & 0 \end{bmatrix}$$

From Eqn. (25), we obtain

$$x^{(k+1)} = - (D + L)^{-1}\, Ux^{(k)} + (D + L)^{-1}b \qquad (26)$$

which is of the form (5) with

$H = -(D + L)^{-1}$ U and $c = (D + L)^{-1}b$.

It may again be noted here, that if A is diagonally dominant then the iteration always converges.

Gauss-Seidel method will generally converge if the Jacobi method converges, and will converge at a faster rate. For symmetric A, it can be shown that

r (Gauss-Seidel iteration method) = [r (Jacobi iteration method)]$^2$

Hence the rate of convergence of the Gauss-Seidel method is twice the rate of convergence of the Jacobi method. This result is usually true even when A is not symmetric.

We shall illustrate this fact through examples.

**Example 6**: Perform four iterations (rounded to four decimal places) using the Gauss-Seidel method for solving the system of equations

$$\begin{bmatrix} 8 & 1 & 1 \\ 1 & -5 & 1 \\ 1 & 1 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 16 \\ 7 \end{bmatrix} \qquad (27)$$

with $x^{(0)} = 0$. The exact solution is $x = (-1\ -4\ -3)^T$.

**Solution**: The Gauss-Seidel method, for the system (25) is

$$\begin{aligned} x_1^{(k+1)} &= \frac{1}{8}[x_2^{(k)} + x_3^{(k)} - 1] \\ x_2^{(k+1)} &= \frac{1}{5}[x_1^{(k+1)} + x_3^{(k+1)} - 16] \qquad (28) \\ x_3^{(k+1)} &= \frac{1}{4}[x_1^{(k+1)} + x_2^{(k+1)} - 7],\ k = 0, 1, .... \end{aligned}$$

Taking $x^{(0)} = 0$, we obtain the following iterations.

k = 0

$$\begin{aligned} x_1^{(1)} &= \frac{1}{8}[0 + 0 - 1] = -0.125 \\ x_2^{(1)} &= \frac{1}{5}[-0.125 + 0 - 16] = -3.225 \\ x_3^{(1)} &= \frac{1}{4}[-0.125 - 3.225 - 7] = -2.5875 \end{aligned}$$

k = 1

$$\begin{aligned} x_1^{(2)} &= \frac{1}{8}[-3.225 - 2.5875 - 1] = -0.8516 \\ x_2^{(2)} &= \frac{1}{5}[-0.8516 - 2.5875 - 16] = 3.8878 \end{aligned}$$

$$x_3^{(2)} = \frac{1}{4}[-0.8516 - 3.8878 - 7] = -2.9349$$

k = 2

$$x_1^{(3)} = \frac{1}{8}[-3.8878 - 2.9349 - 1] = -0.9778$$

$$x_2^{(3)} = \frac{1}{5}[-0.9778 - 2.9349 - 16] = -3.9825$$

$$x_3^{(3)} = \frac{1}{4}[-0.9778 - 3.9825 - 7] = 2.9901$$

k = 3

$$x_1^{(4)} = \frac{1}{8}[-3.9825 - 2.9901 - 1] = 0.9966$$

$$x_2^{(4)} = \frac{1}{5}[-0.9966 - 2.9901 - 16] = -3.9973$$

$$x_3^{(4)} = \frac{1}{4}[-0.996 - 3.9973 - 7] = -2.9985$$

which is a good approximation to the exact solution x = $(-1\ -4\ -3)^T$ with maximum absolute error 0.0034. Comparing with the results obtained in Example 1, we find that the values of $x_i$, i = 1, 2, 3 obtained here are better approximation to the exact solution than the one obtained in Example 1.

**Example 7**: Gauss-Seidel method is used to solved the system of equations

$$\begin{bmatrix} 4 & -1 & 1 \\ 4 & -8 & 1 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} x \\ x \\ x \end{bmatrix} = \begin{bmatrix} 7 \\ 21 \\ 15 \end{bmatrix} \qquad (29)$$

Determine the rate of convergence of the method and the number of iterations needed to make $\max_i |\hat{I}_i^{(k)}|$ , $10^{-2}$. Perform these number of iterations with $x^{(0)} = [1\ 2\ 2]^T$ and compare the results with the exact solution x = $[2\ 4\ 3]^T$.

**Solution**: The Gauss-Seidel method (26) when applied to the system of Eqns. (29) gives the iteration matrix.

$$H = -\begin{bmatrix} 4 & 0 & 0 \\ 4 & -8 & 0 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Since the inverse of a lower triangular matrix let

$$L = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 4 & -8 & 0 \\ 2 & 1 & 5 \end{bmatrix}$$

then

$$\begin{bmatrix} 4 & 0 & 0 \\ 4 & -8 & 0 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$\backslash \;\; 41_{11} = 1,\ 1_{11} = \frac{1}{4}$

$41_{11}\ \text{-}81_{21} = 0,\ 1_{21}\frac{1}{8}$

$\text{-}81_{22} = 1,\ 1_{22} = \text{-}\frac{1}{8}$

$\text{-}21_{11} + 1_{21} + 51_{31} = 0,\ 1_{31} = \frac{3}{40}$

$\text{-}1_{22} + 51_{32} = 0,\ 1_{32} = \frac{1}{40}$

$51_{33} = 1,\ 1_{33} = \frac{1}{5}$

$$\backslash \;\; L = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ \frac{1}{8} & \frac{-1}{8} & 0 \\ \frac{3}{40} & \frac{1}{40} & \frac{-1}{5} \end{bmatrix}$$

Hence

$$H = \begin{bmatrix} \frac{-1}{4} & 0 & 0 \\ \frac{-1}{8} & \frac{1}{8} & 0 \\ \frac{3}{40} & \frac{-1}{40} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{8} & 0 \\ 0 & \frac{3}{40} & \frac{-1}{10} \end{bmatrix}$$

The eigenvalues of the matrix H are the roots of the characteristic equation

$$\det (H - l\ I) = \begin{vmatrix} -l & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{8} - l & 0 \\ 0 & \frac{3}{40} & -(\frac{1}{10} + l) \end{vmatrix} = 0$$

We have

$$l\ (80l^2 - 2l - 1) = 0$$

which gives

$$l = 0, 0.125, -0.1$$

Therefore, we have

$$r\ (H) = 0.125$$

The rate of convergence of the method is given by

$$n = -\log_{10}(0.125) = 0.9031$$

The number of iterations needed for obtaining the desired accuracy is given by

$$k = \frac{2}{n} = \frac{2}{0.9031} » 3$$

The Gauss-Seidel method when applied to the system of Eqns. (29) becomes

$$\begin{aligned} x_1^{(k+1)} &= \frac{1}{4}\ [7 - x_3^{(k)} + x_2^{(k)}] \\ x_2^{(k+1)} &= \frac{1}{8}\ [-21 - 4x_1^{(k+1)} - x_3^{(k)}] \qquad (30) \\ x_3^{(k+1)} &= \frac{1}{5}\ [15 + 2x_1^{(k+1)} - x_2^{(k+1)}] \end{aligned}$$

The successive iterations are obtained as

$$\begin{aligned} x^{(1)} &= [1.75 \quad 3.75 \quad 2.95]^T \\ x^{(2)} &= [1.95 \quad 3.9688 \quad 2.95]^T \\ x^{(3)} &= [1.9956 \quad 3.9961 \quad 2.9990]^T \end{aligned}$$

which is an approximation to the exact solution after three iterations. Comparing the results obtained in Example 2, we conclude that the Gauss-Seidel method converges faster than the Jacobi method.

**Example 8**: Use the Gauss-Seidel method for solving the following system of equations.

$$\begin{bmatrix} 2 & -1 & 0 & 1 \\ 1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \qquad (31)$$

with $x^{(0)} = [0.5 \;\; 0.5 \;\; 0.5 \;\; 0.5]^T$. Compare the results with those obtained in Example 3 after four iterations. The exact solution is $x = [1 \;\; 1 \;\; 1 \;\; 1]^T$.

**Solution**: Use the Gauss-Seidel method, when applied to the system of Eqns. (31) becomes

$$\begin{aligned} x_1^{(k+1)} &= \frac{1}{2}\,[1 + x_2^{(k)}] \\ x_2^{(k+1)} &= \frac{1}{2}\,[x_1^{(k+1)} + x_3^{(k)}] \qquad (32) \\ x_3^{(k+1)} &= \frac{1}{2}\,[x_2^{(k+1)} + x_4^{(k)}] \\ x_4^{(k+1)} &= \frac{1}{2}[1 + x_3^{(k+1)}],\ k = 0, 1, ... \end{aligned}$$

Starting with the initial approximation $x^{(0)} = [0.5 \;\; 0.5 \;\; 0.5 \;\; 0.5]^T$, we obtain the following iterates

$x^{(1)} = [0.75 \quad 0.625 \quad 0.5625 \quad 0.7813]^T$<br>
$x^{(2)} = [0.8125 \quad 0.6875 \quad 0.7344 \quad 0.8672]^T$<br>
$x^{(3)} = [0.8438 \quad 0.7891 \quad 0.8282 \quad 0.9141]^T$<br>
$x^{(4)} = [0.8946 \quad 0.8614 \quad 0.8878 \quad 0.9439]^T$

In Example 3, the result obtained after four iterations by the Jacobi method was

$x^{(4)} = [0.8438 \quad 0.75 \quad 0.75 \quad 0.8438]^T$

Remark: The matrix formulations of the Jacobi and Gauss-Seidel methods are used whenever we want to check whether the iterations converge or to find the rate of convergence. If we wish to iterate and find solutions of the systems, we shall use the equation form of the methods.

## 4.0 CONCLUSION

We now end this unit by giving a summary of what we have covered in it.

## 5.0 SUMMARY

In this unit, we have covered the following:

a Iterative methods for solving linear system of equations
$Ax = b$ (see Eqn. (1))
where A is an n ´ n, non-singular matrix. Iterative methods are generally used when the system is large and the matrix A is sparse. The process is started using an initial approximation and lead to successively better approximations.

b General iterative method for solving the linear system of Eqn. (1) can be written in the form
$x^{(k+1)} = Hx^{(k)} + c$, k = 0, 1, ............(see Eqn. (5))
where $x^{(k)}$ and $x^{(k+1)}$ are the approximation to the solution vector x at the kth and the (k + 1)th iterations respectively. H is the iteration matrix which depends on A and is generally a constant matrix. c is a column vector and depends on both A and b.

c Iterative method of the form given in 2) above converges for any initial vector, if ||H|| < 1, which is a sufficient condition for convergence. The necessary and sufficient condition for convergence is r (H) <, where r (H) is the spectral radius of H.

d In the Jacobi iteratin method or the method of simultaneous displacements.
$H = -D^{-1}(L + U)$; $c = D^{-1}b$
where D is a diagonal matrix, L and U are respectively the lower and upper triangular matrices with zero diagonal elements.

e In the Gauss-Seidel iteration method or the method of successive displacements
$H = -(D + L)^{-1}U$ and $c = (D + L)^{-1}b$.

f) If the matrix A in Eqn. (1) is strictly diagonally dominant then the Jacobi and Gauss-Seidel methods converge Gauss-Seidel method converges faster than the Jacobi method.

## 6.0 TUTOR-MARKED ASSIGNMENT (TMA)

i Perform five iteration of the Jacobi method for solving the system of equations given in Example 4 with $x^{(0)} = [1 \;\; 1 \;\; 1]^T$.

ii Perform four iterations of the Jacobi method for solving the system of equations

$$\begin{bmatrix} 5 & 2 & 2 \\ 2 & 5 & 3 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ 4 \end{bmatrix}$$

with $x^{(0)} = 0$. Exact solution is $x = (1 \; -1 \; -1)^T$

iii Perform four iterations of the Jacobi method for solving the system of equations

$$\begin{bmatrix} 5 & -1 & -1 & -1 \\ 1 & 10 & -1 & -1 \\ 1 & -1 & 5 & -1 \\ 1 & -1 & -1 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 12 \\ 8 \\ 34 \end{bmatrix}$$

with $x^{(0)} = 0$. The exact solution is $x = [1\ 2\ 3\ 4]^T$

iv Set up the Jacobi method in matrix form for solving the system of equations

$$\begin{bmatrix} 1 & 0 & -1/4 & -1/4 \\ 0 & 1 & -1/4 & -1/4 \\ 1/4 & -1/4 & 1 & 0 \\ 1/4 & -1/4 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

and perform four iterations. Exact solution is $x = (\ 1\ 1\ 1\ 1)^T$. Take $x^{(0)} = 0$.

v Perform four iterations of the Gauss-Seidel method for solving the system of equations given in no. 3.

vi Perform four iterations of the Gauss-Seidel method for solving the system of equations given in no. 4.

vii Gauss-Seidel method is used to solve the system of equations given in no. 4. Determine the rate of convergence and the number of iterations needed to make $\max_{1} |\hat{I}_{i}^{(k)}|$ , $10^{-2}$. Perform four iterations and compare the results with the exact solution.

## 7.0 REFERENCES/FURTHER READINGS

Wrede, R.C. and Spegel M. (2002). Schaum's and Problems of Advanced Calculus. McGraw – Hill N.Y.

Keisler, H.J. (2005). Elementary Calculus. An Infinitesimal Approach. 559 Nathan Abbott, Stanford, California, USA

# UNIT 4 EIGENVALUES AND EIGENVECTORS

## CONTENTS

## 1.0 INTRODUCTION

In Unit 7, you have seen that eigenvalues of the iteration matrix play a major role in the study of convergence of iterative methods for solving linear system of equations. Eigenvalues are also of great importance in many physical problems. The stability of an aircraft is determined by the location of the eigenvalues of a certain matrix in the complex plane. The natural frequencies of the vibrations of a beam are actually eigenvalues of a matrix. Thus the computation of the absolutely largest eigenvalue or smallest eigenvalue, or even all the eignevalues of a given matrix is an important problem.

For a given system of equation of the form

$Ax = l\ x$ (1)
or
$(A - l\ I)x = 0$ (2)
the values of the parameter $l$ , for which the system of Eqn. (2) has a nonzero solution, are called the eigenvalues of A. Corresponding to these eigenvalues, the nonzero solutions of Eqn. (2) i.e. the vectors x, are called the eigenvectors of A. The problem of finding the eigenvalues and the eigenvectors of a square matrix A is known as the eigenvalue problem. In this unit, we shall discuss the eigenvalue problem. To begin with, we shall give you some definitions and properties related to eigenvalues.

## 2.0 OBJECTIVES

At the end of this unit, you should be able to:

- solve simple eigenvalue problems
- obtain the largest eigenvalue in magnitude and the corresponding eigenvector of a given matrix by using the power method

- obtain the smallest eigenvalue in magnitude and an eigenvalue closest to any chosen number along with the corresponding eigenvector of a given matrix by using the inverse power method.

## 3.0 MAIN CONTENT

### 3.1 The Eigenvalue Problem

In the previous three units, we were concerned with the non-homogeneous system of linear equations, Ax = b. We know that this system has a unique solution iff the matrix A is nonsingular. But, if the vector b = 0, then the system reduces to the homogeneous system

$$Ax = 0 \qquad (3)$$

If the coefficient matrix A, in Eqn. (3) is nonsingular, then system has only the zero solution, x = 0. for the homogeneous system (3) to have a nonzero solution is not unique.

The homogeneous system of Eqn. (2) will have a nonzero solution only when the coefficient matrix (A - l I) is singular, that is,

$$\det (A - l\ I) = 0 \qquad (4)$$

If the matrix A is an n ´ n matrix then Eqn. (4) gives a polynomial of degree n in l . This polynomial is called the characteristic equation of A. The n roots $l_1$, $l_2$, ....,$l_n$ of this polynomial are the eigenvalues of A. for each eigenvalue $l_i$, there exists a vector $x_i$ (the eigenvector) which is the nonzero solution of the system of equations

$$(A - l_i)x_i = 0 \qquad (5)$$

The eigenvalues have a number of interesting properties. We shall now state and prove a few of these properties which we shall be using frequently.

P1: A matrix A is singular if and only if it has a zero eigenvalue.

**Proof**: If A has a zero eigenvalue then

det (A – 0 I) = 0
Þ det (A) = 0
Þ A is singular.

Conversely, if A is singular then

det (A) = 0
Þ det (A – 0 I) = 0
Þ 0 is an eigenvalue of the matrix A.

P2: A and $A^T$ have the same eigenvalues.

**Proof**: If l is an eigenvalue of A then

det (A - l I) = 0<br>
Þ det $(A - l\ I)^T = 0$<br>
Þ det $(A^T - l\ I^T) = 0$<br>
Þ det $(A^T - l\ I) = 0$<br>
Þ l is an eigenvalue of $A^T$<br>
Hence the result.

However, the eigenvectors and A and $A^T$ are not the same.

P3: If the eigenvalue of a matrix A are $l_1, l_2, ...., l_n$ then the eigenvalues of $A^m$, m any positive integer, are $l_1^m, l_2^m, ...., l_n^m$. Also both the matrices A and $A^m$ have the same set of eigenvectors.

**Proof**: Since $l_i$ (i = 1, 2, ..., n) are the eigenvalues of A, we have

$$Ax = l_i x, i = 1, 2, ...., n \qquad (6)$$

Pre-multiplying Eqn. (6) by A on both sides, we get

$$A^2x = A\ l_i x = l_i(Ax) = l_i^2 x \qquad (7)$$

which implies that $l_1^2, l_2^2, ...., l_n^2$ are the eigenvalues of $A^2$. further, A and $A^2$ have the same eigenvectors. Pre-multiplying Eqn. (7) (m – 1) times by A on both sides the general result follows.

P4: If $l_1, l_2, ....., l_n$ are the eigenvalues of A, then $1/l_1, 1/l_2, ...., 1/l_n$ are the eigenvalues pf $A^{-1}$. Also both the matrices A and $A^{-1}$ have the same set of eigenvectors.

**Proof**: Since $l_i$ (i = 1, 2, ....., n), are the eigenvalues of A, we have

$$Ax = l_i x, i = 1, 2, ..., n \qquad (8)$$

Pre-multiplying Eqn. (8) on both sides by $A^{-1}$, we get

$A^{-1}A\ x = l_i\ A^{-1}x$

which gives

$x = l_i\ A^{-1}x$

or $A^{-1}x = \dfrac{1}{l_i}x$

and hence the result.

P5: If $l_1$, $l_2$, .....,$l_n$ are the eigenvalues of A, then $l_i - q$, i = 1, 2, ...., n are the eigenvalues of A – qI for any real number q. Both the matrices A and A – qI have the same set of eigenvectors.

**Proof**: Since $l_i$ is an eigenvalues of A, we have

$$Ax = l_i x, \; i = 1, 2, ....., n \qquad (9)$$

Subtracting q x from both sides of Eqn. (9), we get

$$Ax - qx = l_i x - qx$$

which gives

$$(A - qI)x = (l_i - q)x$$

and the results follows.

P6: If $l_i$, i = 1, 2, ....., n are the eigenvalues of A then $\frac{1}{l_i - q}$, i = 1, 2, ...., n are the eigenvalues of $(A - qI)^{-1}$ for any real number q. Both the matrices A and $(A - qI)^{-1}$ have the same set of eigenvectors.

P6 can be proved by combining P4 and P5. we leave the proof to you.

We now give you a direct method of calculating the eigenvalues and eigenvectors of a matrix.

**Example 1**: Find the eigenvalues of the matrix

a) $A =; \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

b) $A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{bmatrix}$

c) $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$

**Solution**:

a) Using Eqns. (4), we obtain the characteristic equations as

$$\det(A - l\,I) = \begin{bmatrix} 1-l & 0 & 0 \\ 0 & 2-l & 0 \\ 0 & 0 & 3-l \end{bmatrix} = 0$$

which gives $(1 - l)(2 - l)(3 - l) = 0$.

and hence the eigenvalues of A are $l_1 = 1$, $l_2 = 2$, $l_3 = 3$.

b) $\det(A - l\ I) = \begin{bmatrix} 1-l & 0 & 0 \\ 2 & 3-l & 0 \\ 4 & 5 & 6-l \end{bmatrix} = 0$

which gives $(1 - l\ )(3 - l\ )(6 - l\ ) = 0$.
and hence the eigenvalues of A are $l_{\ 1} = 1$, $l_{\ 2} = 3$, $l_{\ 3} = 6$.

c) $\det(A - l\ I) = \begin{bmatrix} 1-l & 2 & 3 \\ 0 & 4-l & 5 \\ 0 & 0 & 6-l \end{bmatrix} = 0$

Therefore, $(1 - l\ )(4 - l\ )(6 - l\ ) = 0$.

Eigenvalues of A are $l_{\ 1} = 1$, $l_{\ 2} = 4$, $l_{\ 3} = 6$.

Remark: Observe that in Example 1 (a), the matrix A is diagonal and in parts (b) and (c), it is lower and upper triangular respectively. In these cases the eigenvalues of A are the diagonal elements. This is true for any diagonal, lower triangular or upper triangular matrix. Formally, we give the result in the following theorem.

**Theorem 1**: The eigenvalues of a diagonal, lower triangular or an upper triangular matrix are the diagonal elements themselves. Let us consider another example.

**Example 2**: Find the eigenvalues and the corresponding eigenvectors of the matrices.

a) $\begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$;

b) $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

and

c) $\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$

**Solution**:

a) Using Eqns. (4), we obtain the characteristic equation as

$|A - l\ I| = \begin{vmatrix} 2 - l & 2 \\ 1 & 3 - l \end{vmatrix} = 0,$

which gives the polynomial

$l^{\ 2} - 5l\ + 4 = 0$

i.e., $(l\ - 1)(l\ - 4) = 0$

The matrix A has two distinct real eigenvalues $l_1 = 1$, $l_2 = 4$. To obtain the corresponding eigenvectors we solve the system of Eqn. (5) for each value of l .

For $l = 1$, we obtain the system of equations

$$x_1 + 2x_2 = 0$$
$$x_1 + 2x_2 = 0$$

which redices to a single equation

$$x_1 + 2x_2 = 0$$

Taking $x_2 = k$, we get $x_1 = -2k$, k being arbitrary nonzero constant. Thus, the eigenvector is of the form

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \begin{bmatrix} k \\ k \end{bmatrix} = k \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

For $l = 4$, we obtain the system of equations

$$-2x_1 + 2x_2 = 0$$
$$x_1 - x_2 = 0$$

which reduces to a single equation

$$x_1 - x_2 = 0$$

Taking $x_2 = k$, we get $x_1 = k$ and the corresponding eigenvector is

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Note: In practice we usually omit k and say that $[-2\ 1]^T$ and $[1\ \ 1]^T$ are the eigenvectors of A corresponding to the eigenvalues $l = 1$ and $l = 4$ respectively. Moreover, the eigenvectors in this case are linearly independent.

b) The characteristic equation in this case becomes

$$(l - 1)^2 = 0$$

Therefore, the matrix A has a repeated real eigenvalue. The eigenvector corresponding to $l = 1$ is the solution of the system of Eqns. (5), which reduces to a single equation

$$x_2 = 0$$

Taking $x_1 = k$, we obtain the eigenvector as

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = k \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Note: that, in this case of repeated eigenvalues, we got linearly dependent eigenvectors.

c) The characteristic equation in this case becomes

$$l^2 - 2l + 5 = 0$$

which gives two complex eigenvalues $l - 1 \pm 2i$.

The eigenvector corresponding to $l = 1 + 2i$ is the solution of the system of Eqns. (5). In this case we obtain the following equations

$ix_1 + x_2 = 0$
$x_1 - ix_2 = 0$
which reduces to the single equation
$x_1 - ix_2 = 0$
Taking $x_2$ = k, we get the eigenvector

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Similarly, for l = 1 – 2i, we obtain the eigenvector

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

In the above problem you may note that corresponding to complex eigenvalues, we got complex eigenvectors. Let us now consider an example of 3 ´ 3 matrix.

**Example 3**: Determine the eigenvalues and the corresponding eigenvectors for the matrices

a) $$A = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix};$$

$$A = \begin{bmatrix} 6 & -2 & 2 \\ 2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$$

**Solution**:

a) The characteristic equation in this case becomes

$$\begin{bmatrix} 2-1 & -1 & 0 \\ -1 & 2-1 & -1 \\ 0 & -1 & 2-1 \end{bmatrix} = 0$$

which gives the polynomial
$(2 - 1)(1^2 - 41 + 2) = 0$
Therefore, the eigenvalues of A are 2, $2 + \sqrt{2}$ and $2 - \sqrt{2}$.

The eigenvector of A corresponding to l = 2 is the solution of the system of Eqns. (5), which reduces to
$x_2 = 0$
$x_1 + x_3 = 0$
Taking $x_3$ = k, we obtain the eigenvector

$$\begin{bmatrix} X1 \\ X2 \\ X3 \end{bmatrix} = \mathrm{k} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

The eigenvector of A corresponding to $1 = 2 + \sqrt{2}$ is the solution of the system of equations

$$\begin{bmatrix} \sqrt{2} & -1 & 0 \\ -1 & \sqrt{2} & -1 \\ 0 & -1 & \sqrt{2} \end{bmatrix} \begin{bmatrix} X1 \\ X2 \\ X3 \end{bmatrix} = \mathrm{k} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad (10)$$

To find the solution of system of Eqns. (10), we use Gauss elimination method.

Performing $R_2 - \frac{1}{\sqrt{2}} R_1$, we get

$$\begin{bmatrix} \sqrt{2} & -1 & 0 \\ 0 & -1/\sqrt{2} & -1 \\ 0 & -1 & -\sqrt{2} \end{bmatrix} \begin{bmatrix} X1 \\ X2 \\ X3 \end{bmatrix} = \mathrm{k} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Again performing $R_3 - \sqrt{2} R_2$, we get

$$\begin{bmatrix} \sqrt{2} & -1 & 0 \\ 0 & -1/\sqrt{2} & -1 \\ 0 & 0 & 0 \end{bmatrix}_0 \begin{bmatrix} X1 \\ X2 \\ X3 \end{bmatrix} = \mathrm{k} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Which give the equations
$-\sqrt{2} x_1 – x_2 = 0$
$-x_2 - \sqrt{2}\ x_3 = 0$

Taking $x_3 = k$, we obtain the eigenvector

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \mathrm{k} \begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix}$$

Similarly, corresponding to the eigenvalue $1 = 2 - \sqrt{2}$, the eigenvector is the solution of system of equations

$$\begin{bmatrix} \sqrt{2} & -1 & 0 \\ -1 & \sqrt{2} & -1 \\ 0 & -1 & \sqrt{2} \end{bmatrix} \begin{bmatrix} X1 \\ X2 \\ X3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Using the Gauss elimination method, the system reduces to the equations

$\sqrt{2}x_1 - x_2 = 0$
$x_2 - \sqrt{2}\ x_3 = 0$

Taking $x_3 = k$, we obtain the eigenvector

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = k \begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix}$$

b) The characteristic equation in this case becomes
$(l\ - 8)\ (l\ - 2)^2 = 0$
Therefore the matrix A has the real eigenvalues 8, 2 and 2. The eigenvalue 2 is repeated two times.
The eigenvector corresponding to $l\ = 8$ is solution of system of Eqns. (5), which reduces to
$x_1 + x_2 - x_3 = 0$
$2x_1 + 5x_2 + x_3 = 0$ (11)
$2x_1 - x_2 - 5x_3 = 0$
Subtracting the last equation of system (11) from the second equation we obtain the system of equations
$x_1 + x_2 - x_3 = 0$
$x_2 + x_3 = 0$
Taking $x_3 = k$, the eigenvector is

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = k \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

The eigenvector corresponding to $l\ = 2$ is the solution of system of Eqns. (5), which reduces to a single equation.

$2x_1 - x_2 + x_3 = 0$ (12)

We can take any values for $x_1$ and $x_2$ which need not be related to each other. The two linearly independent solutions can be written as:

$$k\begin{bmatrix}1\\0\\2\end{bmatrix} \text{ or } k\begin{bmatrix}0\\1\\1\end{bmatrix}$$

Note that in Eqn. (12), it is not necessary that we always assign values to $x_1$ and $x_2$. we can assign values to any of the two variables and obtain the corresponding value of the third variable.

On the basis of Example 2 and 3, we can make in general, the following observations:

For a given n ´ n matrix A, the characteristic Eqn. (4) is a polynomial of degree n in l . The n roots of this polynomial $l_1$, ......,$l_n$, called the eigenvalues of A may be real or complex, distinct or repeated. Then,

i) For distinct, real eigenvalues we, obtain linearly independent eigenvectors. (Examples 2(a) and 3(a))
ii) For a repeated eigenvalue, there may or may not be linearly independent eigenvectors. (Examples 2(b) and 3(b))
iii) For a complex eigenvalue, we obtain a complex eigenvector.
iv) An eigenvector is not unique. Any non-zero multiple of it is again an eigenvector.

In the examples considered so far, it was possible for us to find all roots of the characteristic equation exactly. But this may not always be possible. This is particularly true for n > 3. In such cases some iterative method like Newton-Raphson method may have to be used to find a particular eigenvalue or all the eigenvalues from the characteristic equation. However, in many practical problems, we do not require all the eigenvalues but need only a selected eigenvalue. For example, when we use iterative methods for solving a non-homogeneous system of linear equations Ax = b, we need to know only the largest eigenvalue in magnitude of the iteration matrix H, to find out whether the method converges or not. One iterative method, which is frequently used to determine the largest eigenvalue in magnitude (also called the dominant eigenvalue) and the corresponding eigenvector for a given square matrix A is the power method. In this method we do not find the characteristic equation. This method is applicable only when all the eigenvalues are real and distinct. If the magnitude of two or more eigenvalues is the same then the method converges slowly.

## 3.2 The Power Method

Let us consider the eigenvalue problem

$$Ax = l\ x.$$

Let $l_1$, $l_2$, ......,$l_n$ be the n real and distinct eigenvalues of A such that

$$|l_1| > |l_2| > ... > |l_n|$$

Therefore, l $_1$ is the dominant eigenvalue of A.

In this method, we start with an arbitrary nonzero vector $y^{(0)}$ (not an eigenvector), and form a sequence of vectors $(y^{(k)})$

$$y^{(k+1)} = Ay^{(k)},\ k = 0, 1, .... \qquad (13)$$

In the limit as k® ¥ , $y^{(k)}$ converges to the eigenvector corresponding to the dominant eigenvalue of the matrix A. we can stop the iteration when the largest element in magnitude in $y^{(k+1)} - y^{(k)}$ is less than the predefined error tolerance. For simplicity, we usually take the initial vector $y^{(0)}$ with all its elements equal to one.

Note that in the process of multiplying the matrix A with the vector $y^{(k)}$, the elements of the vector $y^{(k+1)}$ may become very large. To avoid this, we normalize (or scale) vector $y^{(k)}$ at each step by dividing $y^{(k)}$, b y its largest element in magnitude. This will make the largest element in magnitude in the vector $y^{(k+1)}$ as one and the remaining elements less than one.

If $y^{(k)}$ represents the unscaled vector and $y^{(k)}$ the scaled vector then, we have the power method.

$$y^{(k+1)} = Av^{(k)} \qquad (14)$$

$$v^{(k+1)} = \frac{1}{m_{m+1}} y^{(k+1)},\ k = 0, 1, ... \qquad (15)$$

with, $v^{(0)} = y^{(0)}$ and $m_{k+1}$ being the largest element in magnitude of $y^{(k+1)}$. We then obtain the dominant eigenvalue by taking the limit

$$l_1 = \lim_{k\text{® ¥}} \frac{(y^{(k+1)})r}{(v^{(k)})r} \qquad (16)$$

where r represents the rth component of that vector. Obviously, there are n ratios of numbers. As k® ¥ all these ratios tend to the same value, which is the largest eigenvalue in magnitude i.e., l $_1$. The iteration is stopped when the magnitude of the difference of any two ratios is less than the prescribed tolerance.

The corresponding eigenvector is then $v^{(k+1)}$ obtained at the end of the last iteration performed.

We now illustrate the method through an example.

**Example 4**: Find the dominant eigenvalue and the corresponding eigenvector correct to two decimal places of the matrix

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

Using the power method.

**Solution**: We take

$y^{(0)} = v^{(0)} = (1 \;\; 1 \;\; 1)^T$

Using Eqn. (14), we obtain

$$y^{(1)} = Av^{(0)} = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Now $m_1 = 1$ and $v^{(1)} = \frac{1}{m_1} y^{(1)} = (1 \;\; 0 \;\; 1)^T$.

Again,

$$y^{(2)} = Av^{(1)} = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

$m_2 = 2$ and $v^{(2)} = \frac{1}{m_2}\; y^{(2)} = (1 \;\; -1 \;\; 1)^T$.

Proceding in this manner, we have

$y^{(3)} = Av^{(2)} = [3 \;\; -4 \;\; 3]^T$

$m_3 = 4$

$v^{(3)} = \frac{1}{4} y^{(3)} = [0.75 \;\; -1 \;\; 0.75]^T$

$y^{(4)} = Av^{(3)} = [2.5 \;\; -3.5 \;\; 2.5]^T$

$m_4 = 3.5$

$v^{(4)} = \frac{1}{3.5} y^{(4)} = [0.7143 \;\; -1 \;\; 0.7143]^T$

$y^{(5)} = Av^{(4)} = [2.4286 \;\; -3.4286 \;\; 2.4286]^T$

$m_5 = 3.4286$

$v^{(5)} = \frac{1}{3.4286} y^{(5)} = [0.7083 \;\; -1 \;\; 0.7083]^T$

$y^{(6)} = Av^{(6)} = [2.4166 \;\; -3.4166 \;\; 2.4166]^T$

$m_6 = 3.4166$

$v^{(6)} = \frac{1}{3.4166} y^{(6)} = [0.7073 \;\; -1 \;\; 0.7073]^T$

$y^{(7)} = Av^{(6)} = [2.4146 \;\; -3.4146 \;\; 2.4146]^T$

$m_7 = 3.4146$

$$v^{(7)} = \frac{1}{3.4146} y^{(7)} = [0.7071 \ \ -1 \ \ 0.7071]^T$$

After 7 iterations, the ratios $\frac{(y^{(7)})r}{(v^{(6)})r}$ are given as 3.4138, 3.4146 and 3.4138. The maximum error in these ratios is 0.0008. Hence the dominant eigenvalue can be taken as 3.414 and the corresponding eigenvector is $[0.7071 \ \ -1 \ \ 0.7071]^T$

Note that the exact dominant eigenvalue of A as obtained in Example 3 was $2 + \sqrt{2} =$ 3.4142 and the corresponding eigenvector was $[1 \ - \sqrt{2} \ \ 1]^T$ which can also be written as $[\frac{1}{\sqrt{2}} \ \ -1 \ \ \frac{1}{\sqrt{2}}]^T = [0.7071 \ \ -1 \ \ 0.7071]^T$

You must have realized that an advantage of the power method is that the eigenvector corresponding to the dominant eigenvalue is also generated at the same time. Usually, for most of the methods of determining eigenvalues, we need to do separate computations to obtain the eigenvector.

In some problems, the most important eigenvalue is the least magnitude. We shall discuss now the inverse power method which gives the least eigenvalue in magnitude.

We first note that if $l$ is the smallest eigenvalue in magnitude of A, then $\frac{1}{l}$ is the largest eigenvalue in magnitude of $A^{-1}$. The corresponding eigenvectors are same. If we apply the power method to $A^{-1}$, we obtain its largest eigenvalue and the corresponding eigenvector.
This eigenvalue is then the smallest eigenvalue in magnitude of A and the eigenvector is same. Since power method is applied to $A^{-1}$, it is called the inverse power method.

Consider the method

$$y^{(k+1)} = A^{-1}v^{(k)}, \ k = 0, 1, 2, \ldots\ldots\ldots \qquad (17)$$

$$v^{(k+1)} = \frac{1}{m_{k+1}} y^{(k+1)} \text{with } v^{(0)} = y^{(0)}$$

where $y^{(0)}$ is an arbitrary nonzero vector different from the eigenvector of A.

However, algorithm (17) is not in suitable form, as one has to find $A^{-1}$. Alternately, we write Eqn. (17) as

$$Ay^{(k+1)} = v^{(k)}$$

$$v^{(k+1)} = \frac{1}{m_{k+1}} y^{(k+1)}, \ k = 0, 1, 2, \ldots\ldots\ldots \qquad (18)$$

We now need to solve a system of equations for $y^{(k+1)}$, which can be obtained using any of the method discussed in the previous units. The largest eigenvalue of $A^{-1}$ is again given by

$$m = \lim_{k\text{®}\,\text{¥}} \frac{(y^{(k+1)})r}{(v^{(k)})r}$$

The corresponding eigenvector is $v^{(k+1)}$.

We now illustrate the method through an example.

**Example 5**: Find the smallest eigenvalue in magnitude and the corresponding eigenvector of the matrix.

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

using four iterations of the inverse power method.

**Solution**: Taking $v^{(0)} = [1 \;\; 1 \;\; 1]^T$, we write

First iteration

$Ay^{(1)} = v^{(0)}$

or

$$\begin{bmatrix} 2 & -1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad (19)$$

For solving the system of Eqns. (19), we use the LU decomposition method. We write

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = LU = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix} \qquad (20)$$

comparing the coefficient on both sides of Eqns. (20), we obtain

$$A = LU = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 3/2 & 0 \\ 0 & -1 & 4/3 \end{bmatrix} \begin{bmatrix} 1 & -1/2 & 0 \\ 0 & 1 & -2/3 \\ 0 & -1 & 4/3 \end{bmatrix}$$

Solving $Lz = v^{(0)}$

and then $Uy^{(1)} = z$

we obtain

$y^{(1)} = \left[3/2 \quad 2 \quad 3/2\right] = [1.5 \quad 2.0 \quad 1.5]^T$
$m_1 = 2.0$
\ $v^{(1)} = \frac{1}{m_1} y^{(1)} = [0.75 \quad 1.0 \quad 0.75]^T$

Second iteration
$Ay^{(2)} = v^{(1)}$
Solving $Lz = v^{(1)}$
and $Uy^{(2)} = z$
we obtain
$y^{(2)} = [1.25 \quad 1.75 \quad 1.25]^T$
$m_2 = 1.75$
$v^{(2)} = \frac{1}{m_2} y^{(2)} = [0.7143 \quad 1 \quad 0.7143]^T$

Thirditeration
$Ay^{(3)} = v^{(2)}$
$y^{(3)} = [1.2143 \quad 1.7143 \quad 1.2143]^T$
$m_3 = 1.7143$
$v^{(3)} = \frac{1}{m_3} y^{(3)} = [0.7083 \quad 1 \quad 0.7083]^T$
Fourthiteration
$Ay^{(4)} = v^{(3)}$
$y^{(4)} = [1.2083 \quad 1.7083 \quad 1.2083]^T$
$m_4 = 1.7083$
$v^{(4)} = \frac{1}{m_4} y^{(4)} = [0.7073 \quad 1 \quad 0.7073]^T$

after 4 iterations, the ratios $\frac{(y^{(4)})r}{(v^{(3)})r}$ are given as 1.7059, 1.7083, 1.7059. The maximum error in these ratios is 0.0024. hence the dominant eigenvalue of $A^{-1}$ can be taken as 1.70. Therefore, $\frac{1}{1.70}$ = 0.5882 is the smallest eigenvalue of A in magnitude and the corresponding eigenvector is given by $[0.7073 \quad 1 \quad 0.7073]^T$.

Note that the smallest eigenvalue in magnitude of A as calculated in Example 3 was 2 - √2 = 0.5858 and the corresponding eigenvector was $[1 \quad \sqrt{2} \quad 1]^T$ or $[0.\ 7071 \quad 1 \quad 0.7071]^T$.

The inverse power method can be further generalized to find some other selected eigenvalues of A. For instance, one may be interested to find the eigenvalue of A which is nearest to some chosen number q. You know from P6 of Sec. 3.1 that the matrices A and A - qI have the same set of eigenvectors. Further, for each eigenvalue $l_i$ of A, $l_i - q$ is the eigenvalue of A – qI.

We can therefore use the iteration

$$y^{(k+1)} = (A - qI)^{-1}v^{(k)} \qquad (21)$$

with scaling as described in Eqns. (14) – (16). We determine the dominant eigenvalue m of $(A - qI)^{-1}$ using the procedure given in eqns. (18), i.e.

$$(A - qI)\, y^{(k+1)} = v^{(k)}$$

$$v^{(k+1)} = \frac{1}{m_{k+1}}\, y^{(k+1)} \qquad (22)$$

Using P6, we have the relation

$m = \dfrac{1}{l - q}$, where $l$ is an eigen value of A.

$$\text{i.e., } l = \frac{1}{m} + q \qquad (23)$$

Now since m is the largest eigenvalue in magnitude of $(A - qI)^{-1}$, $\frac{1}{m}$ must be the smallest eigenvalue in magnitude of A – qI. Hence, the eigenvalue $\frac{1}{m} + q$ of A is closest to q.

**Example 6**: Find the eigenvalue of the matrix A, nearest to 3 and also the corresponding eigenvector using four iterations of the inverse power method where,

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

**Solution**: In this case q = 3. Thus we have

$$A - 3I = \begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & -1 \\ 0 & -1 & -1 \end{bmatrix}$$

To find $y^{(k+1)}$, we need to solve the system

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & -1 \\ 0 & -1 & -1 \end{bmatrix} y^{(k+1)} = v^{(k)} \qquad (24)$$

and normalize $y^{(k+1)}$ as given in Eqn. (22).

First iteration

Starting with $v^{(0)} = [1 \;\; 1 \;\; 1]^T$ and using the Gauss elimination method to solve the system (24), we obtain

$y^{(1)} = [0 \;\; -1 \;\; 0]^T$

$m_1 = 1$

$v^{(1)} = \frac{1}{m_1} y^{(1)} = [0 \;\; -1 \;\; 0]^T$

Second iteration

$Ay^{(2)} = v^{(1)}$

$y^{(2)} = [1 \;\; -1 \;\; 1]^T$

$m_2 = 1$

$v^{(2)} = \frac{1}{m_2} y^{(2)} = [1 \;\; -1 \;\; 1]^T$

Thirditeration

$Ay^{(3)} = v^{(2)}$

$y^{(3)} = [2 \;\; -3 \;\; 2]^T$

$m_3 = 3$

$v^{(3)} = \frac{1}{m_3} y^{(3)} = [\frac{2}{3} \;\; -1 \;\; \frac{2}{3}]^T$

Fourthiteration

$Ay^{(4)} = v^{(3)}$

$y^{(4)} = [\frac{5}{3} \;\; -\frac{7}{3} \; \frac{5}{3}]^T$

$m_4 = \frac{7}{3} = 2.333$

$v^{(4)} = \frac{1}{m_4} y^{(4)} = [\frac{5}{7} \;\; -1 \;\; \frac{5}{7}]^T$

After four iterations, the ratios $\frac{(y^{(4)})r}{(v^{(3)})r}$ are given as 2.5, 2.333, 2.5. The maximum error in these ratios is 0.1667. Hence the dominant eigenvalue of $(A - 31)^{-1}$ can be taken as 2. Thus the eigenvalue 1 of A closest to 3 as given by Eqn. (23) is

$$1 = \frac{1}{m} + 3$$

$$= \frac{1}{2} + 3 = \frac{7}{2} = 3.5$$

and the corresponding eigenvector is $v^{(4)} = \left[5/7 \quad -1 \quad 5/7\right]$ = $[0.7143 \quad -1 \quad 0.7143]^T$. Note that the eigenvalue of A closest to 3 as obtained in Examplee 3 was 2 + √2 = 3.4142. The eigenvector corresponding to this eigenvalue was $[0.7071 \quad -1 \quad 0.7071]^T$

The eigenvalues of a given matrix can also be estimated. That is, for a given matrix A, we can find the region in which all its eigenvalues lie. This can be done as follows:

Let $l_i$ be an eigenvalue of A and $x_i$ be the corresponding eigenvector, i.e.,
$Ax_i = l_i x_i$ (25)

or

$$\begin{array}{l} a_{11}x_{i,1} + a_{12}x_{i,2} + \ldots\ldots + a_{1n}x_{i,n} = l_i x_{i,1} \\ a_{21}x_{i,1} + a_{22}x_{i,2} + \ldots\ldots + a_{2n}x_{i,n} = l_i x_{i,2} \\ . \qquad . \qquad . \qquad . \\ . \qquad . \qquad . \qquad . \\ . \qquad . \qquad . \qquad . \\ a_{k1}x_{i,1} + a_{k2}x_{i,2} + \ldots\ldots + a_{kn}x_{i,n} = l_i x_{i,k} \\ . \qquad . \qquad . \qquad . \\ . \qquad . \qquad . \qquad . \\ . \qquad . \qquad . \qquad . \\ a_{n1}x_{i,1} + a_{n2}x_{i,2} + \ldots\ldots + a_{nn}x_{i,n} = l_i x_{i,n} \end{array} \qquad (26)$$

Let $|x_{i,k}|$ be the largest element in magnitude of the vector $[x_{i,1}, x_{i,2}, \ldots\ldots, x_{i,n}]^T$. Consider the kth equation of the system (26) and divide it by $x_{i,k}$. We then have

$$a_{k1}\left(\frac{x_{i,1}}{x_{i,k}}\right) + a_{k2}\left(\frac{x_{i,2}}{x_{i,k}}\right) + \ldots. + a_{kk} + \ldots. + a_{kn}\left(\frac{x_{i,n}}{x_{i,k}}\right) = l_i \quad (27)$$

Taking the magnitudes on both sides of Eqn. (27), we get

$$|l_i| \;,\; |a_{k1}|\left|\frac{x_{i,1}}{x_{i,k}}\right| + |a_{k2}|\left|\frac{x_{i,2}}{x_{i,k}}\right| + \ldots.. + |a_{kk}| + \ldots. + |a_{kn}|$$

$$, \;\; |a_{k1}| + a_{k2}| + \ldots.. + |a_{kk}| + \ldots. + |a_{kn}| \qquad (28)$$

since $\left|\frac{x_{i,j}}{x_{i,k}}\right|$ , 1 for j = 1, 2, ...... n.

Since eigenvalues of A and $A^T$ are same Ref. P2), Eqn. (28) can also be written as

$$|l_i| \;,\; |a_{1k}| + |a_{2k}| + \ldots.. + |a_{kk}| + \ldots.. + |a_{nk}| \qquad (29)$$

Since $|x_{i,k}|$, the largest element in magnitude, is unknown, we approximate Eqns. (28) and (29) by

$$|1\ |\ ,\ \max_{\mathrm{i}} \sum_{\substack{i=1\\ j=i}}^{n} |a_{ij}| \text{ (maximum absolute row sum)} \qquad (30)$$

and

$$|1\ |\ f\ \max_{\mathrm{j}} \sum_{\substack{i=1\\ j=i}}^{n} |a_{ij}| \text{ (maximum absolute column sum)} \qquad (31)$$

We can also rewrite Eqn. (27) in the form

$$|1\ _{\mathrm{i}} - a_{kk}| = a_{k1}\left(\frac{x_{i,1}}{x_{i,k}}\right) + a_{k2}\left(\frac{x_{i,2}}{x_{i,k}}\right) + .... + a_{kn}\left(\frac{x_{i,n}}{x_{i,k}}\right)$$

and taking magnitude on both sides, we get

$$|1\ _{\mathrm{i}} - a_{kk}|\ ,\ \sum_{\substack{i=1\\ j=i}}^{n} |a_{ij}| \qquad (32)$$

Again, since A and $A^T$ have the same eigenvalues Eqn. (32) can be written as

$$|1\ _{\mathrm{i}} - a_{kk}|\ ,\ \sum_{\substack{i=1\\ j=i}}^{n} |a_{ij}| \qquad (33)$$

Note that since the eigenvalues can be complex, the bounds (30), (31), (32) and (33) represents circles in the complex plane. If the eigenvalues are real, then they represent intervals. For example, when A is symmetric then the eigenvalues of A are real.

Again in Eqn. (32), since k is not known, we replace the circle by the union of the n circle

$$|1\ _{\mathrm{i}} - a_{ii}|\ ,\ \sum_{\substack{i=1\\ j=i}}^{n} |a_{ij}|,\ \mathrm{i} = 1, 2, ........., \mathrm{n}. \qquad (34)$$

Similarly from Eqn. (33), we have that eigenvalues of A lie in the union of circles

$$|1\ _{\mathrm{i}} - a_{ii}| \sum_{\substack{i=1\\ j=i}}^{n} |a_{ij}|\ ,\ \ ,\ \mathrm{i} = 1, 2, ........., \mathrm{n}. \qquad (35)$$

The bounds derived in Eqns. (30), (31), (34) and (35) for eigenvalues are all independent bounds. Hence the eigenvalues must lie in the intersection of these bounds. The circles derived above are called the Gerschgorin circles and the bounds are called the Gerschgorin bounds.

Let us now consider the following examples:

**Example 7**: Estimate the eigenvalues of the matrix

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 3 \\ 1 & 3 & 2 \end{bmatrix}$$

using the Gerschgorin bounds.

**Solution**: The eigenvalues of A lie in following regions:

i) absolute row sums are 4, 6 and 6. Hence
|l | ,   max [4, 6, 6] = 6 (36)

ii) absolute column sums are 4, 5 and 7. Hence
|l | ,   7 (37)

iii) union of the circles [using (35)]
|l  - 1| ,   3
|l  - 1| ,   4
|l  - 2| ,   5
union of circles in (iii) is |l  - 1| ,   5 (38)
union of circles in (iv) is |l  - 2| ,   5 (39)

The eigenvalues lie in all circles (36), (37), (38) and (39) i.e., in the intersection of these circles as shown by shaded region in Fig. 1.

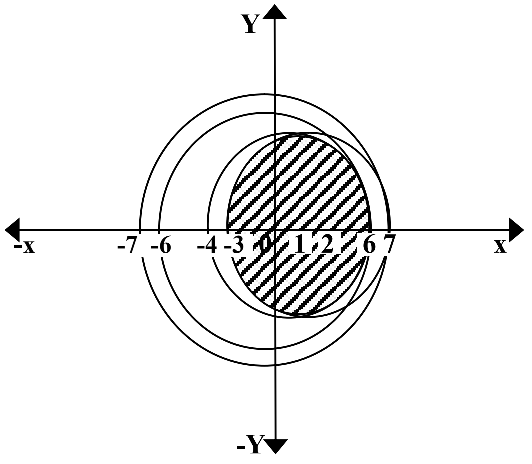

**Fig. 1**

**Example 8**: Estimate the eigenvalues of the symmetric matrix

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & -2 \end{bmatrix}$$

by the Gerschgorin bounds.

**Solution**: The eigenvalues lie in the following regions:

i) | l | , max [4, 4, 6] = 6

ii) union of the circles

a) | l - 1| , 3

b) | l - 1| , 3

c) | l + 1| , 4

Since A is symmetric, it has real eigenvalues. Therefore, the eigenvalues lie in the intervals

i) -6 , l , 6

ii) union of

a) -3 , l -1 , 3, i.e. -2 , l , 4

b) -4 , l +2 , 4, i.e. -6 , l , 2

union of (a) and (c) is -6 , l , 4.

Intersection of (i) and (ii) is -6 , l , 4. Hence the eigenvalues of A lie i the interval -6 , l , 4.

Note that in Example 8, since the matrix A is symmetric, the bounds (30) and (31) are same and also the bounds (34) and (35) are same.

You may now try the following self assessment exercise.

## 4.0 CONCLUSION

We can now conclude as in summary

## 5.0 SUMMARY

In this unit, we have covered the following:

a) For a given system of equations of the form
Ax = l x (see Eqn. (1)).
the values of l for which Eqn. (1) has a nonzero solution are called the eigenvalues and the corresponding nonzero solutions (which are not unique) are called the eigenvectors of the matrix A.

b) The following are the steps involved in solving an eigenvalue problem
  i) Find the nth degree polynomial (called the characteristic equation) inl from det (A - l I) = 0.
  ii) Find the n roots l $_i$, i = 1, 2, ...., n of the characteristic equation.
  iii) Find the eigenvectors corresponding to each l $_i$.

c) For n ƒ 3, it may not be possible to find the roots of the characteristic equation exactly. In such cases, we use some iterative method like Newton Raphson method to find these roots. However,
  i) when only the largest eigenvalue in magnitude is to be obtained, we use the power method. In this method we obtain a sequence of vectors {$y^{(k)}$}, using the iiteative scheme
  $y^{(k+1)} = A\ y^{(k)}$, k = 0, 1, ... (see Eqn. (13))

  which in the limit as k® ¥ , converges to the eigenvector corresponding to the dominant eigenvalue of the matrix A. The vector $y^{(0)}$ is an arbitrary non-zero vector (different from with the eigenvector of A).
  ii) we use the inverse power method with the iteration scheme
  $y^{(k+1)} = (A - qI)^{-1}\ v^{(k)}$,
  i.e., $(A - qI)^{(k+1)} = v^{(k)}$, k = 0, 1, 2, ......
  where $y^{(0)} = v^{(0)}$ is an arbitrary non-zero vector (not an eigenvector)
    a) with q = 0, if only the least eigenvalue of A in magnitude and the corresponding eigenvector are to be obtained and
    b) with any q, if the eigenvalue of A, nearest to some chosen number q and the corresponding eigenvector are to be obtained.

## 6.0 TUTOR-MARKED ASSIGNMENT (TMA)

i Determine the Eigenvalues and the corresponding eigenvectors of the following

$$A = \begin{bmatrix} 1 & \sqrt{2} & 2 \\ \sqrt{2} & 3 & \sqrt{2} \\ 2 & \sqrt{2} & 1 \end{bmatrix}$$

ii $A = \begin{bmatrix} 15 & 4 & 3 \\ 10 & -12 & 6 \\ 20 & -4 & 2 \end{bmatrix}$

iii $A = \begin{bmatrix} 2 & 2 & -3 \\ 2 & 1 & -6 \\ 1 & -2 & 0 \end{bmatrix}$

iv $A = \begin{bmatrix} 2 & -1 & -1 \\ 3 & -2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

v $A = \begin{bmatrix} 1 & \sqrt{2} & 2 \\ \sqrt{2} & 3 & \sqrt{2} \\ 2 & \sqrt{2} & 1 \end{bmatrix}$

vi $A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ 1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$

vii Find the smallest eigenvalue in magnitude and the corresponding eigenvector of the matrix

$A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$

with $v^{(0)}$ – [-1 1]$^T$, using four iterations of the power method.

viii Find the eigenvalue which is nearest to -1 and the corresponding eigenvector for the matrix

$A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$

with $v^{(0)}$ = [-1 1]$^T$, using four iterations of the inverse power method.

ix Using four iterations of the inverse power method, find the eigenvalue which is nearest to 5 and the corresponding eigenvector for the matrix

$A = \begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}$ (exact eigenvalues are = 1 and 6)

with $v^{(0)}$ = [1 1]$^T$

x Estimate the eigenvalues of the matrix A given in Example 3(a) and 3(b), using the Gerschgorin bounds.

## 7.0 REFERENCES/FURTHER READINGS

Wrede, R.C. and Spegel M. (2002). Schaum's and Problems of Advanced Calculus. McGraw – Hill N.Y.

Keisler, H.J. (2005). Elementary Calculus. An Infinitesimal Approach. 559 Nathan Abbott, Stanford, California, USA