Indian Institute of Technology Delhi-Abu Dhabi
Course: AMTL 101 (Linear Algebra)
Semester II, 2024-2025
Instructor: Prof. Amit Priyadarshi

Contents

System of Linear Equations

1 Introduction

In this chapter, we study the fundamental concepts of solving systems of linear equations. We will explore the existence and uniqueness of solutions, as well as methods for determining these solutions when they exist. To illustrate these concepts, we will examine several examples of linear systems, providing a comprehensive understanding of the underlying principles.

A **linear equation** in the variables $x_1, x_2, \ldots, x_n$ is an equation that can be written in the form

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

where b and the coefficient $a_1, a_2, \ldots, a_n$ are real numbers.

Let's start with a system of two linear equations in two variables. Consider the equations

$$a_1x + b_1y = c_1 \text{ and } a_2x + b_2y = c_2$$

We say that (x_0, y_0) is a solution of the system if it satisfies both of the equations. Geometrically, if one of the coefficients, a or b is non-zero, then these linear equations represent a line in $\mathbb{R}^2$. Thus for the system

$$a_1x + b_1y = c_1 \text{ and } a_2x + b_2y = c_2$$

the set of solutions is given by the points of intersection of the two lines. Since in $\mathbb{R}^2$, two straight lines either intersect at a unique point or never intersect. So, the following three cases are possible:

1. Unique Solution:
   For example, consider $x+2y=1$ and $x+3y=1$. The unique solution is $(x_0, y_0) = (1, 0)$.
   Observe that in this case, $a_1b_2 - a_2b_1 \neq 0$.

2. Infinite Number of Solutions:
   consider $x+2y=1$ and $2x+4y=2$. The set of solutions is

$$(x_0, y_0) = (1-2y, y) = (1, 0) + y(-2, 1)$$

   with y arbitrary. In other words, both the equations represent the same line.
   Observe that in this case, $a_1b_2 - a_2b_1 = 0, a_1c_2 - a_2c_1 = 0$ and $b_1c_2 - b_2c_1 = 0$.

3. No Solution:
   consider $x+2y=1$ and $2x+4y=3$. The equations represent a pair of parallel lines and hence there is no point of intersection.
   Observe that in this case, $a_1b_2 - a_2b_1 = 0$ but $a_1c_2 - a_2c_1 \neq 0$.

2 Definition and Example

Definition 1. *(Linear System) A linear system of m equations in n unknowns $x_1, x_2, \ldots, x_n$ is a set of equations of the form*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{2.1}$$

where $a_{ij}, b_i \in \mathbb{R}$ for $1 \leq i \leq m$, and $1 \leq j \leq n$.

- Linear System (2.1) is called homogeneous if $b_1 = 0 = b_2 = \cdots = b_m$ and non-homogeneous otherwise.

The above system can also be written in the form $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \text{ and } \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

The matrix A is called the coefficient matrix, and the block matrix $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$ is called the augmented matrix of the linear system (2.1).

Observe that the i^{th} row of the augmented matrix $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$ represents the i^{th} equation and the j^{th} column of the coefficient matrix A corresponds to coefficients of the j^{th} variable x_j. That is, for $1 \leq i \leq m$ and $1 \leq j \leq n$, the entry a_{ij} of the coefficient matrix A corresponds to the i^{th} equation and j^{th} variable x_j.

Definition 2. *For a system of linear equations $A\mathbf{x} = \mathbf{b}$, the system $A\mathbf{x} = \mathbf{0}$ is called the* ***associated homogeneous system****.*

Definition 3. *(Solution of a Linear System) A solution of the linear system $A\mathbf{x} = \mathbf{b}$ is a column vector $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$ such that $Ay = b$. That is, the linear system (2.1) is satisfied by substituting y_i in place of x_i.*

- Note that $\mathbf{x} = \mathbf{0}$ is always a solution of the system $A\mathbf{x} = \mathbf{0}$, and is called the trivial solution. A non-zero n-tuple $\mathbf{x}$, if it satisfies $A\mathbf{x} = \mathbf{0}$, is called a **non-trivial** solution.

Example 1. Solve the following linear system

$$\begin{aligned} 2x + y - z &= 3 \\ x - y - z &= 0 \\ x + y + 3z &= 12. \end{aligned}$$

Solution Observe that the above linear system and the linear system

$$\begin{aligned} x - y - z &= 0 \\ 2x + y - z &= 3 \\ x + y + 3z &= 12 \end{aligned}$$

have the same set of solutions.

Eliminating x from 2^{nd} and 3^{rd} equations, we get the linear system

$$\begin{aligned} x - y - z &= 0 \\ 3y + z &= 3 \\ 2y + 4z &= 12. \end{aligned}$$

It can be easily shown that this system and the above system have the same set of solutions.

Now, eliminating y from the last equation, we get the system

$$\begin{aligned} x - y - z &= 0 \\ 3y + z &= 3 \\ 10z &= 30 \end{aligned}$$

which has the same set of solutions as the above system has.

Now, from equation 3, we get $z = 3$, substituting the value of z in equation 2, we get $y = 0$, and from equation 1, we get $x = y + z = 3$.

Therefore, the system has a unique solution $(3, 0, 3)$. ■

3 Row Operations and Equivalent Systems

Definition 4. *(Elementary Operations) The following three operations are called elementary operations, where E_i represents the i^{th} equation in the linear system.*

1. *Interchanging any two equations ($E_i \leftrightarrow E_j$).*
2. *Multiply a non-zero constant throughout an equation ($E_i \to kE_i,\ k \neq 0$).*
3. *Replace an equation by itself plus a constant multiple of another equation ($E_i \to E_i + kE_j$).*

Remark 1. *Observe that, in the above example, we applied finitely many elementary operations, and the system we got at the end was easily solvable.*

Definition 5. *(Equivalent Linear Systems) Two linear systems are said to be equivalent if one can be obtained from the other by a finite number of elementary operations.*

Now, we show that equivalent linear systems have the same set of solutions.

Theorem 1. *Let $A\mathbf{x} = \mathbf{b}$ be a linear system of m equations in n unknowns. Let $C\mathbf{x} = \mathbf{d}$ be the linear system obtained from the linear system $A\mathbf{x} = \mathbf{b}$ by a single elementary operation. Then the linear systems $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ have the same set of solutions.*

Proof. We prove the result for the third elementary operation where the k^{th} equation is replaced by k^{th} equation plus c times the j^{th} equation, i.e. $E_k \to E_k + cE_j$. The proof for the other two elementary operations is straight forward.

In this case, the systems $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ vary only in the k^{th} equation. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ be a solution of the linear system $A\mathbf{x} = b$. Then substituting for α_i 's in place of x_i 's in the k^{th} and j^{th} equations, we get

$$a_{k1}\alpha_1 + a_{k2}\alpha_2 + \cdots a_{kn}\alpha_n = b_k, \text{ and } a_{j1}\alpha_1 + a_{j2}\alpha_2 + \cdots a_{jn}\alpha_n = b_j$$

Therefore,

$$(a_{k1} + ca_{j1})\,\alpha_1 + (a_{k2} + ca_{j2})\,\alpha_2 + \cdots + (a_{kn} + ca_{jn})\,\alpha_n = b_k + cb_j$$

But then the k^{th} equation of the linear system $C\mathbf{x} = \mathbf{d}$ is

$$(a_{k1} + ca_{j1})\,x_1 + (a_{k2} + ca_{j2})\,x_2 + \cdots + (a_{kn} + ca_{jn})\,x_n = b_k + cb_j.$$

Therefore, $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is also a solution for the linear system $C\mathbf{x} = \mathbf{d}$.

Similarly, if $(\beta_1, \beta_2, \ldots, \beta_n)$ is a solution of the linear system $C\mathbf{x} = \mathbf{d}$ then it is also a solution of the linear system $A\mathbf{x} = \mathbf{b}$.

Hence, both linear systems have the same set of solutions. $\square$

Using mathematical induction and the above theorem, we have the following result:

Theorem 2. *Two equivalent systems have the same set of solutions.*

Note that, while solving the system of linear equations, we are doing the calculation with the coefficients only, therefore, in place of looking at the system of equations as a whole, we just need to work with the coefficients. These coefficients when arranged in a rectangular array gives us the augmented matrix $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$.

Definition 6. *(Elementary Row Operations) The following three operations are called elementary row operations, where R_i represents the i^{th} row of the matrix.*

1. *Interchanging any two rows ($R_i \leftrightarrow R_j$).*

2. *Multiply a non-zero constant throughout a row ($R_i \to kR_i,\ k \neq 0$).*

3. *Replace a row by itself plus a constant multiple of another row ($R_i \to R_i + kR_j$).*

Definition 7. *(Row Equivalent Matrices) Two matrices are said to be row-equivalent if one can be obtained from the other by a finite number of elementary row operations.*

Example 2. The three matrices given below are row equivalent.

$$\begin{bmatrix} 0 & 2 & 3 & 4 \\ 2 & 5 & 7 & 3 \\ 3 & 0 & 1 & 9 \end{bmatrix} \xrightarrow[R_1 \leftrightarrow R2]{} \begin{bmatrix} 2 & 5 & 7 & 3 \\ 0 & 2 & 3 & 4 \\ 3 & 0 & 1 & 9 \end{bmatrix} \xrightarrow[R_1 \to (1/2)R_1]{} \begin{bmatrix} 1 & \frac{5}{2} & \frac{7}{2} & \frac{3}{2} \\ 0 & 2 & 3 & 4 \\ 3 & 0 & 1 & 9 \end{bmatrix}.$$

Whereas the matrix $\begin{bmatrix} 0 & 2 & 3 & 4 \\ 2 & 5 & 7 & 3 \\ 3 & 0 & 1 & 9 \end{bmatrix}$ is not row equivalent to the matrix $\begin{bmatrix} 2 & 5 & 3 & 4 \\ 0 & 2 & 7 & 3 \\ 3 & 0 & 1 & 9 \end{bmatrix}$.

4 Row Reduction and Echelon Forms

Definition 8. *(Row Echelon Form) A rectangular matrix is in echelon form if it has the following three properties:*

1. *All non-zero rows are above any rows of all zeros.*
2. *Each leading entry of a row (leftmost nonzero entry of that row) is in a column to the right of the leading entry of the row above it.*
3. *All entries in a column below a leading entry are zero.*

Definition 9. *(Row Reduced Echelon Form) If a matrix in echelon form satisfies the following additional conditions, then it is in reduced echelon form.*

1. *The leading entry in each nonzero row is 1.*
2. *Each leading 1 is the only nonzero entry in its column.*

Method to get the row-reduced echelon form of a given matrix A:
Let A be an $m \times n$ matrix. Then the following method is used to obtain the row-reduced echelon form of the matrix A.

Step 1: Consider the first column of the matrix A.
If all the entries in the first column are zero, move to the second column.
Else, find a row, say i^{th} row, which contains a non-zero entry in the first column. Now, interchange the first row with the i^{th} row. Suppose the non-zero entry in the (1,1)-position is $\alpha \neq 0$. Divide the whole row by α so that the $(1,1)$-entry of the new matrix is 1 . Now, use the 1 to make all the entries below this 1 equal to 0.

Step 2: If all entries in the first column after the first step are zero, start with the lower $(m-1) \times (n-1)$ submatrix of the matrix obtained in the first step and proceed as in step 1.

Step 3: Keep repeating this process till we reach a stage where all the entries below a particular row, say r, are zero. Suppose at this stage we have obtained a matrix C. Then C has the following form:

1. the first non-zero entry in each row of C is 1 . These 1 's are the leading terms of C and the columns containing these leading terms are the leading columns.

2. the entries of C below the leading term are all zero.

Step 4: Now use the leading term in the r^{th} row to make all entries in the r^{th} leading column equal to zero.

Step 5: Next, use the leading term in the $(r-1)^{\text{th}}$ row to make all entries in the $(r-1)^{\text{th}}$ leading column equal to zero and continue till we come to the first leading term or column.

The final matrix is the row-reduced echelon form of the matrix A.

Example 3. Apply elementary row operations to transform the following matrix first into echelon form and then into reduced echelon form:

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

Solution

$$\begin{bmatrix} 0 & 3 & -6 & 6 \\ 3 & -7 & 8 & -5 \\ 3 & -9 & 12 & -9 \end{bmatrix} \xrightarrow[R_1 \leftrightarrow R3]{} \begin{bmatrix} 3 & -9 & 12 & -9 \\ 3 & -7 & 8 & -5 \\ 0 & 3 & -6 & 6 \end{bmatrix}$$

$$\xrightarrow[R_2 \to R_2 - R1]{} \begin{bmatrix} 3 & -9 & 12 & -9 \\ 0 & 2 & -4 & 4 \\ 0 & 3 & -6 & 6 \end{bmatrix}$$

$$\xrightarrow[R_3 \to R_3 - \frac{3}{2}R_2]{} \begin{bmatrix} 3 & -9 & 12 & -9 \\ 0 & 2 & -4 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

the matrix A is in echelon form now. To get the row reduced echelon form, we apply

$$\xrightarrow[\substack{R_1 \to \frac{1}{3}R_1 \\ R_2 \to \frac{1}{2}R_2}]{} \begin{bmatrix} 1 & -3 & 4 & -3 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow[R_1 \to R_1 + 3R_2]{} \begin{bmatrix} 1 & 0 & -2 & 3 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

which is the row reduced echelon form of A. ■

Remark 2. *Observe that all the matrices obtained in the process to find the row reduced echelon form are row equivalent to each other.*

5 Gauss Elimination Method

Definition 10. *(Forward/Gauss Elimination Method) Gaussian elimination is a method of solving a linear system $A\mathbf{x} = \mathbf{b}$ (consisting of m equations in n unknowns) by converting the augmented matrix*

$$\begin{bmatrix} A & \mathbf{b} \end{bmatrix} = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

to an upper triangular form or echelon form

$$\left[\begin{array}{cccc|c} c_{11} & c_{12} & \cdots & c_{1n} & d_1 \\ 0 & c_{22} & \cdots & c_{2n} & d_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & c_{mn} & d_m \end{array}\right].$$

This elimination process is also called the forward elimination method.

The following examples illustrate the Gauss elimination procedure.

Example 4. Solve the linear system by Gauss elimination method.

$$\begin{aligned} x+y+z &= 3 \\ x+2y+2z &= 5 \\ 3x+4y+4z &= 11 \end{aligned}$$

Solution In this case, the augmented matrix is $\begin{bmatrix} 1 & 1 & 1 & 3 \\ 1 & 2 & 2 & 5 \\ 3 & 4 & 4 & 11 \end{bmatrix}$ and the method proceeds as follows:

1. Add -1 times the first equation to the second equation.

$$\begin{array}{ll} x+y+z & =3 \\ y+z & =2 \\ 3x+4y+4z & =11 \end{array} \quad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 3 & 4 & 4 & 11 \end{bmatrix}.$$

2. Add -3 times the first equation to the third equation.

$$\begin{array}{ll} x+y+z & =3 \\ y+z & =2 \\ y+z & =2 \end{array} \quad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 2 \end{bmatrix}$$

3. Add -1 times the second equation to the third equation.

$$\begin{array}{ll} x+y+z & =3 \\ y+z & =2 \end{array} \quad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Thus, for any $z \in \mathbb{R}$, $(x, y, z)' = (1, 2-z, z)'$ is a solution of the linear system. In other words, the system has infinite number of solutions. ■

Remark 3. *Note that to solve a linear system,* $A\mathbf{x} = \mathbf{b}$, *one needs to apply only the elementary row operations to the augmented matrix* $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$ *to obtain the echelon form.*

Definition 11. *(Pivot or leading entry, Pivot column) For a matrix in echelon form or reduced echelon form, the first non-zero entry of any row is called a pivot entry. The columns containing the pivot entries are called the pivot columns.*

Definition 12. *(Basic, Free Variables) Consider the linear system* $A\mathbf{x} = \mathbf{b}$ *in* n *variables and* m *equations. Let* $\begin{bmatrix} C & \mathbf{d} \end{bmatrix}$ *be the row-reduced echelon matrix obtained by applying the Gauss elimination method to the augmented matrix* $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$*. Then the variables corresponding to the pivot columns in the first* n *columns of* $\begin{bmatrix} C & \mathbf{d} \end{bmatrix}$ *are called the basic variables. The variables which are not basic are called free variables.*

The free variables are called so as they can be assigned arbitrary values and the value of the basic variables can then be written in terms of the free variables.

Observation: In Example 4, the solution set was given by

$$(x, y, z)' = (1, 2 - z, z)' = (1, 2, 0)' + z(0, -1, 1)', \text{ with } z \text{ arbitrary.}$$

That is, we had two basic variables, x and y, and z as a free variable.

Remark 4. *Observe that, if for the system* $Ax = b$*, there are* r *non-zero rows in the row-reduced echelon form of the matrix* A *then there will be* r *pivots. That is, there will be* r *pivot columns. Therefore, If There are* r *pivots and* n *variables, then there will be* r *basic variables and* $n - r$ *free variables.*

5.1 Gauss-Jordan Elimination

The elimination process applied to obtain the row reduced echelon form of the augmented matrix is called the Gauss-Jordan elimination.

That is, the Gauss-Jordan elimination method consists of both the forward elimination and the backward substitution.

Example 5. In example 4, we can apply the Gauss-Jordan elimination. For that we just need that the pivots are the only nonzero entries in their columns.

After the step 3, adding -1 times the second equation to the first equation, we get

$$\begin{aligned} x \quad &= 1 \\ y + z \quad &= 2 \end{aligned} \qquad \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

From the above matrix, we directly have the set of solution as

$$\{(x, y, z)' = (1, 2 - z, z)';\ z \in \mathbb{R}\}.$$