CHAPTER 5 MATRICES AND SYSTEMS OF LINEAR EQUATIONS

In this Chapter, you will learn:

- what is a system of linear equations,
- the differences between a vector and a matrix,
- the different types of solution for linear systems,
- how to solve a system of linear equations of the form $A\mathbf{x} = \mathbf{b}$,
- how to solve a system of linear equations of the form $A\mathbf{x} = \lambda\mathbf{x}$.

1. INTRODUCTION TO SYSTEMS OF LINEAR EQUATIONS

A **linear equation** in the variables $x_1, \ldots, x_n$ is an equation that can be written in the form of

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where b and $a_1, \ldots, a_n$ (known as the coefficients) are real or complex numbers, usually being defined in advance. The subscripts n may be any positive integer.

<u>*Example 1*:</u> Is this equation linear?

$4x_1 - 5x_2 + 2 = x_1$

It is linear because it can be arranged algebraically as $3x_1 - 5x_2 = -2$

<u>Self-Test:</u>

$x_2 = 2(\sqrt{6} - x_1) + x_3$

Is this equation linear?

<u>*Example 2*:</u> Is this equation linear?

$4x_1 - 5x_2 = x_1x_2$

This equation is not linear because of the presence of the term x_1x_2.

<u>Self-Test:</u>

$x_2 = 2\sqrt{x_1} - 6$

Is this equation linear? If it is not, state the reason?

A **system of linear equations** (or a linear system) is a collection of one or more linear equations involving the same variables of $x_1, \dots x_n$.

$$a_{11}x_{11} + a_{12}x_{12} + \cdots + a_{1n}x_{1n} = b_1$$

$$a_{21}x_{21} + a_{22}x_{22} + \cdots + a_{2n}x_{2n} = b_2$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$$

$$a_{m1}x_{m1} + a_{m2}x_{m2} + \cdots + a_{mn}x_{mn} = b_m$$

This system of linear equations consists of *n* number of variables and *m* number of linear equations.

Notice the changes in the subscripts of the coefficients and *x*?

2. MATRICES AND VECTORS

The essential information of a system of linear equations can be recorded compactly in a rectangular array called a **matrix**. The plural form of a matrix is known as **matrices**.

Example 3:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

with the coefficients of each variable aligned in columns, the **coefficient matrix** is,

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -8 \\ -4 & 5 & 9 \end{bmatrix}$$

The second linear equation is written as $0x_1 + 2x_2 - 8x_3 = 8$

and the **augmented matrix** of the system is,

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 1 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array}\right]$$

The size of the matrix tells how many rows and columns it has. This matrix has 3 linear equations (given by the number of rows) and 3 variables (given by the number of columns).

A matrix with only one column is called a **column vector**. Similarly a matrix with one row is called a **row vector**. Both are simply known as a vector.

Notice that a vector is written in bold and non-italic. Vectors can also be written as $\vec{u}$ and $\vec{v}$ or $\tilde{u}$ and $\tilde{v}$.

Example 4:

$\mathbf{u} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$ $\mathbf{v} = \begin{bmatrix} 5 & 7 & 9 \end{bmatrix}$

Vector **u** is a row vector in R^2 and vector **v** is a column vector in R^3. The superscript of R indicates the size of the space that the vector is in.

A vector in R^2 has 2 entries and are said to be ordered pairs of real numbers. A vector in R^3 has 3 entries. Thus, a vector in R^n has n entries. Two vectors are said to be equal if their corresponding entries are the same. Thus, a vector in R^3 can never be equal to a vector in R^2.

Self-Test:

Given $\mathbf{u} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 7 \\ 4 \end{bmatrix}$. Are these vectors equal?

The sum of two vectors, **u** + **v** can be obtained by adding the respective corresponding entries of the two vectors, **u** and **v**. A scalar multiple of vector **u** by *c* is when a vector **u** and a real number *c* is multiplied together becoming *c***u**.

c is also known as a scalar.

Example 5:

Given $\mathbf{u} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 2 \\ -4 \end{bmatrix}$, find $4\mathbf{u}$, $(-2)\mathbf{v}$, and $4\mathbf{u} + (-2)\mathbf{v}$.

$4\mathbf{u} = \begin{bmatrix} 4 \\ -8 \end{bmatrix}$, $(-2)\mathbf{v} = \begin{bmatrix} -4 \\ 8 \end{bmatrix}$

and $4\mathbf{u} + (-2)\mathbf{v} = \begin{bmatrix} 4 \\ -8 \end{bmatrix} + \begin{bmatrix} -4 \\ 8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

A vector whose entries are all zero is called a **zero vector** and is usually denoted by **0**.

A system of linear equations can be written as a vector equation involving a linear combination of vectors.

Example 6:

$$x_1 - 2x_2 + x_3 = 0$$
$$x_2 - 8x_3 = 8$$
$$-4x_1 + 5x_2 + 9x_3 = -9$$

This system of linear equations is equivalent to:

$$x_1 \begin{bmatrix} 1 \\ 0 \\ -4 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ -8 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ -9 \end{bmatrix}$$

Side Notes: What is Linear Combinations?

Given vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in R^n and given scalars $c_1, c_2, \ldots, c_n$, the vector **y** defined by

$$\mathbf{y} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

is called a **linear combination** of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ with weights $c_1, c_2, \ldots, c_n$.

3. THE MATRIX EQUATION *A*x = b

Using the fundamental idea of vectors and matrices, a system of linear equations can be rephrase to become a matrix equation in the form of *A***x** = **b**.

Definition:

If *A* is an *m* x *n* matrix, with columns $a_1, \ldots, a_n$, and if **x** is in R^n, then the product of *A* and **x** denoted by $A\mathbf{x} = \mathbf{b}$, is the linear combination of the columns of *A* using the corresponding entries in **x** as weights; that is,

$$A\mathbf{x} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_{n-1} \quad \mathbf{a}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

Example 7:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

is equivalent to these vector equations:

$$x_1 \begin{bmatrix} 1 \\ 0 \\ -4 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ -8 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ -9 \end{bmatrix}$$

and is equivalent to

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -8 \\ -4 & 5 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ -9 \end{bmatrix}$$

in that it becomes a matrix equation of the form $A\mathbf{x} = \mathbf{b}$.

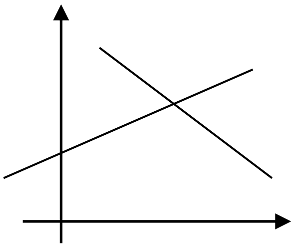
Self-Test:

So by now you should know that a system of linear equations can be written in three compact ways. Try writing all three on the following linear system:

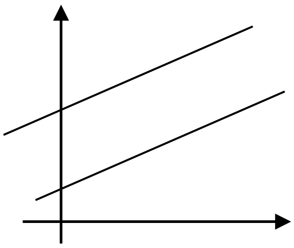
$$\begin{aligned} x_1 + 2x_2 - x_3 &= 4 \\ -5x_2 + 3x_3 &= 1 \end{aligned}$$

4. SOLUTION SETS OF LINEAR SYSTEMS

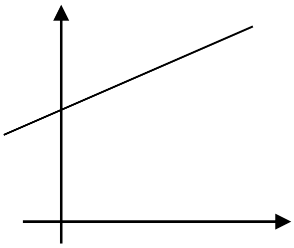
A system of linear equations (or linear system) has either:



One Solution



No Solution



Infinitely Many Solutions

A **solution** of the system of linear equations is a list $(s_1, s_2, \ldots, s_n)$ of numbers that makes each equation a true statement when the values $s_1, s_2, \ldots, s_n$ are substituted into $x_1, \ldots, x_n$, respectively. The set of all possible solutions is called the solution set of the linear system.

Finding the solution set of a system with two linear equations in two variables is equivalent to finding the intersection of two lines. Equivalently, finding the solution set of a system with three linear equations in three variables is similar to finding the intersection between two planes in a 3D space.

Example 8: What is the solution for this system of linear equations?

$$2x_1 - x_2 + \frac{3}{2}x_3 = 8$$

$$x_1 - 4x_3 = -7$$

$s_1 = 5, s_2 = 6.5,$ and $s_3 = 3$ are the solutions because when these values are substituted into x_1, x_2, x_3, respectively where $x_1 = s_1, x_2 = s_2,$ and $x_3 = s_3$, the equations simplify to $8 = 8$ and $-7 = -7$.

Self-Test:

Try to substitute the values into the linear system.

A system of linear equations is said to be **homogeneous** if it can be written in the form of $A\mathbf{x} = \mathbf{0}$, where *A* is an *m* x *n* matrix and **0** is the zero vector in R^m.

Example 9: The following is a homogeneous linear system:

$$\begin{aligned} 3x_1 + 5x_2 - 4x_3 &= 0 \\ -3x_1 - 2x_2 + 4x_3 &= 0 \\ 6x_1 + x_2 - 8x_3 &= 0 \end{aligned}$$

A linear system of $A\mathbf{x} = \mathbf{0}$ *always* has at least one solution, namely $\mathbf{x} = \mathbf{0}$ (the zero vector in R^n). When such system $A\mathbf{x} = \mathbf{0}$ has $\mathbf{x} = \mathbf{0}$ as its only solution, the vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ are said to be **linearly independent**. This solution of $\mathbf{x} = \mathbf{0}$ is known as a trivial solution.

When there exist other nonzero vectors **x** that satisfies $A\mathbf{x} = \mathbf{0}$. Those solutions are known as the nontrivial solution of the homogeneous linear system. The vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ are said to be **linearly dependent**.

Definition:

An indexed set of vectors $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ in R^n is said to be **linearly independent** if the vector equation
$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0}$

has only the trivial solution, where $\mathbf{x} = \mathbf{0}$.

Otherwise, the set of vectors $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ is said to be **linearly dependent** if there exist $x_1, \dots, x_n$, not all zero.

A **nonhomogeneous** linear system is a linear system of $A\mathbf{x} = \mathbf{b}$. This linear system has many solutions. The solution set of $A\mathbf{x} = \mathbf{b}$ is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, where $\mathbf{p}$ is a particular solution and $\mathbf{v}_h$ is any solution of the homogeneous equation $A\mathbf{x} = \mathbf{0}$. This relation is true for all consistent equation of $A\mathbf{x} = \mathbf{b}$. If a linear system has no solution, then it is said to be inconsistent.

Theorem:
If *A* is an *m* x *n* matrix, with columns $a_1, \dots, a_n$, and if **b** is in R^m, the matrix equation

$$A\mathbf{x} = \mathbf{b}$$

has the same solution set as the vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is

$$[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n \quad | \quad \mathbf{b}]$$

5. ELIMINATION PROCESS: ROW REDUCTION, ECHELON FORMS, PIVOTING

The elimination procedure is a systematic procedure for solving linear systems. The basic strategy is to replace one system with an equivalent system that is easier to solve.

Three basic operations are used to simplify a linear system:

1. Replace one equation by the sum of itself and a multiple of another equation.
2. Interchange two equations.
3. Multiply all the terms in an equation by a nonzero constant.

Given a matrix, the basic idea is to perform those three basic operations on the rows of a matrix. Two matrices are said to be **row equivalent** if there is a sequence of elementary row operations that transforms one matrix into the other. If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.

Example 10:

$$\begin{aligned} x+y+2z&=9 \\ 2x+4y-3z&=1 \\ 3x+6y-5z&=0 \end{aligned}$$

Sometimes the variables x_1, x_2, x_3 can be replaced by the variables x, y, z.

This linear system can be rewritten as an augmented matrix,

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{array}\right]$$

and by means of the three basic row operations, it is row equivalent to,

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

This can be written as

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

This symbol ~ has the meaning of being row equivalent.

The simpler augmented matrix that results from the elimination procedure is known as a **reduced echelon matrix**. A reduced echelon matrix is a matrix that is said to be in a **reduced row echelon form**. An *echelon* is a “steplike” pattern that moves down and to the right through the matrix.

Definition:

A rectangular matrix is in **row echelon form** if it has the following three properties:

1. All nonzero rows are above any rows of all zeros.
2. Each leading entry of a row is in a column to the right of the leading entry of the row above it. A **leading entry** of a row refers to the leftmost nonzero entry in a nonzero row.
3. All entries in a column below a leading entry are zeros.

If a matrix in echelon form satisfies the following additional conditions, then it is in **reduced row echelon form**:

4. The leading entry in each nonzero row is 1.
5. Each leading 1 is the only nonzero entry in its column.

Theorem: Each matrix is row equivalent to one and only one reduced row echelon form.

Example 11: The following matrices are in row echelon form. The leading entries (#) may have any nonzero value; the starred entries (*) may have any values including zero.

$$\begin{bmatrix} \# & * & * & * \\ 0 & \# & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & \# & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \# & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \# & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \# & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \# & * \end{bmatrix}.$$

Example 12: The following matrices are in reduced row echelon form because the leading entries are 1's and there are 0's below *and above* each leading 1.

$$\begin{bmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 1 & * & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 1 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Self-Test:

Which of these are in row echelon form and which are in reduced row echelon form?

$$\begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & -2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

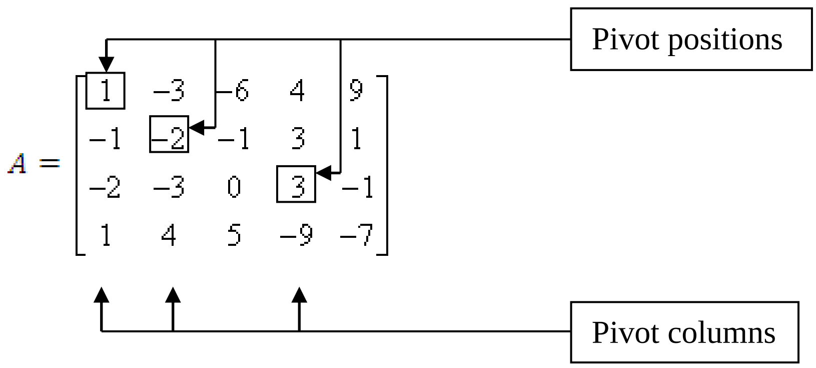
$$\begin{bmatrix} 1 & 4 & 3 & 7 \\ 0 & 1 & 6 & 2 \\ 0 & 0 & 1 & 5 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 & 6 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

When row operations on a matrix produce an echelon form, further row operations to obtain the reduced row echelon form do not change the positions of the leading entries. Since the reduced echelon form is unique, *the leading entries are always in the same positions in any echelon form obtained from a given matrix*. These leading entries correspond to leading 1's in the reduced row echelon form.

Definition:

A **pivot position** in a matrix A is a location in A that corresponds to a leading 1 in the reduced row echelon form of A. A **pivot column** is a column of A that contains a pivot position. A **pivot** is a nonzero number in a pivot position that is used as needed to create zeros via row operations.

Example 13:



6. NAÏVE GAUSS ELIMINATION (GAUSSIAN ELIMINATION WITH BACK- SUBSTITUTION)

Naïve Gauss Elimination is an elimination procedure to change any given matrix into a **row echelon form** and performs **back-substitution** on the resultant linear combination of vector equation to solve for x.

Example 14: Apply Naïve Gauss Elimination to solve the following linear system.

$$\begin{aligned} 2x_2 - 8x_3 &= 8 \\ x_1 - 2x_2 + x_3 &= 0 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

Step 1: Create an augmented matrix for the linear system.

$$\begin{bmatrix} 0 & 2 & -8 & | & 8 \\ 1 & -2 & 1 & | & 0 \\ -4 & 5 & 9 & | & -9 \end{bmatrix}$$

Step 2: Begin with the leftmost nonzero column. This is a pivot column 1. The pivot position is at the top (row 1).

$$\begin{bmatrix} 0 & 2 & -8 & | & 8 \\ 1 & -2 & 1 & | & 0 \\ -4 & 5 & 9 & | & -9 \end{bmatrix}$$

Pivot column

Step 3: Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry into the pivot position.

> At the beginning the pivot is at row 2 of the pivot column. After that row 1 and row 2 are interchanged so that the pivot moves to the first topmost of the pivot column.

Pivot Pivot

$$\begin{bmatrix} 0 & 2 & -8 & | & 8 \\ 1 & -2 & 1 & | & 0 \\ -4 & 5 & 9 & | & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & | & 0 \\ 0 & 2 & -8 & | & 8 \\ -4 & 5 & 9 & | & -9 \end{bmatrix}$$

Step 4: Use row replacement operations to create zeros in all positions below the pivot.

Here, you want to keep x_1 in the row 1 and eliminate it from row 3.

Pivot

$$\begin{bmatrix} 1 & -2 & 1 & | & 0 \\ 0 & 2 & -8 & | & 8 \\ -4 & 5 & 9 & | & -9 \end{bmatrix}$$

To do so, this is what you do:	It is equivalent to:
$4\times[\text{row } 1]$	$4x_1 - 8x_2 + 4x_3 = 0$
$+[\text{row } 3]$	$-4x_1 + 5x_2 + 9x_3 = -9$
[new row 3]	$0x_1 - 3x_2 + 13x_3 = -9$

The result of this calculation is written in place of the original row 3:

$$\begin{bmatrix} 1 & -2 & 1 & | & 0 \\ 0 & 2 & -8 & | & 8 \\ 0 & -3 & 13 & | & -9 \end{bmatrix}$$

Step 5: Apply steps 2-4 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

Here, you want to keep x_2 and eliminate it from the row below it (row 3).

Pivot

$$\begin{bmatrix} 1 & -2 & 1 & | & 0 \\ 0 & 2 & -8 & | & 8 \\ 0 & -3 & 13 & | & -9 \end{bmatrix}$$

Pivot column

To do so, this is what you do:	It is equivalent to:
	$\left(\frac{3}{2}\right)2x_2 - \left(\frac{3}{2}\right)8x_3 = \left(\frac{3}{2}\right)8$
$\frac{3}{2}\times$[row 2]	$-3x_2 + 13x_3 = -9$
+[row 3]	$3x_2 - 12x_3 = 12$
[new row 3]	$-3x_2 + 13x_3 = -9$
	$0x_2 + x_3 = 3$

The result of this calculation is written in place of the original row 3:

$$\begin{bmatrix} 1 & -2 & 1 & | & 0 \\ 0 & 2 & -8 & | & 8 \\ 0 & 0 & 1 & | & 3 \end{bmatrix}$$

The augmented matrix is now in a row echelon form. The **Naïve Gauss Elimination** steps end here. Proceed to Step 6.

Step 6: Apply linear combination to the vector equation. Notice that the new system has a *triangular* form.

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ x_3 &= 3 \end{aligned}$$

Step 7: Apply back-substitution for solution.
Since $x_3 = 3$, substitute x_3 into equation of row 2 to get x_2.

Once you have got $x_2 = 16$, substitute both x_2 and x_3 into equation of row 1 to get x_1.

$$\begin{aligned} &x_3 = 3 \\ &2x_2 - 8(3) = 8 \Rightarrow x_2 = 16 \\ &x_1 - 2(16) + (3) = 0 \Rightarrow x_1 = 29 \end{aligned}$$

This system has *one unique solution*, since there is only one values for each *x*.

7. GAUSS-JORDAN ELIMINATION

Gauss-Jordan Elimination is an elimination procedure to change any given matrix into a **reduced row echelon form.** The resultant linear combination of vector equation directly gives out the solution of x.

Example 15: Apply Gauss-Jordan Elimination to solve the following linear system.

$$\begin{aligned} 3x_2 - 6x_3 + 6x_4 + 4x_5 &= -5 \\ 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 &= 9 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 &= 15 \end{aligned}$$

Step 1: Create an augmented matrix for the linear system.

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & | & -5 \\ 3 & -7 & 8 & -5 & 8 & | & 9 \\ 3 & -9 & 12 & -9 & 6 & | & 15 \end{bmatrix}$$

Step 2: Begin with the leftmost nonzero column. This is a pivot column. The pivot position is at the top.

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & | & -5 \\ 3 & -7 & 8 & -5 & 8 & | & 9 \\ 3 & -9 & 12 & -9 & 6 & | & 15 \end{bmatrix}$$

Pivot column

Step 3: Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry into the pivot position.

Pivot Pivot

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & | & -5 \\ 3 & -7 & 8 & -5 & 8 & | & 9 \\ 3 & -9 & 12 & -9 & 6 & | & 15 \end{bmatrix} \sim \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & | & 15 \\ 3 & -7 & 8 & -5 & 8 & | & 9 \\ 0 & 3 & -6 & -6 & 4 & | & -5 \end{bmatrix}$$

Here rows 1 and 3 have been interchanged. Or you could interchange rows 1 and 2.

Step 4: Use row replacement operations to create zeros in all positions below the pivot.

Pivot

$$\left[\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & -6 & 4 & -5 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & -6 & 4 & -5 \end{array}\right]$$

$-1\times$[row 1]
$+$[row 2]
[new row 2]

Step 5: Apply steps 2-4 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

Pivot

$$\left[\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & -6 & 4 & -5 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right]$$

$-\frac{3}{2}\times$[row 2]
$+$[row 3]
[new row 3]

Step 6: Beginning with the rightmost pivot and working upward and to the left, create zeros **above** each pivot. If a pivot is not 1, make it 1 by a division (scaling) operation.

Pivot

$$\left[\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right]$$

$-6\times$[row 3]
$+$[row 1]
[new row 1]

$-2\times$[row 3]
$+$[row 2]
[new row 2]

Step 7: Repeat Step 6 for the next rightmost pivot until the matrix reaches a reduced row echelon form.

$$\left[\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right] \quad \frac{\frac{1}{2}\times[\text{row 2}]}{[\text{new row 2}]}$$

Pivot Pivot

$$\left[\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right] \quad \frac{9\times[\text{row 2}] + [\text{row 1}]}{[\text{new row 1}]}$$

Pivot Pivot

$$\left[\begin{array}{ccccc|c} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right] \quad \frac{\frac{1}{3}\times[\text{row 1}]}{[\text{new row 1}]}$$

The augmented matrix is now in a reduced row echelon form. The **Gauss-Jordan Elimination** steps end here. Proceed to Step 8.

Step 9: Apply linear combination to the vector equation.

$$\begin{aligned} x_1 - 2x_3 + 3x_4 &= -24 \\ x_2 - 2x_3 + 2x_4 &= -7 \\ x_3 + x_5 &= 4 \end{aligned}$$

Step 10: Obtain the solution.

$$\begin{aligned} x_1 &= -24 + 2x_3 - 3x_4 \\ x_2 &= -7 + 2x_3 - 2x_4 \\ x_5 &= 4 - x_3 \end{aligned}$$

x_3 and x_4 are called free variables.

Free variables indicate that the solution is *not unique*. Each different choice of x_3 and x_4 determines a different solution of the system. Thus, the system has infinitely many solutions.

Do you know?

This linear system is known as an underdetermined system, where there are more variables than there are equations. An overdetermined system is the vice versa, where there are more equations than variables.

Do you know?

The elimination steps to get a row echelon form (as in Naïve Gauss Elimination) are known as the forward phase. The proceeding steps from there to get a reduced row echelon form (as in Gauss-Jordan Elimination) are known as the backward phase. In general, the forward phase of row reduction takes much longer than the backward phase.

8. SOLVING A SYSTEM OF LINEAR EQUATIONS BY MATRIX INVERSION

Theorem:

If *A* is an invertible *n* x *n* matrix, then for each **b** in R^n, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

By this theorem, a system of linear equations can be solved via the **inverse** of the coefficient matrix *A*. An *n* x *n* matrix is said to be **invertible** if there is an *n* x *n* matrix *C* such that

$$CA = I \quad \text{and} \quad AC = I$$

where $I = I_n$, the *n* x *n* identity matrix. In this case, *C* is an inverse of *A*. This unique inverse is denoted by A^{-1}, so that

$$A^{-1}A = I \quad \text{and} \quad AA^{-1} = I$$

<u>Theorem:</u>

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

If $ad - bc = 0$, then A is not invertible.

The quantity $ad - bc$ is called the **determinant** of A. It is denoted as $\det A$.

<u>Example 16</u>: Use the matrix inversion technique to solve the following linear system.

$$3x_1 + 4x_2 = 3$$
$$5x_1 + 6x_2 = 7$$

Step 1: Construct the coefficient matrix A.

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

Step 2: Find the inverse of A.

Here, $a = 3, b = 4, c = 5, d = 6$, so

$$A^{-1} = \frac{1}{3(6) - 4(5)}\begin{bmatrix} 6 & -4 \\ -5 & 3 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} -3 & 2 \\ \frac{5}{2} & -\frac{3}{2} \end{bmatrix}$$

Step 3: Multiply A^{-1} to **b** to get **x**.

Given that $\mathbf{b} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$,

$$\mathbf{x} = \begin{bmatrix} -3 & 2 \\ \frac{5}{2} & \frac{3}{2} \end{bmatrix}\begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} -3(3) + 2(7) \\ \frac{5}{2}(3) + \frac{3}{2}(7) \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix} \begin{matrix} \Rightarrow x_1 = 5 \\ \Rightarrow x_2 = -3 \end{matrix}$$

Solving a linear system by matrix inversion is seldom used since numerically the elimination method is nearly always faster. An exception is the 2 x 2 case as in this example. For a general $n \times n$ matrix, the following theorem states that A^{-1} can be found by the row reduction technique.

> Theorem:
> An $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n, and in this case, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}.

If A and I are placed side-by-side to form an augmented matrix $[A \quad | \quad I]$, then row operations on this matrix produce identical operations on A and on I. If A is row equivalent to I, then $[A \quad | \quad I]$ is row equivalent to $[I \quad | \quad A^{-1}]$. Otherwise, A has no inverse.

Example 17: Find the inverse of the matrix below, if it exists:

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

Step 1: Construct augmented matrix $[A \quad | \quad I]$.

$$\begin{bmatrix} 0 & 1 & 2 & | & 1 & 0 & 0 \\ 1 & 0 & 3 & | & 0 & 1 & 0 \\ 4 & -3 & 8 & | & 0 & 0 & 1 \end{bmatrix}$$

Step 2: Apply row reduction technique till the matrix A is in a reduced row echelon form.

$$\left[\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array}\right]$$

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

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{3}{2} & -2 & \frac{1}{2} \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{9}{2} & 7 & -\frac{3}{2} \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & \frac{3}{2} & -2 & \frac{1}{2} \end{array}\right]$$

$$A^{-1} = \begin{bmatrix} -\frac{9}{2} & 7 & -\frac{3}{2} \\ -2 & 4 & -1 \\ \frac{3}{2} & -2 & \frac{1}{2} \end{bmatrix}$$

Do you know?

In practice, A^{-1} is seldom computed, unless the entries of A^{-1} are needed. Computing both A^{-1} and $A^{-1}\mathbf{b}$ takes about three times as many arithmetic operations as solving $A\mathbf{x} = \mathbf{b}$ by row reduction.

Sometimes a matrix with no inverse is known as a singular matrix. You might occasionally encounter a "nearly singular" or **ill-conditioned matrix**. This type of matrix is an invertible matrix but it can become singular if some of its entries are changed over so slightly. In this case, row reduction may produce fewer than *n* pivot positions, as a result of roundoff error. Also, roundoff error can sometimes make a singular matrix appear to be invertible.

For a $n \times n$ (square) matrix, a **condition number** can be computed. The condition number of a $n \times n$ matrix *A* is

$$cond\ A = \|A\|\,\|A^{-1}\|$$

where $\| \quad \|$ indicates the matrix norm on *A* and A^{-1}.

When solving linear system of $A\mathbf{x} = \mathbf{b}$, the condition number indicates the accuracy of **x**. The larger the condition number, the closer the matrix is to being singular. The condition number of the identity matrix is 1. A singular matrix has an infinite condition number.

9. LU FACTORIZATION

A *factorization* of a matrix A is an equation that expresses A as a product of two or more matrices. It aims to solve a sequence of equations, all with the same coefficient matrix. The LU factorization is described below:

$$A\mathbf{x} = \mathbf{b}_1, \quad A\mathbf{x} = \mathbf{b}_2, \quad \dots, \quad A\mathbf{x} = \mathbf{b}_p$$

When A is invertible, one could compute A^{-1} and then compute $A^{-1}\mathbf{b}_1, A^{-1}\mathbf{b}_2$, and so on. However, it is more efficient to solve $A^{-1}\mathbf{b}_1$ by row reduction and obtained an LU factorization of A at the same time. Thereafter, the remaining equations are solved with the LU factorization.

At first, assume that A is an m x n matrix that can be row reduced to echelon form, *without row interchanges*. Then A can be written in the form of $A = LU$, where L is an m x m lower triangular matrix with 1's on the diagonal and U is an m x n echelon form of A. Such factorization is called an **LU factorization** of A. The matrix L is invertible and is called a *unit* triangular matrix.

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} \# & * & * & * & * \\ 0 & \# & * & * & * \\ 0 & 0 & 0 & \# & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{U}$$

The leading entries (#) may have any nonzero value; the starred entries (*) may have any values including zero.

Such factorization is useful since whenever $A = LU$, the equation $A\mathbf{x} = \mathbf{b}$ can be written as $L(U\mathbf{x}) = \mathbf{b}$. Writing $\mathbf{y}$ for $U\mathbf{x}$, $\mathbf{x}$ can be found by solving the *pair* of equations,

$$L\mathbf{y} = \mathbf{b}$$

$$U\mathbf{x} = \mathbf{y}$$

First solve for $L\mathbf{y} = \mathbf{b}$ for $\mathbf{y}$, and then solve $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$. Each equation is easy to solve because L and U are triangular.

Do you know?

Whereas matrix multiplication involves a *synthesis* of data (combining the effect of two or more matrices into a single matrix), matrix factorization is an analysis of data. In the language of computer science, the expression of *A* as a product amounts to *preprocessing* of the data in *A*, organizing that data into two or more parts whose structures are more useful in some way, perhaps more accessible to computation.

Example 18: Find an LU factorization of,

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix}$$

Step 1: Construct matrix *L*.
Since *A* has four rows, *L* should be 4 x 4. The first column of *L* is the first column of *A* divided by the top pivot entry.

Pivot

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

$\frac{1}{2} \times$[pivot entry 1]

[new column 1]

Compare the first columns of *A* and *L*. *The row operations that create zeros in the first column of A will also create zeros in the first column of L*. You want this same correspondence of row operations to hold for the rest of *L*.

Step 2: Perform a row reduction of A to an echelon form U.

$$A=\begin{bmatrix}2&4&-1&5&-2\\-4&-5&3&-8&1\\2&-5&-4&1&8\\-6&0&7&-3&1\end{bmatrix}\sim\begin{bmatrix}2&4&-1&5&-2\\0&3&1&2&-3\\0&-9&-3&-4&10\\0&12&4&12&-5\end{bmatrix}$$

$$\sim\begin{bmatrix}2&4&-1&5&-2\\0&3&1&2&-3\\0&0&0&2&1\\0&0&0&4&7\end{bmatrix}\sim\begin{bmatrix}2&4&-1&5&-2\\0&3&1&2&-3\\0&0&0&2&1\\0&0&0&0&5\end{bmatrix}$$

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

Step 3: Complete matrix L. At each pivot column, divide the entries by the pivot of that column and place the result into L.

From the first pivot column $\begin{bmatrix}2\\-4\\2\\-6\end{bmatrix}\div 2$

From the second pivot column $\begin{bmatrix}3\\-9\\12\end{bmatrix}\div 3$

From the third pivot column $\begin{bmatrix}2\\4\end{bmatrix}\div 2$

From the fourth pivot column $[5]\div 5$

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

<u>Self-Test:</u>

Verify the results of L and U by performing the matrix multiplication of LU and check that it should be equal to matrix A.

For solving a linear system, once the matrices L and U have been found, proceed to solving the pair of equations,

$$L\mathbf{y} = \mathbf{b} \qquad U\mathbf{x} = \mathbf{y}$$

Example 19: Solve the linear system, given that the LU factorization matrix is as below:

$$\begin{aligned} 3x_1 - 7x_2 - 2x_3 + 2x_4 &= -9 \\ -3x_1 + 5x_2 + x_3 &= 5 \\ 6x_1 - 4x_2 - 5x_4 &= 7 \\ -9x_1 + 5x_2 - 5x_3 + 12x_4 &= 11 \end{aligned}$$

$$LU = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Step 1: Solve $L\mathbf{y} = \mathbf{b}$ by using forward substitution to get **y.**

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix} \sim \Rightarrow y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} -9 \\ -4 \\ 5 \\ 1 \end{bmatrix}$$

Step 2: Solve $U\mathbf{x} = \mathbf{y}$ by using back substitution row to get **x**.

$$\begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -9 \\ -4 \\ 5 \\ 1 \end{bmatrix} \sim \Rightarrow \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ -6 \\ -1 \end{bmatrix}$$

$\therefore x_1 = 3, x_2 = 4, x_3 = -6, x_4 = -1$

10. SOLVING LINEAR SYSTEM BY ITERATIVE METHODS

Gaussian Elimination is a finite sequence of $O(n^3)$ floating point operations that result in a solution. For that reason, Gaussian Elimination is called a **direct method** for solving systems of linear equations. Direct methods, in theory, give the exact solution within a finite number of steps. So-called **iterative methods** also can be applied to solving systems of linear equations. The methods begin with an initial guess and refine the guess at each step, converging to the solution vector.

Direct methods based on Gaussian Elimination provide the user a finite number of steps that terminate in the solution. What is the reason for pursuing iterative methods, which are only approximate and may require several steps for convergence?

There are two major reasons for using iterative methods. Both reasons stem from the fact that one step of an iterative method requires only a fraction of the floating point operations of a full LU factorization. A single step of Jacobi's Method, for example, requires about n^2 multiplications and about the same number of additions. The question is how many steps will be needed for convergence within the user's tolerance.

One particular circumstance that argues for an iterative technique is when a good approximation to the solution is already known. For example, suppose that a solution to $A\mathbf{x} = \mathbf{b}$ is known, after which A and/or **b** change by a small amount. If the solution to the previous problem is used as a starting guess for the new, but similar, problem, fast convergence can be expected.

Such technique is known as **polishing**, because the method begins with an approximate solution, which could be the solution from a previous, related problem, and the merely refines the approximate solution to make it more accurate. Polishing is common in real-time applications where the same problem needs to be re-solved repeatedly with data that is updated as time passes. If the system is large and time is short, it may be impossible to run an entire Gaussian Elimination or even a back-substitution in the allotted time. If the solution has not changed too much, a few steps of a relatively cheap iterative method might keep sufficient accuracy as the solution moves through time.

The second major reason to use iterative methods is to solve sparse systems of equations. A coefficient matrix is called **sparse** if many of the matrix entries are known to be zero. A **full** matrix is the opposite, where few entries may be assumed to be zero.

The iterative-method produces a sequence of approximate solution vectors $\boldsymbol{x}^{(1)}$, $\boldsymbol{x}^{(2)}$….for the linear system $\boldsymbol{A}\mathbf{x}=\boldsymbol{b}$. The iterative method starts with selecting the nonsingular matrix $\boldsymbol{Q}$ and having a starting vector $\boldsymbol{x}^{(0)}$, then generate vectors $\boldsymbol{x}^{(1)}$, $\boldsymbol{x}^{(2)}$….recursively from the equation

$$\boldsymbol{Q}\,\boldsymbol{x}^{(n+1)}=(\boldsymbol{Q}-\boldsymbol{A})\,\boldsymbol{x}^{(n)}+\boldsymbol{b} \quad \text{where } n=0,1,2,3,4\ldots\ldots$$

11. JACOBI METHOD

The Jacobi Method is a form of fixed-point iteration for a system of linear equations ***A*x**=***b***. In Jacobi iteration, ***Q*** is taken to be the diagonal of ***A***. We write the equation of Jacobi method as

$$\boldsymbol{x}^{(n+1)}=\boldsymbol{B}\boldsymbol{x}^{(n)}+\boldsymbol{h}$$

where $\boldsymbol{B} = \boldsymbol{I} - \boldsymbol{Q}^{-1}\boldsymbol{A}$, $\boldsymbol{h}= \boldsymbol{Q}^{-1}\boldsymbol{b}$ and n=0,1,2,3,4……

For example, $\boldsymbol{A}=\begin{bmatrix}3 & 1\\ 1 & 2\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}5\\ 5\end{bmatrix}$

$$\boldsymbol{Q}=\begin{bmatrix}3 & 0\\ 0 & 2\end{bmatrix},\ \boldsymbol{Q}^{-1}\boldsymbol{A}=\begin{bmatrix}\frac{1}{3} & 0\\ 0 & \frac{1}{2}\end{bmatrix}\begin{bmatrix}3 & 1\\ 1 & 2\end{bmatrix}=\begin{bmatrix}1 & \frac{1}{3}\\ \frac{1}{2} & 1\end{bmatrix}$$

$$\boldsymbol{B}=\boldsymbol{I}-\boldsymbol{Q}^{-1}\boldsymbol{A}=\begin{bmatrix}0 & -\frac{1}{3}\\ -\frac{1}{2} & 0\end{bmatrix},\ \boldsymbol{h}=\boldsymbol{Q}^{-1}\boldsymbol{b}=\begin{bmatrix}\frac{5}{3}\\ \frac{5}{2}\end{bmatrix}$$

In order to approximate the solution $\boldsymbol{x}^{(1)}$, $\boldsymbol{x}^{(2)}$….for the linear system ***A*x**=***b***, we iterate the equation of the Jacobi method, starting with an initial guess(starting vector) $\boldsymbol{x}^{(0)}$. The iterations are stopped when the absolute relative approximate error is less than a pre-specified tolerance. The formula for absolute relative approximate error is given as below:

$$|\in_{n+1}|_x=\left|\frac{x^{n+1}-x^{n}}{x^{n+1}}\right|\times 100 \qquad where \qquad n\geq 0$$

<u>*Example 22*</u>: Apply the Jacobi Method to the system,

$$u+2v=5$$
$$3u+v=5$$

with initial guess $[u_0, v_0]=[0,0]$.

Remember that the variables x_1, x_2 can be replaced by the variables u, v.

Step 1: Begin by constructing the first equation for *u* and the second equation for *v*.

$$u_{k+1}=5-2v_k$$
$$v_{k+1}=5-3u_k$$

Step 3: Use the initial guess $[u_0, v_0] = [0,0]$ and iterate.

First iteration:

$$u_1 = 5 - 2(0) = 5$$

$$v_1 = 5 - 3(0) = 5$$

After the first iteration, the absolute relative approximate errors are

$$|\in_1|_u = \left|\frac{5-0}{5}\right| \times 100 = 100.00\%$$

$$|\in_1|_v = \left|\frac{5-0}{5}\right| \times 100 = 100.00\%$$

Repeating more iterations, the following values are obtained

Iteration	u	$\|\in_{n+1}\|_u$ (%)	v	$\|\in_{n+1}\|_v$ (%)
0	0		0	
1	5	100	5	100
2	-5	200	-10	150
3	25	120	20	150

As seen in the table above, the estimated solutions are not converging to the exact solution of which is $u = 1$ and $v = 2$. Why? This is because the coefficient matrix of the system of equations is not diagonally dominant. In other words, if a system of equations has coefficient matrix that is not diagonally dominant, it may not converge.

Convergence Theorems for Iterative Methods

A simple way to determine the convergence is to inspect the diagonal elements.
All of the diagonal elements must be non-zero. Convergence is guaranteed if the system is diagonally dominant.

Definition:

The $n \times n$ matrix $A = (a_{ij})$ is **strictly diagonally dominant** if, for each $1 \leq i \leq n$, $|a_{ij}| > \sum_{j \neq i} |a_{ij}|$. In other words, each main diagonal entry dominates its row in the sense that it is greater in magnitude than the sum of magnitudes of the remainder of the entries in its row.

Theorem 1:

If the $n \times n$ matrix A is strictly diagonally dominant, then

1. A is a non-singular matrix, and
2. for every vector $\mathbf{b}$ and every starting guess, the Iterative Methods applied to $A\mathbf{x} = \mathbf{b}$ converges to the (unique) solution.

Theorem 2: Spectral Radius Theorem

In order that the equations generated by $\boldsymbol{Q}\,\boldsymbol{x}^{(n)} = (\boldsymbol{Q}-\boldsymbol{A})\,\boldsymbol{x}^{(n-1)} + \boldsymbol{b}$ to converge, no matter what initial guess $\boldsymbol{x}^{(0)}$ is selected, it is necessary that all eigenvalues of $\boldsymbol{B} = (\boldsymbol{I} - \boldsymbol{Q}^{-1}\boldsymbol{A})$ satisfy the condition $\rho(\boldsymbol{I} - \boldsymbol{Q}^{-1}\boldsymbol{A}) < 1$. Where ρ is the spectral radius of $\boldsymbol{B}$ and ρ=maximum of |eigenvalues $\boldsymbol{\lambda}_i$ of $\boldsymbol{B}$|.

Example 20: Matrix A is strictly diagonally dominant because $|3| > |1|$ in row 1 and $|2| > |1|$ in row 2; but matrix C is not.

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

Convergence is guaranteed for matrix A when Jacobi Method is applied.

Note that strict diagonally dominance is only a sufficient condition. The Jacobi Method may still converge in its absence.

Example 21: Determine whether the matrices are strictly diagonally dominant.

$$A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & -5 & 2 \\ 1 & 6 & 8 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & 2 & 6 \\ 1 & 8 & 1 \\ 9 & 2 & -2 \end{bmatrix}$$

Matrix A is diagonally dominant because $|3| > |1| + |-1|$ in row 1, $|-5| > |2| + |2|$ in row 2, and $|8| > |1| + |6|$ in row 3.

Matrix B is not, because, for example, $|3| > |2| + |6|$ is not true. However, matrix B can be made diagonally dominant if the first and third rows are exchanged to become,

$$B = \begin{bmatrix} 9 & 2 & -2 \\ 1 & 8 & 1 \\ 3 & 2 & 6 \end{bmatrix}$$

Now, row 1 of B is $|9| > |2| + |-2|$, row 2 is $|8| > |1| + |1|$, and row 3 is $|6| > |3| + |2|$.

Example 22: Apply the Jacobi Method to the system,

$$3u + v = 5$$
$$u + 2v = 5$$

with initial guess $[u_0, v_0] = [0,0]$

Step 1: Construct the coefficient matrix and check if it is diagonally dominant. If necessary, exchange rows to fulfil this condition.

$$\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$$

Step 2: Begin by solving the first equation for u and the second equation for v.

$$u = \frac{5 - v}{3} \quad \Rightarrow \quad u_{k+1} = \frac{5 - v_k}{3}$$

$$v = \frac{5 - u}{2} \quad \Rightarrow \quad v_{k+1} = \frac{5 - u_k}{2}$$

Step 3: Use the initial guess $[u_0, v_0] = [0,0]$ and iterate.

First iteration:

$$u_1 = \frac{5 - (0)}{3} = 1.6667$$

$$v_1 = \frac{5 - (0)}{2} = 2.5$$

After the first iteration, the absolute relative approximate errors are

$$|\epsilon_1|_u = \left| \frac{1.6667 - 0}{1.6667} \right| \times 100 = 100.00\%$$

$$|\epsilon_1|_v = \left| \frac{2.5 - 0}{2.5} \right| \times 100 = 100.00\%$$

The maximum absolute relative approximate error after the first iteration is 100%. Is the solution diverging? Repeating more iterations, the following values are obtained

Iteration	u	$\lvert\in_{n+1}\rvert_u$ (%)	v	$\lvert\in_{n+1}\rvert_v$ (%)
0	0		0	
1	1.6667	100	2.5	100
2	0.8333	100	1.6667	50
3	1.1111	25	2.0833	20
4	0.9722	14.2857	1.9444	7.1429
5	1.0185	4.5455	2.0139	3.4483
6	0.9954	2.3256	1.9907	1.1628
7	1.0031	0.7692	2.0023	0.5780

From the table above, further steps of Jacobi show convergence toward the solution, which is $u = 1$ and $v = 2$. In this example, the convergence is guaranteed as the system is diagonally dominant.

> Do you know?
> Jacobi Method obeys linear convergence. What is a linear convergence?
>
> Definition:
> Let e_i denote the error at step *i* of an iterative method. If
>
> $$\lim_{i\to\infty} \frac{e_i + 1}{e_i} = S < 1$$
>
> the method is said to obey **linear convergence** with rate *S*.

12. GAUSS-SEIDEL METHOD

Closely related to the Jacobi Method is an iteration called Gauss-Seidel Method. The only difference between Gauss-Seidel and Jacobi is that in the former, the most recently updated values of the unknowns are used at each step, even if the updating occurs in the current step. Therefore, in Gauss-Seidel iteration, $\boldsymbol{Q}$ is taken as a lower triangular part of $\boldsymbol{A}$, including the diagonal. We write the Gauss-Seidel method as

$$\boldsymbol{x}^{(n+1)}=\boldsymbol{B}\boldsymbol{x}^{(n)}+\boldsymbol{h}$$

where $\boldsymbol{B} = \boldsymbol{I} - \boldsymbol{Q}^{-1}\boldsymbol{A}$, $\boldsymbol{h} = \boldsymbol{Q}^{-1}\boldsymbol{b}$ and n=0,1,2,3,4……

For example, $\boldsymbol{A}=\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix} 5 \\ 5 \end{bmatrix}$

$$\boldsymbol{Q} = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}$$

Gauss-Seidel often converges faster than Jacobi if the method is convergent. Like Jacobi, Gauss-Seidel Method converges to the solution as long as the coefficient matrix is strictly diagonally dominant. The Spectral Radius Theorem is also applicable to the Gauss-Seidel method to prove that Gauss-Seidel method converges for all the initial guesses

Example 23: Apply the Gauss-Seidel Method to the system

$$12u + 3v - 5w = 1$$
$$u + 5v + 3w = 28$$
$$3u + 7v + 13w = 76$$

With an initial guess of

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Step 1: Check if the coefficient matrix is strictly diagonally dominant. If necessary, exchange rows to fulfil this condition.

$$|12| > |3| + |-5| = 8$$
$$|5| > |1| + |3| = 4$$
$$|13| > |3| + |7| = 10$$

Step 2: Construct the equations for u, v, and w.

Solve equation of row 1 for *u*.
Equation of row 1,

$$u_{k+1} = \frac{1 - 3v_k + 5w_k}{12}$$

Solve equation of row 2 for *v*.
Equation of row 2,

$$v_{k+1} = \frac{28 - u_{k+1} - 3w_k}{5}$$

Solve equation of row 3 for *w*.
Equation of row 3,

$$w_{k+1} = \frac{76 - 3u_{k+1} - 7v_{k+1}}{13}$$

Step 3: Use the initial guess and iterate the constructed equations.

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$u_1 = \frac{1 - 3(0) + 5(1)}{12} = 0.50000$$

$$v_1 = \frac{28 - (0.5) - 3(1)}{5} = 4.9000$$

$$w_1 = \frac{76 - 3(0.50000) - 7(4.9000)}{13} = 3.0923$$

After the first iteration, the absolute relative approximate errors are

$$|\in_1|_u = \left|\frac{0.50000 - 1.0000}{0.50000}\right| \times 100 = 100.00\%$$

$$|\in_1|_v = \left|\frac{4.9000 - 0}{4.9000}\right| \times 100 = 100.00\%$$

$$|\in_1|_w = \left|\frac{3.0923 - 1.0000}{3.0923}\right| \times 100 = 67.662\%$$

The maximum percentage relative error after the first iteration is 100%. Is the solution diverging? No, as you conduct more iterations, the solution converges as follows. This is because the system is strictly diagonally dominant, and therefore the iteration will converge to the exact solution

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$$

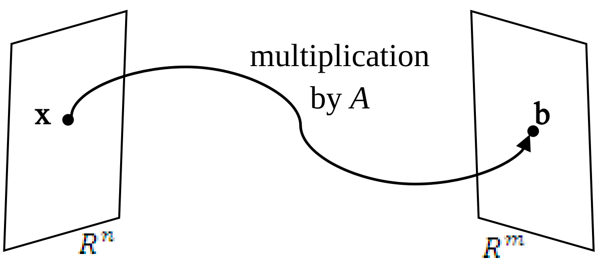
Iteration	u	$\left\|\in_{n+1}\right\|_u$ (%)	v	$\left\|\in_{n+1}\right\|_v$ (%)	w	$\left\|\in_{n+1}\right\|_w$ (%)
0	1		0		1	
1	0.50000	100.00	4.9000	100.00	3.0923	67.662
2	0.14679	240.61	3.7153	31.889	3.8118	18.876
3	0.74275	80.236	3.1644	17.408	3.9708	4.0042
4	0.94675	21.546	3.0281	4.4996	3.9971	0.65772
5	0.99177	4.5391	3.0034	0.82499	4.0001	0.074383
6	0.99919	0.74307	3.0001	0.10856	4.0001	0.00101

Note the difference between Gauss-Seidel and Jacobi:

The definition of v_{k+1} uses the expression of u_{k+1} not u_k. Similarly, the definition of w_{k+1} uses the expression of both u_{k+1} and v_{k+1}.

13. INTRODUCTION TO LINEAR TRANSFORMATIONS

The difference between a matrix equation $A\mathbf{x} = \mathbf{b}$ and the associated vector equation $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$ is merely a matter of notation. However, a matrix equation $A\mathbf{x} = \mathbf{b}$ can arise in applications such as computer graphics and image processing in a way that is not directly connected with linear combinations of vectors. This happens when the matrix A becomes an object that "acts" on a vector $\mathbf{x}$ by multiplication to produce a new vector called $A\mathbf{x}$. In such a case, it can be said that a multiplication by A transforms $\mathbf{x}$ into $\mathbf{b}$.



From this point of view, solving the equation $A\mathbf{x} = \mathbf{b}$ amounts to finding all vectors $\mathbf{x}$ in R^n that are transformed into the vector $\mathbf{b}$ in R^m under the "action" of multiplication by A. A **transformation** (or **function** or **mapping**) T from R^n to R^m is a rule that assigns to each vector **x** in R^n a vector $T(\mathbf{x})$ in R^m.

The correspondence from **x** to $A\mathbf{x}$ is a **function** from one set of vectors to another. The set R^n is called the **domain** of T, and the set R^m is called the **codomain** of T. The notation $T : R^n \to R^m$ indicates that the domain of T is R^n and the codomain in R^m. For **x** in R^n, the vector $T(\mathbf{x})$ in R^m is called the **image** of **x** (under the action of T). The set of all images $T(\mathbf{x})$ is called the **range** of T. Observe that the domain of T is when A has n columns and the codomain of T is when each column of A has m entries. For simplicity, sometimes **matrix transformation** is denoted by $\mathbf{x} \mapsto A\mathbf{x}$.

Every matrix transformation is a linear transformation. Linear transformation preserves the operations of vector addition and scalar multiplication.

Definition:

A transformation (or mapping) T is **linear** if:

1. $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all **u**, **v** in the domain of T;
2. $T(c\mathbf{u}) = cT(\mathbf{u})$ for all **u** and all scalars of c.

Theorem:

Let $T : R^n \to R^m$ be a linear transformation. Then there exists a unique matrix A such that

$$T(\mathbf{x}) = A\mathbf{x} \text{ for all } \mathbf{x} \text{ in } R^n$$

In fact, A is the m x n matrix whose jth column is the vector $T(\mathbf{e}_j)$, where $\mathbf{e}_j$ is the jth column of the identity matrix in R^n:

$$A = [\, T(\mathbf{e}_1) \ldots T(\mathbf{e}_n)]$$

Matrix A is called the **standard matrix for the linear transformation** T.

14. EIGENVECTORS AND EIGENVALUES: LINEAR SYSTEM OF $A\mathbf{x} = \lambda\mathbf{x}$.

The topic on eigenvectors {pronounce as **ˈīgən ˌvektər** (ei-gen-vec-tor)} dissects the action of a linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ into elements that are easily visualized. Although such transformation may move vectors in a variety of directions, it often happens that there are special vectors on which the action of A is quite simple. Eigenvectors are vectors that are transformed by A into a scalar multiple of themselves. Such system amounts to solving a linear system of $A\mathbf{x} = \lambda\mathbf{x}$.

> Definition:
> An **eigenvector** of an n x n matrix A is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ. A scalar λ is called an **eigenvalue** of A is there is a nontrivial solution $\mathbf{x}$ of $A\mathbf{x} = \lambda\mathbf{x}$; such an $\mathbf{x}$ is called an *eigenvector corresponding to* λ.

Intuitively, to solve the linear system of $\mathbf{x} = \lambda\mathbf{x}$, one can performed,

$$A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0}$$

Since A is usually a square n x n matrix and λ is a scalar, to solve the equation requires ingenuity where λ is multiplied to an identity matrix I_n as,

$$A\mathbf{x} - \lambda I\mathbf{x} = \mathbf{0} \quad \Rightarrow \quad (A - \lambda I)\mathbf{x} = \mathbf{0}$$

This matrix equation is a type of a homogeneous linear system, which *always* has a trivial solution of $\mathbf{x} = \mathbf{0}$. It involves *two* unknowns, λ and $\mathbf{x}$.

By the definition above, an eigenvector $\mathbf{x}$ *must be nonzero*, but an eigenvalue may be zero. Then $\mathbf{x} = \mathbf{0}$ is not the solution of the linear system. The solution is the nontrivial one, where, it must be that $(A - \lambda I) = \mathbf{0}$, where this relates to matrix inversion and determinant.

Recall that if a matrix A is not invertible, its det A = 0. The equation $(A - \lambda I) = \mathbf{0}$, can be solved by finding the non invertible matrix $(A - \lambda I)$ using det$(A - \lambda I)$ = 0. Such scalar equation det$(A - \lambda I) = 0$ is known as the characteristic equation of the linear system.

The set of all solutions of the linear system $(A - \lambda I)\mathbf{x} = \mathbf{0}$ is just the null space of the matrix $(A - \lambda I)$. So this set is a subspace of R^n and is called the **eigenspace** of A

corresponding to λ. The eigenspace consists of the zero vector and all the eigenvectors corresponding to λ.

Side Notes: What is an eigenspace?

An eigenspace is a subspace (subset) of a vector space. A **vector space** V is a nonempty set of objects, called vectors, on which two operations are defined, called addition and multiplication by scalars (real numbers), subject to 10 axioms (rules) listed below. The axioms must hold for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and for all scalars c and d.

1. The sum of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u} + \mathbf{v}$, is in V.
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
4. There is a zero vector $\mathbf{0}$ in V such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
5. For each $\mathbf{u}$ in V, there is a vector $-\mathbf{u}$ in V such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. The scalar multiple of $\mathbf{u}$ by c, denoted by $c\mathbf{u}$, is in V.
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$.
10. $1\mathbf{u} = \mathbf{u}$.

A **subspace** of R^n is any set H in R^n that has these three properties:

1. The zero vector is in set H.
2. For each $\mathbf{u}$ and $\mathbf{v}$ in H, the sum $\mathbf{u} + \mathbf{v}$ is in H.
3. For each $\mathbf{u}$ in H and each scalar c, the vector $c\mathbf{u}$ is in H.

The **null space** of a matrix A, is the set of all solutions to the homogeneous equation $A\mathbf{x} = \mathbf{0}$. The null space is a subspace and denoted as **Nul** $\boldsymbol{A}$.

Because an eigenspace (a subspace) typically contains an infinite number of vectors, some problems involving a subspace are handled best by working with a small finite set of vectors that span the subspace; the smaller the set, the better. Such small finite set of vectors is known as a basis (plural form is bases). It would be useful then to express the solution of the linear system $(A - \lambda I)\mathbf{x} = \mathbf{0}$ by its bases.

Definition:

A basis for a subspace H of R^n is a linearly independent set in H that spans H.

Example 24: Find the eigenvalues of *A*.

$$A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$$

Find all the scalars λ such that the matrix equation, $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution.

$A - \lambda I = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix}$

det $(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix} = 0$

Recall that $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc,$

$$\det (A - \lambda I) = (2-\lambda)(-6-\lambda) - (3)(3) = -12 + 6\lambda - 2\lambda + \lambda^2 - 9$$

$$-12 + 6\lambda - 2\lambda + \lambda^2 - 9 = \lambda^2 + 4\lambda - 21 = (\lambda - 3)(\lambda + 7) = 0$$

$$\Rightarrow \lambda_1 = 3, \quad \lambda_2 = -7$$

The eigenvalues of *A* are 3 and -7.

Self-Test: Is 7 an eigenvalue of $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$?

Example 25: Find the eigenvalues and its corresponding eigenvectors, and bases for matrix *A*.

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

Step 1: Construct the characteristic equation and find the eigenvalues.

$A - \lambda I = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 3-\lambda & -2 \\ 1 & -\lambda \end{bmatrix}$

$\det(A - \lambda I) = \det \begin{bmatrix} 3-\lambda & -2 \\ 1 & -\lambda \end{bmatrix} = (3-\lambda)(-\lambda) - (-2)(1) = \lambda^2 - 3\lambda + 2$

$\lambda^2 - 3\lambda + 2 = (\lambda - 2)(\lambda - 1)$

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

Step 2: Compute the eigenvector for $\lambda_1 = 2$.

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

$\left[\begin{array}{cc|c} 1 & -2 & 0 \\ 1 & -2 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -2 & 0 \\ 0 & 0 & 0 \end{array}\right]$

$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad x_1 - 2x_2 = 0 \quad \Rightarrow \quad x_1 = 2x_2$

$x_1 = 2x_2 \quad \Rightarrow \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Step 3: Compute the eigenvector for $\lambda_2 = 1$.

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

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

$$x_1 \begin{bmatrix} 2 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow 2x_1 - 2x_2 = 0 \Rightarrow x_1 = x_2$$

$$x_1 = x_2 \Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Step 4: Construct the bases.

Basis for $\lambda_1 = 2$ is $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and basis for $\lambda_2 = 1$ is $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

<u>Do you know?</u>

In many cases, the eigenvalues-eigenvector information contained within a matrix *A* can be displayed in a useful factorization of the form $A = PDP^{-1}$. The factorization enables the computation of A^k quickly for large values of *k* that can be used in performing linear transformation. The *D* in the factorization stands for *diagonal*.

Unfortunately, not all matrices can be factored as $A = PDP^{-1}$. However, a factorization $A = QDP^{-1}$ is possible for any *m* x *n* matrix *A*. This special factorization is called the **singular value decomposition**.

<u>*Example 26*</u>: Find the eigenvalues and determine whether the Jacobi and Gauss-Seidel methods of the *Example 22* converge for all the initial guesses.

$$3u + v = 5$$
$$u + 2v = 5$$

$$\boldsymbol{A} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

For Jacobi method:

$\boldsymbol{Q} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$, $\boldsymbol{Q}^{-1}\boldsymbol{A} = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{3} \\ \frac{1}{2} & 1 \end{bmatrix}$

$\boldsymbol{B} = \boldsymbol{I} - \boldsymbol{Q}^{-1}\boldsymbol{A} = \begin{bmatrix} 0 & -\frac{1}{3} \\ -\frac{1}{2} & 0 \end{bmatrix}$, $B - \lambda I = \begin{bmatrix} 0 & \frac{1}{3} \\ -\frac{1}{2} & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} -\lambda & -\frac{1}{3} \\ -\frac{1}{2} & -\lambda \end{bmatrix}$

$$\det(B - \lambda I) = \det\begin{bmatrix} -\lambda & -\frac{1}{3} \\ -\frac{1}{2} & -\lambda \end{bmatrix} = \lambda^2 - \frac{1}{6}$$

$\lambda = \pm\sqrt{\frac{1}{6}}$, spectral radius $\rho = \sqrt{\frac{1}{6}} < 1$. Thus, by the Spectral Radius Theorem, the Jacobi method succeeds for any starting initial vector in this example.

For Gauss-Seidel method:

$\boldsymbol{Q} = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}$, $\boldsymbol{Q}^{-1}\boldsymbol{A} = \begin{bmatrix} \frac{1}{3} & 0 \\ -\frac{1}{6} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{3} \\ 0 & \frac{5}{6} \end{bmatrix}$

$\boldsymbol{B} = \boldsymbol{I} - \boldsymbol{Q}^{-1}\boldsymbol{A} = \begin{bmatrix} 0 & -\frac{1}{3} \\ 0 & \frac{1}{6} \end{bmatrix}$, $B - \lambda I = \begin{bmatrix} 0 & -\frac{1}{3} \\ 0 & \frac{1}{6} \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} -\lambda & -\frac{1}{3} \\ 0 & \frac{1}{6} - \lambda \end{bmatrix}$

$$\det(B - \lambda I) = \det\begin{bmatrix} -\lambda & -\frac{1}{3} \\ 0 & \frac{1}{6} - \lambda \end{bmatrix} = \lambda^2 - \frac{1}{6}\lambda$$

$\lambda = 0, \frac{1}{6}$, spectral radius $\rho = \frac{1}{6} < 1$. Thus, by the Spectral Radius Theorem, the Gauss-Seidel method succeeds for any starting initial vector in this example.