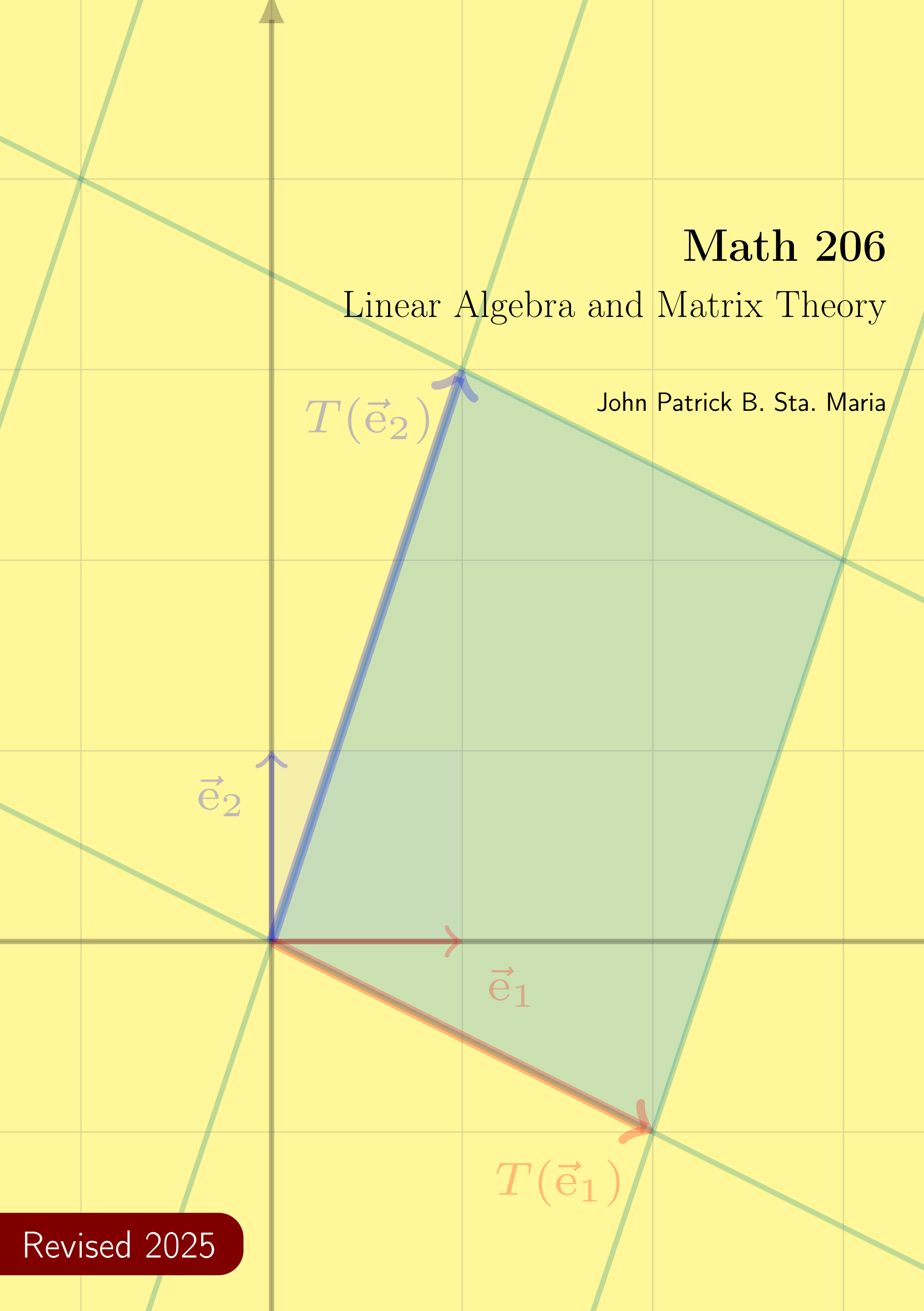
Math 206

Linear Algebra and Matrix Theory

John Patrick B. Sta. Maria



Revised 2025

Unit 2

Linear Systems and Determinants

Lesson 2.1: Solutions to Linear Systems

Learning Outcomes: At the end of the lesson, the students are able to

1. perform Gaussian algorithm to reduce a matrix into its reduced row echelon form.
2. solve a linear system by the Gauss-Jordan method.
3. model real-life problems using linear systems.

Solving a Linear System

We start by defining a linear equation. A **linear equation** in the unknowns $x_1, x_2, \ldots, x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \tag{2.1}$$

where $a_1, \ldots, a_n, b \in \mathbb{R}$. A **solution** to (2.1) is an n-vector

$$\vec{s} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} \in \mathbb{R}^n$$

such that $a_1s_1 + a_2s_2 + \cdots + a_ns_n = b$ is true.

Example 58. Consider the linear equation

$$4x_1 + 3x_2 + x_3 + 2x_4 = 5.$$

The vector $\begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix}$ is a solution since

$$4(1) + 3(-1) + 0 + 2(2) = 5 \text{ is true.}$$

On the other hand, the vector $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ is **not** a solution since

$$4(1) + 3(1) + 1 + 2(1) = 5 \text{ is false.}$$

Example 59. The following equations are **not** linear. (Why?)

- $x_1^2 + 3x_2^2 = 4$
- $4x + 5xy + 9y = 7$
- $\cos x_1 + \sin x_2 = 1$

Degenerate Cases

The equation $0x_1 + 0x_2 + \cdots + 0x_n = b$, is called the **degenerate** linear equation.

1. If $b = 0$, then every vector of $\mathbb{R}^n$ is a solution.

2. If $b \neq 0$, then the given equation has no solution in $\mathbb{R}^n$.

Definition 60: Linear System

Let m and n be positive integers. An m-by-n **linear system** is a conjunction of m linear equations each in n unknowns. We usually represent an m-by-n linear system as

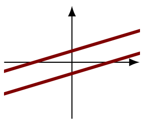
$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases} \tag{2.2}$$

where $a_{ij}, b_i \in \mathbb{R}$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$. A **solution** to the linear system is any vector in $\mathbb{R}^n$ which is a solution to every linear equation in the system. We say that a linear system is **consistent** if it has a solution, otherwise, we say the system is **inconsistent.**

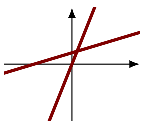
For example, a 2-by-2 linear system in the unknowns x and y is usually presented as

$$\begin{cases} ax + by = c \\ dx + ey = f \end{cases}$$

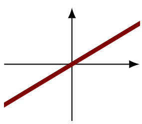
where a, b, c, d, e, f are constants. This linear system represent a pair of lines in the cartesian plane, where a solution is any point of intersection. In this case, there are three possibilities: no solution, unique solution, or infinitely many



no solution



unique solution



infinitely many solutions

Geometrically, every equation in the system represents a **hyperplane**. Hence, a solution to the system are points of intersection of each of the hyperplanes. In the case when $n = 2$, a hyperplane is simply a straight line in the cartesian plane $\mathbb{R}^2$, and when $n = 3$, a hyperplane is simply a flat plane in a 3-dimensional cartesian coordinate system $\mathbb{R}^3$. Two linear systems are said to be **equivalent** if they have the same solution sets. In other words, solving one from a set of equivalent linear system will also serve another.

Example 61. The given pair of linear systems below are equivalent.

$$\begin{cases} x + y = 3 \\ 2x - y = 0 \end{cases} \quad \text{and} \quad \begin{cases} 2x + 2y = 6 \\ \qquad 3y = 6 \end{cases}$$

These linear systems may be solved by the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. When we look at the linear systems above, notice that the second one is already *practically* solved. From the second equation, one yields the value of $y = 2$ and by substituting this value to y in the first equation yields $2x + 4 = 6$ or $x = 1$.

The following theorem allows us to obtain an equivalent linear system from a given linear system. The reason for this is self-explanatory.

Theorem 62: Gaussian Elimination

Suppose that a linear system $(\heartsuit)$ is obtained from a linear system $(\diamondsuit)$ by performing and of the following operations:

1. swapping two equations
2. multiplying both sides of an equation by a nonzero scalar
3. replacing one equation by its sum with a scalar multiple of another

then the linear systems $(\heartsuit)$ and $(\diamondsuit)$ are equivalent.

In other words, doing any of these three operations to a given linear system does not change the solution set.

Example 63. Consider the following linear system

$$\begin{cases} 2x + 5y - 9z = -10 & \text{(E1)} \\ x + 2y - 4z = -4 & \text{(E2)} \\ 3x - 2y + 3z = 11 & \text{(E3)} \end{cases}$$

By labeling the linear equations in the system as above, consider the following sequence of operations obtaining equivalent linear systems. This method is called the **Gaussian elmination** method.

Given:
$$\begin{cases} 2x+5y-9z=-10 & \text{(E1)} \\ x+2y-4z=-4 & \text{(E2)} \\ 3x-2y+3z=11 & \text{(E3)} \end{cases}$$

Swap (E1) and (E2).
$$\begin{cases} x+2y-4z=-4 & \text{(E1)} \\ 2x+5y-9z=-10 & \text{(E2)} \\ 3x-2y+3z=11 & \text{(E3)} \end{cases}$$

By swapping, we get an x as the first term in the first equation. Then using this x to eliminate the 'x' terms from the equations below it.

Replace (E2) by $\text{(E2)}+(-2)\cdot\text{(E1)}$.
$$\begin{cases} x+2y-4z=-4 & \text{(E1)} \\ y-z=-2 & \text{(E2)} \\ 3x-2y+3z=11 & \text{(E3)} \end{cases}$$

Replace (E3) by $\text{(E3)}+(-3)\cdot\text{(E1)}$.
$$\begin{cases} x+2y-4z=-4 & \text{(E1)} \\ y-z=-2 & \text{(E2)} \\ -8y+15z=23 & \text{(E3)} \end{cases}$$

At this point, notice that we have *eliminated* the variable 'x' on the second and third equations. Next, we are going to eliminate the 'y' term from the third equation using the 'y' term from second equation.

Replace (E3) by $\text{(E3)}+8\cdot\text{(E2)}$.
$$\begin{cases} x+2y-4z=-4 & \text{(E1)} \\ y-z=-2 & \text{(E2)} \\ 7z=7 & \text{(E3)} \end{cases}$$

Replace (E3) by $\frac{1}{7}\cdot\text{(E3)}$.
$$\begin{cases} x+2y-4z=-4 & \text{(E1)} \\ y-z=-2 & \text{(E2)} \\ z=1 & \text{(E3)} \end{cases}$$

We now stop the elimination process and solve for the values of x, y, and z using **back substitution**:

Beginning from the third equation, we see that $z = 1$. Substituting this value to the second equation yields:

$$y - 1 = -2 \implies y = -1.$$

Finally, substituting the z and y values to the first equation gives

$$x + 2(-1) - 4(1) = -4 \implies x = 2.$$

Therefore, the given linear system has a unique solution of

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}. \quad ▨$$

In the process presented in the above example, we see that the linear systems

$$\begin{cases} 2x + 5y - 9z = -10 \\ \quad x + 2y - 4z = -4 \\ 3x - 2y + 3z = 11 \end{cases} \text{and} \begin{cases} x + 2y - 4z = -4 \\ \qquad\quad y - \ \ z = -2 \\ \qquad\qquad\quad z = 1. \end{cases}$$

are equivalent.

Linear Systems as Matrices

In performing Gaussian elimination, notice how the roles of the unknowns say x, y, and z are minimal in the sense that they only act as *placeholders* for the numerical coefficients. Numerically speaking, one can argue that the solution is a function of the numerical coefficients.

Definition 64: Augmented Matrix of a Linear System

The **augmented matrix** of the m-by-n linear system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases} \tag{2.3}$$

is defined as the m-by-$(n+1)$ matrix

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

Example 65. The following shows the corresponding augmented matrices of the following equivalent linear systems:

Linear System	**Augmented Matrix**
$\begin{cases} 2x + 5y - 9z = -10 \\ x + 2y - 4z = -4 \\ 3x - 2y + 3z = 11 \end{cases}$	$\left[\begin{array}{ccc\|c} 2 & 5 & -9 & -10 \\ 1 & 2 & -4 & -4 \\ 3 & -2 & 3 & 11 \end{array}\right]$
$\begin{cases} x + 2y - 4z = -4 \\ y - z = -2 \\ -8y + 15z = 23 \end{cases}$	$\left[\begin{array}{ccc\|c} 1 & 2 & -4 & -4 \\ 0 & 1 & -1 & -2 \\ 0 & -8 & 15 & 23 \end{array}\right]$
$\begin{cases} x + 2y - 4z = -4 \\ y - z = -2 \\ z = 1 \end{cases}$	$\left[\begin{array}{ccc\|c} 1 & 2 & -4 & -4 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 1 \end{array}\right]$

Notice how the augmented matrix of the second linear system follows a special form. where entries below a 1 are 0s and the 1s are arranged in a certain 'ladder' form. The vertical line in the matrix indicates the separation of the left hand side of the linear equations from the right hand side.

As we shall see, representing a linear system by augmented matrices not only simplifies the elimination process, but also makes the experience more *aesthetic* and minimalistic, especially when we have to perform the process by hand calculation. This is why the following definition was formulated.

Definition 66: Matrices in Row Echelon Form

A matrix A is said to be in **row echelon form (REF)** if the following are satisfied:

(R1) Rows of zeroes, if there's any, appears at the bottom part of A.

(R2) For each nonzero row, the first nonzero entry from the left is a 1. This is called a **pivot** of the row.

(R3) Each pivot is in a column to the right of the pivot in the previous row.

A matrix in REF is said to be in **reduced row echelon form (RREF)** if (R4): every pivot is the only nonzero entry in its column.

Example 67. Consider the following matrices:

$$A = \begin{bmatrix} 1 & 4 & 0 & -2 & 1 \\ 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$
$$D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix}, E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The table on the right shows which among properties (R1) - (R4) are satisfied (✓) or not (×):

matrix	(R1)	(R2)	(R3)	(R4)
A	✓	✓	✓	✓
B	✓	✓	✓	×
C	×	✓	×	×
D	✓	×	−	−
E	✓	✓	×	✓

By looking at the table, we see that only matrices A and B are in row echelon form (REF) since they satisfy properties (R1), (R2), and (R3). However, only matrix A is in reduced row echelon form (RREF).

We remark that some matrices may satisfy (R4), yet they are not in RREF since it may be the case that it is not in REF. Take note that being in RREF **requires** one to be in REF. For example, look at matrix E.

Also, note that in order for (R3) and (R4) to make sense, (R2) must first be satisfied. As we can see from the table, the matrix D as an example, did not satisfy (R2), hence, it is pointless to check (R3) and (R4) (indicated by a $-$ on the table).

We now study how to convert a matrix to matrix in REF or RREF so that the linear systems they represent are equivalent.

Definition 68: Elementary Row Operations

An **elementary row operation** on a matrix is any of the following:

1. [Type I, E_{ij}] swap row i and row j
2. [Type II, $E_i(\alpha)$] multiply row i by a nonzero scalar α
3. [Type III, $E_{ij}(\alpha)$] replace row i by the sum of row i and α times row j

A matrix A is said to be **row equivalent** to a matrix B, expressed as $A \sim B$, if B can be obtained from A by applying a finite sequence of elementary row operations.

Notation. If an elementary row operation E is applied to a matrix A obtaining a matrix B, we write $A \xrightarrow{E} B$.

Example 69. Given the matrix $\begin{bmatrix} 1 & 2 & -1 \\ 3 & 4 & 0 \\ 1 & 3 & -7 \end{bmatrix}$.

- $\begin{bmatrix} 1 & 2 & -1 \\ 3 & 4 & 0 \\ 1 & 3 & -7 \end{bmatrix} \xrightarrow{E_{13}} \begin{bmatrix} 1 & 3 & -7 \\ 3 & 4 & 0 \\ 1 & 2 & -1 \end{bmatrix}$

  This is a type I operation that interchanges row 1 and row 3.

- $\begin{bmatrix} 1 & 2 & -1 \\ 3 & 4 & 0 \\ 1 & 3 & -7 \end{bmatrix} \xrightarrow{E_2(1/4)} \begin{bmatrix} 1 & 3 & -7 \\ 3/4 & 1 & 0 \\ 1 & 2 & -1 \end{bmatrix}$

  This is a type II operation that multiplies a scalar $1/4$ to row 2.

- $\begin{bmatrix} 1 & 2 & -1 \\ 3 & 4 & 0 \\ 1 & 3 & -7 \end{bmatrix} \xrightarrow{E_{21}(-3)} \begin{bmatrix} 1 & 3 & -7 \\ 0 & -2 & 3 \\ 1 & 2 & -1 \end{bmatrix}$

  This is a type III operation that replaces by the sum of row 2 and -3 times row 1.

Example 70. Let $A = \begin{bmatrix} 1 & 2 & -3 \\ 0 & 0 & 1 \\ 0 & 2 & -4 \end{bmatrix}$.

Consider the following sequence of elementary row operations:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 2 & 4 \end{bmatrix} \xrightarrow{E_{23}} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{E_2(1/2)} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

At this point, notice that the last matrix is already a matrix in REF which is row equivalent to A. Should we require a matrix in RREF that is row equivalent to A, we may proceed as follows.

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{E_{12}(-2)} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{E_{13}(1)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{E_{23}(-2)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

As we can see, we eventually end up with the 3-by-3 identity matrix which coincidentally is a matrix in RREF. This example shows that each of the matrices

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

are matrices in REF which is row equivalent to A, but of these that is row equivalent to A, only the identity matrix is in RREF.

Remark.
Every m-by-n matrix A is row equivalent to a matrix in REF. Furthermore, each A is row equivalent to a **unique** matrix in RREF.

The unique matrix in RREF which is row equivalent to A shall be denoted by $\text{RREF}(A)$. The uniqueness of $\text{RREF}(A)$ makes it a little more interesting than matrices in REF (but not in REF). If two people talks about the RREF of a matrix, we are certain that they are talking about the same thing.

Technically speaking, there are various ways we can calculate for $\text{RREF}(A)$. Meanwhile, the following steps are suggested.

1. If A has zero rows, apply type I operations to move the zero rows at the bottom of the matrix. In each step, if in case a zero row appears, move them to the bottom part of the matrix.

2. Create a pivot for each nonzero row starting from the first row. One can achieve this by either using a type II operation on the row or a type I operation interchanging it with a row from a row below it.

3. Once a pivot is created on the first row, eliminate all the entries below it.

4. Ignore the first row and create a pivot on the second row by using either a type II or type I operation on rows below it.

5. Once a pivot is created on the second row, use type III operations to eliminate the nonzero entries above and below the pivot.

6. Continue to repeat the process for the next row, and so on, until we obtain $\text{RREF}(A)$ of the form:

$$\text{RREF}(A) = \begin{bmatrix} 0 & \cdots & \mathbf{1} & * & \cdots & 0 & * & \cdots & 0 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & \cdots & \mathbf{1} & * & \cdots & 0 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & * & \cdots & * \\ & \vdots & & & \vdots & & & \vdots & & & \vdots & \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & \mathbf{1} & * & \cdots & * \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ & \vdots & & & \vdots & & & \vdots & & & \vdots & \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Example 71. Find a matrix in RREF that is row equivalent to $\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 3 \\ 0 & 1 & 1 & 1 \end{bmatrix}$.

Solution. We present the solution in the tabular form as shown:

Operation/Remark	Matrix
Given	$\begin{bmatrix} \mathbf{1} & 1 & 1 & 0 \\ 1^* & 1 & 0 & 3 \\ 0 & 1 & 1 & 1 \end{bmatrix}$
Apply $E_{21}(-1)$ to eliminate the $(2,1)$-entry 1 (marked by $*$) using the pivot on the $(1,1)$ position (in boldface).	$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 3 \\ 0 & 1 & 1 & 1 \end{bmatrix}$
Apply E_{23} to create a pivot on the second row.	$\begin{bmatrix} 1 & 1^* & 1 & 0 \\ 0 & \mathbf{1} & 1 & 1 \\ 0 & 0 & -1 & 3 \end{bmatrix}$
Apply $E_{12}(-1)$ to eliminate the $(1,2)$-entry using the pivot in $(2,2)$-position.	$\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & \mathbf{1} & 1 & 1 \\ 0 & 0 & -1 & 3 \end{bmatrix}$
Apply $E_3(-1)$ to create a pivot on the third row.	$\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1^* & 1 \\ 0 & 0 & \mathbf{1} & -3 \end{bmatrix}$
Apply $E_{23}(-1)$ to create a pivot on the third row.	$\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -3 \end{bmatrix}$

We stop the algorithm since the last matrix is already in RREF. Therefore,

$$\operatorname{RREF}\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 3 \\ 0 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -3 \end{bmatrix}.$$

In the previous example, if we interpret the given matrix as an augmented matrix of a linear system:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 3 \\ 0 & 1 & 1 & 1 \end{array}\right] \iff \begin{cases} x+y+z=0 \\ x+y =3 \\ y+z=1 \end{cases}$$

then we have solved the system by looking at the RREF of the given matrix:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -3 \end{array}\right] \iff \begin{cases} x = -1 \\ y = 4 \\ z = -3 \end{cases}$$

Theorem 72: Gauss-Jordan

Let A and B be the augmented matrices of the linear systems $(\heartsuit)$ and $(\diamondsuit)$, respectively. Then A and B are row equivalent if and only if $(\heartsuit)$ and $(\diamondsuit)$ are equivalent.

We now present the Gauss-Jordan method for solving linear systems.

Gauss-Jordan Method

Given a linear system $(\heartsuit)$.

- **(Step 1)** Find the augmented matrix A of $(\heartsuit)$.
- **(Step 2)** Find the matrix B in RREF which is row equivalent to A by applying a sequence of elementary row operations.
- **(Step 3)** Rewrite B into a linear system and use **back-substitution**, if necessary.

Example 73. Solve the following linear systems by Gauss-Jordan method.

1. $\begin{cases} x+2y-z=3 \\ x+3y+z=5 \\ 3x+8y+4z=17 \end{cases}$

2. $\begin{cases} x-2y+4z=2 \\ 2x-3y+5z=3 \\ 3x-4y+6z=7 \end{cases}$

3. $$\begin{cases} x_1 + 2x_2 - 3x_3 - 2x_4 + 4x_5 = 1 \\ 2x_1 + 5x_2 - 8x_3 - x_4 + 6x_5 = 4 \\ x_1 + 4x_2 - 7x_3 + 5x_4 + 2x_5 = 8 \end{cases}$$

Solutions.

1. For this linear system, the augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 1 & 3 & 1 & 5 \\ 3 & 8 & 4 & 17 \end{array}\right].$$

We now reduce this matrix to its RREF using elimination.

$$\begin{array}{ccc|c} \mathbf{1} & 2 & -1 & 3 \\ 1^* & 3 & 1 & 5 \\ 3^* & 8 & 4 & 17 \end{array} \xrightarrow[E_{31}(-3)]{E_{21}(-1)} \begin{array}{ccc|c} 1 & 2^* & -1 & 3 \\ 0 & \mathbf{1} & 2 & 2 \\ 0 & 2^* & 7 & 8 \end{array}$$

$$\xrightarrow[E_{32}(-2)]{E_{12}(-2)} \begin{array}{ccc|c} 1 & 0 & -5 & -1 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 3 & 4 \end{array}$$

$$\xrightarrow{E_3(1/3)} \begin{array}{ccc|c} 1 & 0 & -5^* & -1 \\ 0 & 1 & 2^* & 2 \\ 0 & 0 & \mathbf{1} & 4/3 \end{array}$$

$$\xrightarrow[E_{23}(-2)]{E_{13}(5)} \begin{array}{ccc|c} 1 & 0 & 0 & 17/3 \\ 0 & 1 & 0 & -2/3 \\ 0 & 0 & 1 & 4/3 \end{array}$$

The last step yields a matrix in RREF, so we stop. This implies that $x = 17/3, y =$

$-2/3,$ and $z = 4/3$ yielding a unique solution

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 17/3 \\ -2/3 \\ 4/3 \end{bmatrix} \overset{\text{or}}{=} \frac{1}{3} \begin{bmatrix} 17 \\ -2 \\ 4 \end{bmatrix}.$$

2. For this example, the augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & -2 & 4 & 2 \\ 2 & -3 & 5 & 3 \\ 3 & -4 & 6 & 7 \end{array}\right].$$

Reducing this matrix to RREF using elimination yields

$$\begin{array}{ccc|c} \mathbf{1} & -2 & 4 & 2 \\ 2^* & -3 & 5 & 3 \\ 3^* & -4 & 6 & 7 \end{array} \xrightarrow[E_{31}(-3)]{E_{21}(-2)} \begin{array}{ccc|c} 1 & -2^* & 4 & 2 \\ 0 & \mathbf{1} & -3 & -1 \\ 0 & 2^* & -6 & -1 \end{array}$$

$$\xrightarrow[E_{32}(-2)]{E_{12}(2)} \begin{array}{ccc|c} 1 & 0 & -2 & 0 \\ 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 1 \end{array}$$

from where we stop the algorithm because the last row of this matrix corresponds to the degenerate case $0x + 0y + 0z = 1$, which has no solution. Consequently, the linear system is inconsistent.

3. For this linear system, the augmented matrix is

$$\left[\begin{array}{ccccc|c} 1 & 2 & -3 & -2 & 4 & 1 \\ 2 & 5 & -8 & -1 & 6 & 4 \\ 1 & 4 & -7 & 5 & 2 & 8 \end{array}\right].$$

We now perform Gaussian elimination to attain its RREF.

$$\left[\begin{array}{ccccc|c} \mathbf{1} & 2 & -3 & -2 & 4 & 1 \\ 2^* & 5 & -8 & -1 & 6 & 4 \\ 1^* & 4 & -7 & 5 & 2 & 8 \end{array}\right] \xrightarrow[E_{31}(-1)]{E_{21}(-2)} \left[\begin{array}{ccccc|c} 1 & 2^* & -3 & -2 & 4 & 1 \\ 0 & \mathbf{1} & -2 & 3 & -2 & 2 \\ 0 & 2^* & -4 & 7 & -2 & 7 \end{array}\right]$$

$$\xrightarrow[E_{32}(-2)]{E_{12}(-2)} \left[\begin{array}{ccccc|c} 1 & 0 & 1 & -8 & 8 & -3 \\ 0 & 1 & -2 & 3 & -2 & 2 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{array}\right]$$

continuing,

$$\left[\begin{array}{ccccc|c} 1 & 0 & 1 & -8^* & 8 & -3 \\ 0 & 1 & -2 & 3^* & -2 & 2 \\ 0 & 0 & 0 & \mathbf{1} & 2 & 3 \end{array}\right] \xrightarrow[E_{23}(-3)]{E_{13}(8)} \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & 24 & 21 \\ 0 & 1 & -2 & 0 & -8 & -7 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{array}\right]$$

We stop from here since we have already achieved a matrix in RREF. As a linear system, this translates to

$$\begin{cases} x_1 \quad + \ x_3 \quad + 24x_5 = 21 \\ \quad x_2 - 2x_3 \quad - \ 8x_5 = -7 \\ \qquad\qquad x_4 + \ 2x_5 = 3 \end{cases}$$

Equivalently, we may write this linear system as

$$\begin{cases} x_1 = 21 - x_3 - 24x_5 \\ x_2 = -7 + 2x_3 + 8x_5 \\ x_4 = 3 - 2x_5 \end{cases}$$

As we can see, the values for the unknowns x_1, x_2, and x_4 are uniquely determined if one sets values to the unknowns x_3 and x_5. If we set $x_3 = r$ and $x_5 = s$, where

$r, s \in \mathbb{R}$ are arbitrary, then we may write the *general form* of the solution as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 21 - r - 24s \\ -7 + 2r + 8s \\ r \\ 3 - 2s \\ s \end{bmatrix} = \begin{bmatrix} 21 \\ -7 \\ 0 \\ 3 \\ 0 \end{bmatrix} + r \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -24 \\ 8 \\ 0 \\ -2 \\ 1 \end{bmatrix},$$

where $r, s \in \mathbb{R}$ are arbitrary constants. Consequently, the linear system has infinitely many solutions. Should one require *particular* solutions, one may set specific values for r and s. For example, if we set $r = 1$ and $s = -1$, we get

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 21 \\ -7 \\ 0 \\ 3 \\ 0 \end{bmatrix} + (1) \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + (-1) \begin{bmatrix} -24 \\ 8 \\ 0 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 44 \\ -13 \\ 1 \\ 5 \\ -1 \end{bmatrix}.$$

In the previous example, the variables x_1, x_2, and x_4 are called the **leading variables** for the reason that their values may be expressed to depend on other variables. The non-leading variables x_3 and x_5 are called the **free variables**, since they are 'free' to assume arbitrary values.

Remarks.

1. The leading variables correspond to columns of the RREF matrix that have pivots. On the other hand, the free variables are the ones that correspond to columns without pivots.

2. The existence of a free variable implies the existence of infinitely many solutions in a linear system.

3. The number of pivots in $\text{RREF}(A)$ is also called as the **rank** of A denoted by $\text{rank}(A)$. Consequently,

$$(\text{no. of free variables}) = (\text{no. of columns of } A) - \text{rank}(A).$$

Example 74. Polynomial Interpolation Problem

Find a polynomial function $p(x)$ of degree 3 satisfying the following table of values.

x	-1	1	2	3
$p(x)$	2	0	-4	-8

Solution. We wish to find a polynomial function

$$p(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3$$

such that $p(-1) = 2$, $p(1) = 0$, $p(2) = -4$, and $p(3) = -8$, where c_i is unknown for $i = 0, 1, 2, 3$.

The condition $p(-1) = 2$ implies

$$c_0 + c_1(-1) + c_2(-1)^2 + c_3(-1)^3 = 2 \implies c_0 - c_1 + c_2 - c_3 = 2.$$

The condition $p(1) = 0$ implies

$$c_0 + c_1(1) + c_2(1)^2 + c_3(1)^3 = 0 \implies c_0 + c_1 + c_2 + c_3 = 0.$$

The condition $p(2) = -4$ implies

$$c_0 + c_1(2) + c_2(2)^2 + c_3(2)^3 = 0 \implies c_0 + 2c_1 + 4c_2 + 8c_3 = -4.$$

The condition $p(3) = -8$ implies

$$c_0 + c_1(3) + c_2(3)^2 + c_3(3)^3 = 0 \implies c_0 + 3c_1 + 9c_2 + 27c_3 = -8.$$

This gives us the linear system

$$\begin{cases} c_0 - c_1 + c_2 - c_3 = 2 \\ c_0 + c_1 + c_2 + c_3 = 0 \\ c_0 + 2c_1 + 4c_2 + 8c_3 = -4 \\ c_0 + 3c_1 + 9c_2 + 27c_3 = -8 \end{cases}$$

whose augmented matrix is

$$\left[\begin{array}{cccc|c} 1 & -1 & 1 & -1 & 2 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 4 & 8 & -4 \\ 1 & 3 & 9 & 27 & -8 \end{array}\right].$$

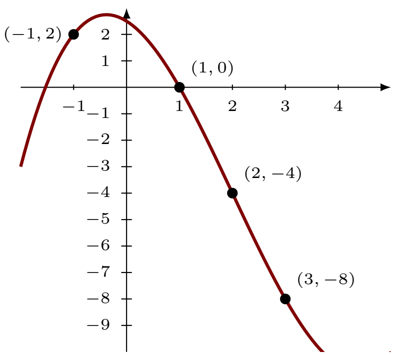
Calculating the RREF of the augmented matrix gives

$$\begin{array}{cccc|c} 1 & -1 & 1 & -1 & 2 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 4 & 8 & -4 \\ 1 & 3 & 9 & 27 & -8 \end{array} \xrightarrow{\text{RREF}} \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 2.5 \\ 0 & 1 & 0 & 0 & -1.25 \\ 0 & 0 & 1 & 0 & -1.5 \\ 0 & 0 & 0 & 1 & 0.25 \end{array}$$

It follows that $c_0 = 2.5$, $c_1 = -1.25$, $c_2 = -1.5$, and $c_3 = 0.25$. Therefore,

$$p(x) = 2.5 - 1.25x - 1.5x^2 + 0.25x^3.$$

The figure below shows the graph of $p(x)$ as it interpolates the given data.



Existence of Solutions and Linear Combinations

Consider the linear system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases} \tag{2.4}$$

Note that the linear system (2.4) may be written alternatively as

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

If let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, then we can conclude that the given linear system is consistent if and only if the column vector

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

is a linear combination of the columns of A.

Theorem 75: Linear Combination Theorem

Let $[A \mid \vec{b}]$ be the augmented matrix of an m-by-n linear system. The linear system is consistent if and only if $\vec{b}$ is a linear combination of the columns of A.

Example 76. Consider the linear system

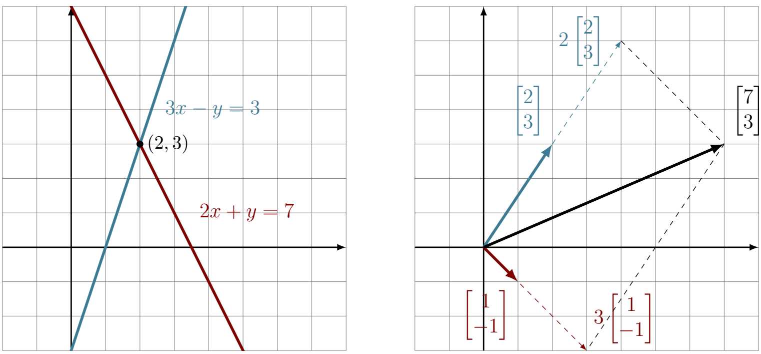
$$\begin{cases} 2x + y = 7 \\ 3x - y = 3 \end{cases}$$

When both equations $2x+y=7$ and $3x-y=3$ are viewed as straight lines in the plane, the solution is the point of intersection of these lines.

On the other hand, one may also interpret the linear system as the linear combination problem:

$$x\begin{bmatrix}2\\3\end{bmatrix} + y\begin{bmatrix}1\\-1\end{bmatrix} = \begin{bmatrix}7\\3\end{bmatrix}.$$

Using the Gauss-Jordan method, one checks that both problems have a unique solution of $x = 2$ and $y = 3$. The figures below show the difference in perspectives of these two approaches to linear systems.



row view column view

The first model is sometimes called as the **row view** of the linear system, while the second model is called the **column view**. As we shall see later, problems in linear algebra are usually phrased in terms of the column view.

Example 77. Prove that every vector in $\mathbb{R}^3$ is a linear combination of the vectors

$$\vec{v_1} = \begin{bmatrix}1\\1\\0\end{bmatrix}, \vec{v_2} = \begin{bmatrix}1\\0\\1\end{bmatrix}, \text{and } \vec{v_3} = \begin{bmatrix}0\\1\\1\end{bmatrix}.$$

Proof. Let $\vec{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$. We need to find scalars c_1, c_2, and c_3 such that

$$\vec{v} = c_1\vec{v_1} + c_2\vec{v_2} + c_3\vec{v_3}. \tag{2.5}$$

Equation (2.5) implies that

$$c_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

This is equivalent to the linear system

$$\begin{cases} c_1 + c_2 \phantom{{}+c_3} = x \\ c_1 \phantom{{}+c_2} + c_3 = y \\ \phantom{c_1 +{}} c_2 + c_3 = z \end{cases}$$

We run the Gauss-Jordan algorithm to solve this linear system.

$$\left.\begin{array}{ccc|c} 1 & 1 & 0 & x \\ 1 & 0 & 1 & y \\ 0 & 1 & 1 & z \end{array}\right. \xrightarrow{E_{21}(-1)} \left.\begin{array}{ccc|c} 1 & 1 & 0 & x \\ 0 & -1 & 1 & y-x \\ 0 & 1 & 1 & z \end{array}\right. \xrightarrow{E_{2}(-1)} \left.\begin{array}{ccc|c} 1 & 1 & 0 & x \\ 0 & 1 & -1 & x-y \\ 0 & 1 & 1 & z \end{array}\right.$$

$$\xrightarrow[E_{32}(-1)]{E_{12}(-1)} \left.\begin{array}{ccc|c} 1 & 0 & 1 & y \\ 0 & 1 & -1 & x-y \\ 0 & 0 & 2 & z-x+y \end{array}\right. \xrightarrow{E_{3}(1/2)} \left.\begin{array}{ccc|c} 1 & 0 & 1 & y \\ 0 & 1 & -1 & x-y \\ 0 & 0 & 1 & \frac{1}{2}(z-x+y) \end{array}\right.$$

$$\xrightarrow[E_{23}(1)]{E_{13}(-1)} \left.\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{2}(x+y-z) \\ 0 & 1 & 0 & \frac{1}{2}(x-y+z) \\ 0 & 0 & 1 & \frac{1}{2}(z-x+y) \end{array}\right.$$

It follows that

$$c_1 = \tfrac{1}{2}(x+y-z),\ c_2 = \tfrac{1}{2}(x-y+z), \text{ and } c_3 = \tfrac{1}{2}(-x+y+z).$$

Consequently, every

$$\vec{v} = \left(\tfrac{x+y-z}{2}\right)\vec{v_1} + \left(\tfrac{x-y+z}{2}\right)\vec{v_2} + \left(\tfrac{-x+y+z}{2}\right)\vec{v_3}$$

is a linear combination of $\vec{v_1}$, $\vec{v_2}$, and $\vec{v_3}$.

Homogeneous Linear Systems

Definition 78

A **homogenous linear system** is a linear system of the form $A\vec{x} = \vec{0}$.

Observe that a homogenous linear system is always consistent as the zero vector $\vec{x} = \vec{0}$ is a solution. This solution is called as the **trivial solution**. This means that for a homogeneous linear system, we ask if it has solutions different from the zero vector.

Example 79. Find the solution set of the homogeneous linear system

$$\begin{cases} x + \ \ y + z + \ \ w = 0 \\ 2x + 3y + z - 2w = 0 \\ 3x + 5y + z \qquad\ \ = 0 \end{cases}$$

Give two non-trivial solutions.

Solution. In matrix form, the linear system may be expressed as

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 1 & -2 \\ 3 & 5 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

We apply the Gauss-Jordan method as follows:

$$\begin{array}{cccc|c} x & y & z & w & \\ \hline 1 & 1 & 1 & 1 & 0 \\ 2 & 3 & 1 & -2 & 0 \\ 3 & 5 & 1 & 0 & 0 \end{array} \xrightarrow{\text{RREF}} \begin{array}{cccc|c} x & y & z & w & \\ \hline 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}.$$

Looking at the RREF matrix, we see that the pivots are located in the columns corresponding to the unknowns x, y, and z, which means these are the leading variables. On the other hand, the column corresponding to z has no pivot, thus, z is a free variable.

- Set $z = r$, where $r \in \mathbb{R}$.
- The first row of the RREF matrix means

$$x + 2z = 0 \implies x = -2z = -2r.$$

- The second row of the RREF matrix means

$$y - z = 0 \implies y = z = r.$$

- The third row of the RREF matrix means

$$w = 0.$$

Therefore, the general solution to the linear system may be expressed as

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -2r \\ r \\ r \\ 0 \end{bmatrix} = r \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \end{bmatrix}, r \in \mathbb{R}$$

which defines the solution set. To give two non-trivial solutions, we set two <u>nonzero</u> values of the parameter r.

- If $r = 1$, then $\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \end{bmatrix}$.
- If $r = -1$, then $\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -1 \\ 0 \end{bmatrix}$.

By looking at the example above, we have reached the following realization.

Theorem 80

A homogeneous linear system with fewer equations than unknowns always has a nontrivial solution.

Proof. Consider the m-by-n linear system $A\vec{x} = \vec{0}$, where $m < n$. Let r denote the number of pivot of RREF(A). Then $r \leq m < n$. This implies that at least one column of RREF(A) receives no pivot, hence, a free variable exists. Consequently, the linear system has infinitely many solutions and that a nontrivial solution exists. ▨

Definition 81: Null Space of a Matrix

The solution set to the homogeneous linear system $A\vec{x} = \vec{0}$ is called the **null space** of A denoted by $N(A)$. In symbols,

$$N(A) = \{\vec{x} \mid A\vec{x} = \vec{0}\}.$$

Example 82. Find the null space of the matrix $A = \begin{bmatrix} 1 & 3 & 2 & -4 \\ 3 & 9 & -5 & -10 \end{bmatrix}$.

Solution. To find $N(A)$, we solve for the homogeneous linear system $A\vec{x} = \vec{0}$ which is equivalent to

$$\begin{cases} x_1 + 3x_2 + 2x_3 - 4x_4 = 0 \\ 3x_1 + 9x_2 - 5x_3 - 10x_4 = 0 \end{cases}.$$

Since

$$\left.\begin{array}{cccc|c} 1 & 3 & 2 & -4 & 0 \\ 3 & 9 & -5 & -10 & 0 \end{array}\right. \xrightarrow{\text{RREF}} \left.\begin{array}{cccc|c} 1 & 3 & 0 & -4 & 0 \\ 0 & 0 & 1 & 2 & 0 \end{array}\right.,$$

it follows that the x_2 and x_4 are free variables and that the general solution may be expressed as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3r + 4s \\ r \\ -2s \\ s \end{bmatrix} = r\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 4 \\ 0 \\ -2 \\ 1 \end{bmatrix},$$

where r and s are arbitrary scalars. Therefore,

$$N(A) = \left\{ r \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 4 \\ 0 \\ -2 \\ 1 \end{bmatrix} : r, s \in \mathbb{R} \right\}.$$

Consider a general linear system $A\vec{x} = \vec{b}$. Suppose that the given system is consistent. Let $\vec{x}_p$ be a particular solution. We ask: what are the other solutions? If $\vec{x}$ is any solution, note that

$$A(\vec{x} - \vec{x}_p) = A\vec{x} - A\vec{x}_p = \vec{b} - \vec{b} = \vec{0}.$$

This means that $\vec{x}_h := \vec{x} - \vec{x}_p \in N(A)$. Therefore, every solution of $A\vec{x} = \vec{b}$ has the form $\vec{x} = \vec{x}_p + \vec{x}_h$, where $\vec{x}_h \in N(A)$.

We have proven the following result.

Theorem 83

Given an m-by-n linear system $A\vec{x} = \vec{b}$ over $\mathbb{R}$. If $A\vec{x} = \vec{b}$ has a solution $\vec{x}_p$, then its solution set is

$$x_p + N(A) := \{x_p + x_h \mid x_h \in N(A)\}.$$

Geometrically speaking, the solution set to $A\vec{x} = \vec{b}$ is a *vector translation* of $N(A)$. Once we know one solution, we all know all other solutions. As a consequence, we have the following result.

Corollary 84

A consistent linear system $A\vec{x} = \vec{b}$ has a unique solution if and only if $N(A) = \{\vec{0}\}$, i.e., the associated homogeneous linear system only has the trivial solution.

