Matrices

Definition of a Matrix

Definition 1.1.1 (Matrix) A rectangular array of numbers is called a matrix.

We shall mostly be concerned with matrices having real numbers as entries.

The horizontal arrays of a matrix are called its ROWS and the vertical arrays are called its COLUMNS. A matrix having m rows and n columns is said to have the order $m \times n$.

A matrix A of ORDER $m \times n$ can be represented in the following form:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

where a_{ij} is the entry at the intersection of the i^{th} row and j^{th} column.

In a more concise manner, we also denote the matrix A by $[a_{ij}]$ by suppressing its order.

Remark 1.1.2 *Some books also use* $\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$ *to represent a matrix.*

Let $A = \begin{bmatrix} 1 & 3 & 7 \\ 4 & 5 & 6 \end{bmatrix}$. Then $a_{11} = 1,\ a_{12} = 3,\ a_{13} = 7,\ a_{21} = 4,\ a_{22} = 5,$ and $a_{23} = 6.$

A matrix having only one column is called a COLUMN VECTOR; and a matrix with only one row is called a ROW VECTOR.

WHENEVER A VECTOR IS USED, IT SHOULD BE UNDERSTOOD FROM THE CONTEXT WHETHER IT IS A ROW VECTOR OR A COLUMN VECTOR.

Definition 1.1.3 (Equality of two Matrices) Two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ having the same order $m \times n$ are equal if $a_{ij} = b_{ij}$ for each $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

In other words, two matrices are said to be equal if they have the same order and their corresponding entries are equal.

Example 1.1.4 The linear system of equations $2x + 3y = 5$ and $3x + 2y = 5$ can be identified with the matrix $\begin{bmatrix} 2 & 3 & : & 5 \\ 3 & 2 & : & 5 \end{bmatrix}$.

1.1.1 Special Matrices

Definition 1.1.5

1. A matrix in which each entry is zero is called a zero-matrix, denoted by $\mathbf{0}$. For example,

$$\mathbf{0}_{2\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \mathbf{0}_{2\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

2. A matrix having the number of rows equal to the number of columns is called a square matrix. Thus, its order is $m \times m$ (for some m) and is represented by m only.

3. In a square matrix, $A = [a_{ij}]$, of order n, the entries $a_{11}, a_{22}, \ldots, a_{nn}$ are called the diagonal entries and form the principal diagonal of A.

4. A square matrix $A = [a_{ij}]$ is said to be a diagonal matrix if $a_{ij} = 0$ for $i \neq j$. In other words, the non-zero entries appear only on the principal diagonal. For example, the zero matrix $\mathbf{0}_n$ and $\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$ are a few diagonal matrices.

   A diagonal matrix D of order n with the diagonal entries $d_1, d_2, \ldots, d_n$ is denoted by $D = \text{diag}(d_1, \ldots, d_n)$.

   If $d_i = d$ for all $i = 1, 2, \ldots, n$ then the diagonal matrix D is called a **scalar matrix**.

5. A square matrix $A = [a_{ij}]$ with $a_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$ is called the identity matrix, denoted by I_n.

   For example, $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, and $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

   The subscript n is suppressed in case the order is clear from the context or if no confusion arises.

6. A square matrix $A = [a_{ij}]$ is said to be an upper triangular matrix if $a_{ij} = 0$ for $i > j$.

   A square matrix $A = [a_{ij}]$ is said to be an lower triangular matrix if $a_{ij} = 0$ for $i < j$.

   A square matrix A is said to be triangular if it is an upper or a lower triangular matrix.

   For example $\begin{bmatrix} 2 & 1 & 4 \\ 0 & 3 & -1 \\ 0 & 0 & -2 \end{bmatrix}$ is an upper triangular matrix. An upper triangular matrix will be represented by $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$.

1.2 Operations on Matrices

Definition 1.2.1 (Transpose of a Matrix) The transpose of an $m \times n$ matrix $A = [a_{ij}]$ is defined as the $n \times m$ matrix $B = [b_{ij}]$, with $b_{ij} = a_{ji}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$. The transpose of A is denoted by A^t.

That is, by the transpose of an $m \times n$ matrix A, we mean a matrix of order $n \times m$ having the rows of A as its columns and the columns of A as its rows.

For example, if $A = \begin{bmatrix} 1 & 4 & 5 \\ 0 & 1 & 2 \end{bmatrix}$ then $A^t = \begin{bmatrix} 1 & 0 \\ 4 & 1 \\ 5 & 2 \end{bmatrix}$.

Thus, the transpose of a row vector is a column vector and vice-versa.

Theorem 1.2.2 For any matrix A, we have $(A^t)^t = A$.

PROOF. Let $A = [a_{ij}]$, $A^t = [b_{ij}]$ and $(A^t)^t = [c_{ij}]$. Then, the definition of transpose gives

$$c_{ij} = b_{ji} = a_{ij} \quad \text{for all} \quad i, j$$

and the result follows. $\square$

Definition 1.2.3 (Addition of Matrices) let $A = [a_{ij}]$ and $B = [b_{ij}]$ be are two $m \times n$ matrices. Then the sum $A + B$ is defined to be the matrix $C = [c_{ij}]$ with $c_{ij} = a_{ij} + b_{ij}$.

Note that, we define the sum of two matrices only when the order of the two matrices are same.

Definition 1.2.4 (Multiplying a Scalar to a Matrix) Let $A = [a_{ij}]$ be an $m \times n$ matrix. Then for any element $k \in \mathbb{R}$, we define $kA = [ka_{ij}]$.

For example, if $A = \begin{bmatrix} 1 & 4 & 5 \\ 0 & 1 & 2 \end{bmatrix}$ and $k = 5$, then $5A = \begin{bmatrix} 5 & 20 & 25 \\ 0 & 5 & 10 \end{bmatrix}$.

Theorem 1.2.5 Let A, B and C be matrices of order $m \times n$, and let $k, \ell \in \mathbb{R}$. Then

1. $A + B = B + A$ (commutativity).
2. $(A + B) + C = A + (B + C)$ (associativity).
3. $k(\ell A) = (k\ell)A$.
4. $(k + \ell)A = kA + \ell A$.

PROOF. Part 1.
Let $A = [a_{ij}]$ and $B = [b_{ij}]$. Then

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] = [b_{ij} + a_{ij}] = [b_{ij}] + [a_{ij}] = B + A$$

as real numbers commute.

The reader is required to prove the other parts as all the results follow from the properties of real numbers. $\square$

Exercise 1.2.6

1. Suppose $A + B = A$. Then show that $B = \mathbf{0}$.
2. Suppose $A + B = \mathbf{0}$. Then show that $B = (-1)A = [-a_{ij}]$.

Definition 1.2.7 (Additive Inverse) Let A be an $m \times n$ matrix.

1. Then there exists a matrix B with $A + B = \mathbf{0}$. This matrix B is called the additive inverse of A, and is denoted by $-A = (-1)A$.
2. Also, for the matrix $\mathbf{0}_{m \times n}$, $A + \mathbf{0} = \mathbf{0} + A = A$. Hence, the matrix $\mathbf{0}_{m \times n}$ is called the additive identity.

Multiplication of Matrices

Definition 1.2.8 (Matrix Multiplication / Product) Let $A = [a_{ij}]$ be an $m \times n$ matrix and $B = [b_{ij}]$ be an $n \times r$ matrix. The product AB is a matrix $C = [c_{ij}]$ of order $m \times r$, with

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj}.$$

Observe that the product AB is defined if and only if
THE NUMBER OF COLUMNS OF $A =$ THE NUMBER OF ROWS OF B.

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

$$AB = \begin{bmatrix} 1+0+3 & 2+0+0 & 1+6+12 \\ 2+0+1 & 4+0+0 & 2+12+4 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 19 \\ 3 & 4 & 18 \end{bmatrix}.$$

Note that in this example, while AB is defined, the product BA is not defined. However, for square matrices A and B of the same order, both the product AB and BA are defined.

Definition 1.2.9 Two square matrices A and B are said to commute if $AB = BA$.

Remark 1.2.10

1. *Note that if A is a square matrix of order n then $AI_n = I_n A$. Also for any $d \in \mathbb{R}$, the matrix dI_n commutes with every square matrix of order n. The matrices dI_n for any $d \in \mathbb{R}$ are called* SCALAR *matrices.*

2. *In general, the matrix product is not commutative. For example, consider the following two matrices $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$. Then check that the matrix product*

$$AB = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = BA.$$

Theorem 1.2.11 Suppose that the matrices A, B and C are so chosen that the matrix multiplications are defined.

1. Then $(AB)C = A(BC)$. That is, the matrix multiplication is associative.

2. For any $k \in \mathbb{R}$, $(kA)B = k(AB) = A(kB)$.

3. Then $A(B + C) = AB + AC$. That is, multiplication distributes over addition.

4. If A is an $n \times n$ matrix then $AI_n = I_n A = A$.

5. For any square matrix A of order n and $D = \text{diag}(d_1, d_2, \ldots, d_n)$, we have
   - the first row of DA is d_1 times the first row of A;
   - for $1 \leq i \leq n$, the i^{th} row of DA is d_i times the i^{th} row of A.

   A similar statement holds for the columns of A when A is multiplied on the right by D.

PROOF. Part 1. Let $A = [a_{ij}]_{m \times n}$, $B = [b_{ij}]_{n \times p}$ and $C = [c_{ij}]_{p \times q}$. Then

$$(BC)_{kj} = \sum_{\ell=1}^{p} b_{k\ell} c_{\ell j} \quad \text{and} \quad (AB)_{i\ell} = \sum_{k=1}^{n} a_{ik} b_{k\ell}.$$

Therefore,

$$\begin{aligned}
\big(A(BC)\big)_{ij} &= \sum_{k=1}^{n} a_{ik}\big(BC\big)_{kj} = \sum_{k=1}^{n} a_{ik}\big(\sum_{\ell=1}^{p} b_{k\ell}c_{\ell j}\big) \\
&= \sum_{k=1}^{n}\sum_{\ell=1}^{p} a_{ik}\big(b_{k\ell}c_{\ell j}\big) = \sum_{k=1}^{n}\sum_{\ell=1}^{p}\big(a_{ik}b_{k\ell}\big)c_{\ell j} \\
&= \sum_{\ell=1}^{p}\big(\sum_{k=1}^{n} a_{ik}b_{k\ell}\big)c_{\ell j} = \sum_{\ell=1}^{t}\big(AB\big)_{i\ell}c_{\ell j} \\
&= \big((AB)C\big)_{ij}.
\end{aligned}$$

Part 5. For all $j = 1, 2, \ldots, n$, we have

$$(DA)_{ij} = \sum_{k=1}^{n} d_{ik}a_{kj} = d_i a_{ij}$$

as $d_{ik} = 0$ whenever $i \neq k$. Hence, the required result follows.

The reader is required to prove the other parts. $\square$

Exercise 1.2.12 1. Let A and B be two matrices. If the matrix addition $A+B$ is defined, then prove that $(A+B)^t = A^t + B^t$. Also, if the matrix product AB is defined then prove that $(AB)^t = B^t A^t$.

2. Let $A = [a_1, a_2, \ldots, a_n]$ and $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$. Compute the matrix products AB and BA.

3. Let n be a positive integer. Compute A^n for the following matrices:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Can you guess a formula for A^n and prove it by induction?

4. Find examples for the following statements.

   (a) Suppose that the matrix product AB is defined. Then the product BA need not be defined.

   (b) Suppose that the matrix products AB and BA are defined. Then the matrices AB and BA can have different orders.

   (c) Suppose that the matrices A and B are square matrices of order n. Then AB and BA may or may not be equal.

1.3 Some More Special Matrices

Definition 1.3.1 1. A matrix A over $\mathbb{R}$ is called symmetric if $A^t = A$ and skew-symmetric if $A^t = -A$.

2. A matrix A is said to be orthogonal if $AA^t = A^tA = I$.

Example 1.3.2 1. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & -1 \\ 3 & -1 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ -2 & 3 & 0 \end{bmatrix}$. Then A is a symmetric matrix and B is a skew-symmetric matrix.

2. Let $A = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{bmatrix}$. Then A is an orthogonal matrix.

3. Let $A = [a_{ij}]$ be an $n \times n$ matrix with $a_{ij} = \begin{cases} 1 & \text{if } i = j+1 \\ 0 & \text{otherwise} \end{cases}$. Then $A^n = \mathbf{0}$ and $A^\ell \neq \mathbf{0}$ for $1 \leq \ell \leq n-1$. The matrices A for which a positive integer k exists such that $A^k = \mathbf{0}$ are called NILPOTENT matrices. The least positive integer k for which $A^k = \mathbf{0}$ is called the ORDER OF NILPOTENCY.

4. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Then $A^2 = A$. The matrices that satisfy the condition that $A^2 = A$ are called IDEMPOTENT matrices.

Exercise 1.3.3 1. Show that for any square matrix A, $S = \frac{1}{2}(A + A^t)$ is symmetric, $T = \frac{1}{2}(A - A^t)$ is skew-symmetric, and $A = S + T$.

2. Show that the product of two lower triangular matrices is a lower triangular matrix. A similar statement holds for upper triangular matrices.

3. Let A and B be symmetric matrices. Show that AB is symmetric if and only if $AB = BA$.

4. Show that the diagonal entries of a skew-symmetric matrix are zero.

5. Let A, B be skew-symmetric matrices with $AB = BA$. Is the matrix AB symmetric or skew-symmetric?

6. Let A be a symmetric matrix of order n with $A^2 = \mathbf{0}$. Is it necessarily true that $A = \mathbf{0}$?

7. Let A be a nilpotent matrix. Show that there exists a matrix B such that $B(I + A) = I = (I + A)B$.

1.3.1 Submatrix of a Matrix

Definition 1.3.4 A matrix obtained by deleting some of the rows and/or columns of a matrix is said to be a submatrix of the given matrix.

For example, if $A = \begin{bmatrix} 1 & 4 & 5 \\ 0 & 1 & 2 \end{bmatrix}$, a few submatrices of A are

$$[1], [2], \begin{bmatrix} 1 \\ 0 \end{bmatrix}, [1\ 5], \begin{bmatrix} 1 & 5 \\ 0 & 2 \end{bmatrix}, A.$$

But the matrices $\begin{bmatrix} 1 & 4 \\ 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 4 \\ 0 & 2 \end{bmatrix}$ are not submatrices of A.

Linear System of Equations

2.1 Introduction

Let us look at some examples of linear systems.

1. Suppose $a, b \in \mathbb{R}$. Consider the system $ax = b$.
   (a) If $a \neq 0$ then the system has a UNIQUE SOLUTION $x = \frac{b}{a}$.
   (b) If $a = 0$ and
      i. $b \neq 0$ then the system has NO SOLUTION.
      ii. $b = 0$ then the system has INFINITE NUMBER OF SOLUTIONS, namely all $x \in \mathbb{R}$.

2. We now consider a system with 2 equations in 2 unknowns.
   Consider the equation $ax + by = c$. If one of the coefficients, a or b is non-zero, then this linear equation represents 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. There are three cases to be considered. Each case is illustrated by an example.

   (a) UNIQUE SOLUTION
   $x + 2y = 1$ and $x + 3y = 1$. The unique solution is $(x, y)^t = (1, 0)^t$.
   Observe that in this case, $a_1b_2 - a_2b_1 \neq 0$.

   (b) INFINITE NUMBER OF SOLUTIONS
   $x + 2y = 1$ and $2x + 4y = 2$. The set of solutions is $(x, y)^t = (1 - 2y, y)^t = (1, 0)^t + y(-2, 1)^t$ 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 $b_1c_2 - b_2c_1 = 0$.

   (c) NO SOLUTION
   $x + 2y = 1$ 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 $a_1c_2 - a_2c_1 \neq 0$.

3. As a last example, consider 3 equations in 3 unknowns.
   A linear equation $ax + by + cz = d$ represent a plane in $\mathbb{R}^3$ provided $(a, b, c) \neq (0, 0, 0)$. As in the case of 2 equations in 2 unknowns, we have to look at the points of intersection of the given three planes. Here again, we have three cases. The three cases are illustrated by examples.

(a) UNIQUE SOLUTION

Consider the system $x+y+z=3, \;\; x+4y+2z=7$ and $4x+10y-z=13$. The unique solution to this system is $(x,y,z)^t = (1,1,1)^t$; *i.e.* THE THREE PLANES INTERSECT AT A POINT.

(b) INFINITE NUMBER OF SOLUTIONS

Consider the system $x+y+z=3, \;\; x+2y+2z=5$ and $3x+4y+4z=11$. The set of solutions to this system is $(x,y,z)^t = (1, 2-z, z)^t = (1,2,0)^t + z(0,-1,1)^t$, with z arbitrary: THE THREE PLANES INTERSECT ON A LINE.

(c) NO SOLUTION

The system $x+y+z=3, \;\; x+2y+2z=5$ and $3x+4y+4z=13$ has no solution. In this case, we get three parallel lines as intersections of the above planes taken two at a time.

The readers are advised to supply the proof.

2.2 Definition and a Solution Method

Definition 2.2.1 (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

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{2.2.1}$$

where for $1 \leq i \leq n$, and $1 \leq j \leq m$; $a_{ij}, b_i \in \mathbb{R}$. Linear System (2.2.1) is called HOMOGENEOUS if $b_1 = 0 = b_2 = \cdots = b_m$ and NON-HOMOGENEOUS otherwise.

We rewrite the above equations 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 $[A \;\; \mathbf{b}]$, is the AUGMENTED matrix of the linear system (2.2.1).

Remark 2.2.2 *Observe that the i^{th} row of the augmented matrix $[A \;\; \mathbf{b}]$ 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..*

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 2.2.3 (Solution of a Linear System) A solution of the linear system $A\mathbf{x} = \mathbf{b}$ is a column vector $\mathbf{y}$ with entries $y_1, y_2, \ldots, y_n$ such that the linear system (2.2.1) is satisfied by substituting y_i in place of x_i.

That is, if $\mathbf{y}^t = [y_1, y_2, \ldots, y_n]$ then $A\mathbf{y} = \mathbf{b}$ holds.

Note: The zero n-tuple $\mathbf{x} = \mathbf{0}$ 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.

2.2.1 A Solution Method

Example 2.2.4 Let us solve the linear system $x + 7y + 3z = 11$, $x + y + z = 3$, and $4x + 10y - z = 13$.
Solution:

1. The above linear system and the linear system

$$\begin{aligned} x + y + z &= 3 && \text{Interchange the first two equations.} \\ x + 7y + 3z &= 11 \\ 4x + 10y - z &= 13 \end{aligned} \tag{2.2.2}$$

   have the same set of solutions. (why?)

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

$$\begin{aligned} x + y + z &= 3 \\ 6y + 2z &= 8 && \text{(obtained by subtracting the first} \\ & && \text{equation from the second equation.)} \\ 6y - 5z &= 1 && \text{(obtained by subtracting 4 times the first equation} \\ & && \text{from the third equation.)} \end{aligned} \tag{2.2.3}$$

   This system and the system (2.2.2) has the same set of solution. (why?)

3. Eliminating y from the last two equations of system (2.2.3), we get the system

$$\begin{aligned} x + y + z &= 3 \\ 6y + 2z &= 8 \\ 7z &= 7 && \text{obtained by subtracting the third equation} \\ & && \text{from the second equation.} \end{aligned} \tag{2.2.4}$$

   which has the same set of solution as the system (2.2.3). (why?)

4. The system (2.2.4) and system

$$\begin{aligned} x + y + z &= 3 \\ 3y + z &= 4 && \text{divide the second equation by 2} \\ z &= 1 && \text{divide the second equation by 2} \end{aligned} \tag{2.2.5}$$

   has the same set of solution. (why?)

5. Now, $z = 1$ implies $y = \dfrac{4-1}{3} = 1$ and $x = 3 - (1+1) = 1$. Or in terms of a vector, the set of solution is $\{\ (x, y, z)^t\ : (x, y, z) = (1, 1, 1)\}$.

2.3 Row Operations and Equivalent Systems

Definition 2.3.1 (Elementary Operations) The following operations 1, 2 and 3 are called elementary operations.

1. interchange of two equations, say "interchange the i^{th} and j^{th} equations";

   (compare the system (2.2.2) with the original system.)

2. multiply a non-zero constant throughout an equation, say "multiply the k^{th} equation by $c \neq 0$";

   (compare the system (2.2.5) and the system (2.2.4).)

3. replace an equation by itself plus a constant multiple of another equation, say "replace the k^{th} equation by k^{th} equation plus c times the j^{th} equation".

   (compare the system (2.2.3) with (2.2.2) or the system (2.2.4) with (2.2.3).)

Observations:

1. In the above example, observe that the elementary operations helped us in getting a linear system (2.2.5), which was easily solvable.

2. Note that at Step 1, if we interchange the first and the second equation, we get back to the linear system from which we had started. This means the operation at Step 1, has an inverse operation. In other words, INVERSE OPERATION sends us back to the step where we had precisely started. It will be a useful exercise for the reader to IDENTIFY THE INVERSE OPERATIONS at each step in Example 2.2.4.

So, in Example 2.2.4, the application of a finite number of elementary operations helped us to obtain a simpler system whose solution can be obtained directly. That is, after applying a finite number of elementary operations, a simpler linear system is obtained which can be easily solved. Note that the three elementary operations defined above, have corresponding INVERSE operations, namely,

1. "interchange the i^{th} and j^{th} equations",

2. "divide the k^{th} equation by $c \neq 0$";

3. "replace the k^{th} equation by k^{th} equation minus c times the j^{th} equation".

It will be a useful exercise for the reader to IDENTIFY THE INVERSE OPERATIONS at each step in Example 2.2.4.

Definition 2.3.2 (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.

The linear systems at each step in Example 2.2.4 are equivalent to each other and also to the original linear system.

Lemma 2.3.3 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 elementary operation "the k^{th} equation is replaced by k^{th} equation plus c times the j^{th} equation." The reader is advised to prove the result for other elementary operations.

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,

$$(a_{k1} + ca_{j1})\alpha_1 + (a_{k2} + ca_{j2})\alpha_2 + \cdots + (a_{kn} + ca_{jn})\alpha_n = b_k + cb_j. \tag{2.3.1}$$

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

$$(a_{k1} + ca_{j1})x_1 + (a_{k2} + ca_{j2})x_2 + \cdots + (a_{kn} + ca_{jn})x_n = b_k + cb_j. \tag{2.3.2}$$

Therefore, using Equation (2.3.1), $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is also a solution for the k^{th} Equation (2.3.2).

Use a similar argument to show that 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, we have the proof in this case. □

Lemma 2.3.3 is now used as an induction step to prove the main result of this section (Theorem 2.3.4).

Theorem 2.3.4 Two equivalent systems have the same set of solutions.

PROOF. Let n be the number of elementary operations performed on $A\mathbf{x} = \mathbf{b}$ to get $C\mathbf{x} = \mathbf{d}$. We prove the theorem by induction on n.

If $n = 1$, Lemma 2.3.3 answers the question. If $n > 1$, assume that the theorem is true for $n = m$. Now, suppose $n = m+1$. Apply the Lemma 2.3.3 again at the "last step" (that is, at the $(m+1)^{\text{th}}$ step from the m^{th} step) to get the required result using induction. □

Let us formalise the above section which led to Theorem 2.3.4. For solving a linear system of equations, we applied elementary operations to equations. It is observed that in performing the elementary operations, the calculations were made on the COEFFICIENTS (numbers). The variables $x_1, x_2, \ldots, x_n$ and the sign of equality (that is, " $=$ ") are not disturbed. 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 $[A \;\; \mathbf{b}]$.

Definition 2.3.5 (Elementary Row Operations) The elementary row operations are defined as:

1. interchange of two rows, say "interchange the i^{th} and j^{th} rows", denoted R_{ij};
2. multiply a non-zero constant throughout a row, say "multiply the k^{th} row by $c \neq 0$", denoted $R_k(c)$;
3. replace a row by itself plus a constant multiple of another row, say "replace the k^{th} row by k^{th} row plus c times the j^{th} row", denoted $R_{kj}(c)$.

Exercise 2.3.6 Find the INVERSE row operations corresponding to the elementary row operations that have been defined just above.

Definition 2.3.7 (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 2.3.8 The three matrices given below are row equivalent.

$$\begin{bmatrix} 0 & 1 & 1 & 2 \\ 2 & 0 & 3 & 5 \\ 1 & 1 & 1 & 3 \end{bmatrix} \overrightarrow{R_{12}} \begin{bmatrix} 2 & 0 & 3 & 5 \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{bmatrix} \overrightarrow{R_1(1/2)} \begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{bmatrix}.$$

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

2.3.1 Gauss Elimination Method

Definition 2.3.9 (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 bringing the augmented matrix

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

to an upper triangular form

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

This elimination process is also called the forward elimination method.

The following examples illustrate the Gauss elimination procedure.

Example 2.3.10 Solve the linear system by Gauss elimination method.

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

Solution: In this case, the augmented matrix is $\begin{bmatrix} 0 & 1 & 1 & 2 \\ 2 & 0 & 3 & 5 \\ 1 & 1 & 1 & 3 \end{bmatrix}$. The method proceeds along the following steps.

1. Interchange 1^{st} and 2^{nd} equation (or R_{12}).

$$\begin{aligned} 2x + 3z &= 5 \\ y + z &= 2 \\ x + y + z &= 3 \end{aligned} \qquad \begin{bmatrix} 2 & 0 & 3 & 5 \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{bmatrix}.$$

2. Divide the 1^{st} equation by 2 (or $R_1(1/2)$).

$$\begin{aligned} x + \tfrac{3}{2}z &= \tfrac{5}{2} \\ y + z &= 2 \\ x + y + z &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{bmatrix}.$$

3. Add -1 times the 1^{st} equation to the 3^{rd} equation (or $R_{31}(-1)$).

$$\begin{aligned} x + \tfrac{3}{2}z &= \tfrac{5}{2} \\ y + z &= 2 \\ y - \tfrac{1}{2}z &= \tfrac{1}{2} \end{aligned} \qquad \begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & 1 & 2 \\ 0 & 1 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

4. Add -1 times the 2^{nd} equation to the 3^{rd} equation (or $R_{32}(-1)$).

$$\begin{aligned} x + \tfrac{3}{2}z &= \tfrac{5}{2} \\ y + z &= 2 \\ -\tfrac{3}{2}z &= -\tfrac{3}{2} \end{aligned} \qquad \begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & 1 & 2 \\ 0 & 0 & -\frac{3}{2} & -\frac{3}{2} \end{bmatrix}.$$

5. Multiply the 3$^{\text{rd}}$ equation by $\frac{-2}{3}$ (or $R_3(-\frac{2}{3})$).

$$\begin{aligned} x + \tfrac{3}{2}z &= \tfrac{5}{2} \\ y + z &= 2 \\ z &= 1 \end{aligned} \qquad \begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

The last equation gives $z = 1$, the second equation now gives $y = 1$. Finally the first equation gives $x = 1$. Hence the set of solutions is $(x, y, z)^t = (1, 1, 1)^t$, A UNIQUE SOLUTION.

Example 2.3.11 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.

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

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

$$\begin{aligned} x + y + z &= 3 \\ y + z &= 2 \\ y + z &= 2 \end{aligned} \qquad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 2 \end{bmatrix}.$$

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

$$\begin{aligned} x + y + z &= 3 \\ y + z &= 2 \end{aligned} \qquad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Thus, the set of solutions is $(x, y, z)^t = (1, 2 - z, z)^t = (1, 2, 0)^t + z(0, -1, 1)^t$, with z arbitrary. In other words, the system has INFINITE NUMBER OF SOLUTIONS.

Example 2.3.12 Solve the linear system by Gauss elimination method.

$$\begin{aligned} x + y + z &= 3 \\ x + 2y + 2z &= 5 \\ 3x + 4y + 4z &= 12 \end{aligned}$$

Solution: In this case, the augmented matrix is $\begin{bmatrix} 1 & 1 & 1 & 3 \\ 1 & 2 & 2 & 5 \\ 3 & 4 & 4 & 12 \end{bmatrix}$ and the method proceeds as follows:

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

$$\begin{aligned} x + y + z &= 3 \\ y + z &= 2 \\ 3x + 4y + 4z &= 12 \end{aligned} \qquad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 3 & 4 & 4 & 12 \end{bmatrix}.$$

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

$$\begin{aligned} x+y+z &= 3 \\ y+z &= 2 \\ y+z &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 3 \end{bmatrix}.$$

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

$$\begin{aligned} x+y+z &= 3 \\ y+z &= 2 \\ 0 &= 1 \end{aligned} \qquad \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The third equation in the last step is

$$0x + 0y + 0z = 1.$$

This can never hold for any value of x, y, z. Hence, the system has NO SOLUTION.

Remark 2.3.13 *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 $[A \;\; \mathbf{b}]$.*

2.4 Row Reduced Echelon Form of a Matrix

Definition 2.4.1 (Row Reduced Form of a Matrix) A matrix C is said to be in the row reduced form if

1. THE FIRST NON-ZERO ENTRY IN EACH ROW OF C IS 1;
2. THE COLUMN CONTAINING THIS 1 HAS ALL ITS OTHER ENTRIES ZERO.

A matrix in the row reduced form is also called a ROW REDUCED MATRIX.

Example 2.4.2
1. One of the most important examples of a row reduced matrix is the $n \times n$ identity matrix, I_n. Recall that the $(i,j)^{\text{th}}$ entry of the identity matrix is

$$I_{ij} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}.$$

δ_{ij} is usually referred to as the Kronecker delta function.

2. The matrices $\begin{bmatrix} 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ are also in row reduced form.

3. The matrix $\begin{bmatrix} 1 & 0 & 0 & 0 & 5 \\ 0 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ is not in the row reduced form. (why?)

Definition 2.4.3 (Leading Term, Leading Column) For a row-reduced matrix, the first non-zero entry of any row is called a LEADING TERM. The columns containing the leading terms are called the LEADING COLUMNS.

Definition 2.4.4 (Basic, Free Variables) Consider the linear system $A\mathbf{x} = \mathbf{b}$ in n variables and m equations. Let $[C \ \ \mathbf{d}]$ be the row-reduced matrix obtained by applying the Gauss elimination method to the augmented matrix $[A \ \ \mathbf{b}]$. Then the variables corresponding to the leading columns in the first n columns of $[C \ \ \mathbf{d}]$ 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 2.3.11, the solution set was given by

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

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

Remark 2.4.5 *It is very important to observe that if there are r non-zero rows in the row-reduced form of the matrix then there will be r leading terms. That is, there will be r leading columns. Therefore,* IF THERE ARE r LEADING TERMS AND n VARIABLES, THEN THERE WILL BE r BASIC VARIABLES AND $n - r$ FREE VARIABLES.

2.4.1 Gauss-Jordan Elimination

We now start with Step 5 of Example 2.3.10 and apply the elementary operations once again. But this time, we start with the 3^{rd} row.

I. Add -1 times the third equation to the second equation (or $R_{23}(-1)$).

$$\begin{aligned} x + \tfrac{3}{2}z &= \tfrac{5}{2} \\ y &= 2 \\ z &= 1 \end{aligned} \qquad \begin{bmatrix} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

II. Add $\frac{-3}{2}$ times the third equation to the first equation (or $R_{13}(-\frac{3}{2})$).

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

III. From the above matrix, we directly have the set of solution as $(x, y, z)^t = (1, 1, 1)^t$.

Definition 2.4.6 (Row Reduced Echelon Form of a Matrix) A matrix C is said to be in the row reduced echelon form if

1. C is already in the row reduced form;
2. The rows consisting of all zeros comes below all non-zero rows; and
3. the leading terms appear from left to right in successive rows. That is, for $1 \le \ell \le k$, let i_ℓ be the leading column of the ℓ^{th} row. Then $i_1 < i_2 < \cdots < i_k$.

Example 2.4.7 Suppose $A = \begin{bmatrix} 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ are in row reduced form. Then the corresponding matrices in the row reduced echelon form are respectively, $\begin{bmatrix} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$.

Definition 2.4.8 (Row Reduced Echelon Matrix) A matrix which is in the row reduced echelon form is also called a row reduced echelon matrix.

Definition 2.4.9 (Back Substitution/Gauss-Jordan Method) The procedure to get to Step II of Example 2.3.10 from Step 5 of Example 2.3.10 is called the back substitution.

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.

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 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, consider the right $m \times (n-1)$ submatrix of the matrix obtained in step 1 and proceed as in step 1.

Else, forget the first row and first column. 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.

Remark 2.4.10 *Note that the row reduction involves only row operations and proceeds from* LEFT TO RIGHT. *Hence, if A is a matrix consisting of first s columns of a matrix C, then the row reduced form of A will be the first s columns of the row reduced form of C.*

The proof of the following theorem is beyond the scope of this book and is omitted.

Theorem 2.4.11 The row reduced echelon form of a matrix is unique.

Exercise 2.4.12 1. Solve the following linear system.

(a) $x+y+z+w=0,\ x-y+z+w=0$ and $-x+y+3z+3w=0$.

(b) $x+2y+3z=1$ and $x+3y+2z=1$.

(c) $x+y+z=3,\ x+y-z=1$ and $x+y+7z=6$.

(d) $x+y+z=3,\ x+y-z=1$ and $x+y+4z=6$.

(e) $x+y+z=3,\ x+y-z=1,\ x+y+4z=6$ and $x+y-4z=-1$.

2. Find the row-reduced echelon form of the following matrices.

$$1.\ \begin{bmatrix} -1 & 1 & 3 & 5 \\ 1 & 3 & 5 & 7 \\ 9 & 11 & 13 & 15 \\ -3 & -1 & 13 & \end{bmatrix}, \qquad 2.\ \begin{bmatrix} 10 & 8 & 6 & 4 \\ 2 & 0 & -2 & -4 \\ -6 & -8 & -10 & -12 \\ -2 & -4 & -6 & -8 \end{bmatrix}$$

2.4.2 Elementary Matrices

Definition 2.4.13 A square matrix E of order n is called an **elementary matrix** if it is obtained by applying exactly one elementary row operation to the identity matrix, I_n.

Remark 2.4.14 *There are three types of elementary matrices.*

1. *E_{ij}, which is obtained by the application of the elementary row operation R_{ij} to the identity matrix, I_n. Thus, the $(k,\ell)^{th}$ entry of E_{ij} is* $(E_{ij})_{(k,\ell)} = \begin{cases} 1 & \text{if } k=\ell \text{ and } \ell \neq i,j \\ 1 & \text{if } (k,\ell)=(i,j) \text{ or } (k,\ell)=(j,i) \\ 0 & \text{otherwise} \end{cases}$.

2. *$E_k(c)$, which is obtained by the application of the elementary row operation $R_k(c)$ to the identity matrix, I_n. The $(i,j)^{th}$ entry of $E_k(c)$ is* $(E_k(c))_{(i,j)} = \begin{cases} 1 & \text{if } i=j \text{ and } i \neq k \\ c & \text{if } i=j=k \\ 0 & \text{otherwise} \end{cases}$.

3. *$E_{ij}(c)$, which is obtained by the application of the elementary row operation $R_{ij}(c)$ to the identity matrix, I_n. The $(k,\ell)^{th}$ entry of $E_{ij}(c)$ is* $(E_{ij})_{(k,\ell)} \begin{cases} 1 & \text{if } k=\ell \\ c & \text{if } (k,\ell)=(i,j) \\ 0 & \text{otherwise} \end{cases}$.

In particular,

$$E_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},\ E_1(c) = \begin{bmatrix} c & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ \text{and } E_{23}(c) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}.$$

Example 2.4.15 1. Let $A = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 0 & 3 & 4 \\ 3 & 4 & 5 & 6 \end{bmatrix}$. Then

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 0 & 3 & 4 \\ 3 & 4 & 5 & 6 \end{bmatrix} \overrightarrow{R_{23}} \begin{bmatrix} 1 & 2 & 3 & 0 \\ 3 & 4 & 5 & 6 \\ 2 & 0 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} A = E_{23}A.$$

That is, interchanging the two rows of the matrix A is same as multiplying on the left by the corresponding elementary matrix. In other words, we see that the left multiplication of elementary matrices to a matrix results in elementary row operations.

2. Consider the augmented matrix $[A\ \mathbf{b}] = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 2 & 0 & 3 & 5 \\ 1 & 1 & 1 & 3 \end{bmatrix}$. Then the result of the steps given below is same as the matrix product

$$E_{23}(-1)E_{12}(-1)E_3(1/3)E_{32}(2)E_{23}E_{21}(-2)E_{13}[A\ \mathbf{b}].$$

$$\begin{aligned}
\begin{bmatrix} 0 & 1 & 1 & 2 \\ 2 & 0 & 3 & 5 \\ 1 & 1 & 1 & 3 \end{bmatrix} \quad & \overrightarrow{R_{13}} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 2 & 0 & 3 & 5 \\ 0 & 1 & 1 & 2 \end{bmatrix} \overrightarrow{R_{21}(-2)} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & -2 & 1 & -1 \\ 0 & 1 & 1 & 2 \end{bmatrix} \overrightarrow{R_{23}} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & -2 & 1 & -1 \end{bmatrix} \\
& \overrightarrow{R_{32}(2)} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 3 & 3 \end{bmatrix} \overrightarrow{R_3(1/3)} \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix} \overrightarrow{R_{12}(-1)} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix} \\
& \overrightarrow{R_{23}(-1)} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}
\end{aligned}$$

Now, consider an $m \times n$ matrix A and an elementary matrix E of order n. Then multiplying by E on the right to A corresponds to applying column transformation on the matrix A. Therefore, for each elementary matrix, there is a corresponding column transformation. We summarize:

Definition 2.4.16 The column transformations obtained by right multiplication of elementary matrices are called **elementary column operations**.

Example 2.4.17 Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & 3 \\ 3 & 4 & 5 \end{bmatrix}$ and consider the elementary column operation f which interchanges the second and the third column of A. Then $f(A) = \begin{bmatrix} 1 & 3 & 2 \\ 2 & 3 & 0 \\ 3 & 5 & 4 \end{bmatrix} = A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = AE_{23}.$

Exercise 2.4.18

1. Let e be an elementary row operation and let $E = e(I)$ be the corresponding elementary matrix. That is, E is the matrix obtained from I by applying the elementary row operation e. Show that $e(A) = EA$.

2. Show that the Gauss elimination method is same as multiplying by a series of elementary matrices on the left to the augmented matrix.

   Does the Gauss-Jordan method also corresponds to multiplying by elementary matrices on the left? Give reasons.

3. Let A and B be two $m \times n$ matrices. Then prove that the two matrices A, B are row-equivalent if and only if $B = PA$, where P is product of elementary matrices. When is this P unique?

2.5 Rank of a Matrix

In previous sections, we solved linear systems using Gauss elimination method or the Gauss-Jordan method. In the examples considered, we have encountered three possibilities, namely

1. existence of a unique solution,

2. existence of an infinite number of solutions, and

3. no solution.

Based on the above possibilities, we have the following definition.

Definition 2.5.1 (Consistent, Inconsistent) A linear system is called CONSISTENT if it admits a solution and is called INCONSISTENT if it admits no solution.

The question arises, as to whether there are conditions under which the linear system $A\mathbf{x} = \mathbf{b}$ is consistent. The answer to this question is in the affirmative. To proceed further, we need a few definitions and remarks.

Recall that the row reduced echelon form of a matrix is unique and therefore, the number of non-zero rows is a unique number. Also, note that the number of non-zero rows in either the row reduced form or the row reduced echelon form of a matrix are same.

Definition 2.5.2 (Row rank of a Matrix) The number of non-zero rows in the row reduced form of a matrix is called the row-rank of the matrix.

By the very definition, it is clear that row-equivalent matrices have the same row-rank. For a matrix A, we write 'row-rank (A)' to denote the row-rank of A.

Example 2.5.3 1. Determine the row-rank of $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix}$.

Solution: To determine the row-rank of A, we proceed as follows.

(a) $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix} \xrightarrow{R_{21}(-2), R_{31}(-1)} \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$.

(b) $\begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \xrightarrow{R_2(-1), R_{32}(1)} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$.

(c) $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \xrightarrow{R_3(1/2), R_{12}(-2)} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

(d) $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_{23}(-1), R_{13}(1)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

The last matrix in Step 1d is the row reduced form of A which has 3 non-zero rows. Thus, row-rank$(A) = 3$. This result can also be easily deduced from the last matrix in Step 1b.

2. Determine the row-rank of $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 0 \end{bmatrix}$.

Solution: Here we have

(a) $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 0 \end{bmatrix} \xrightarrow{R_{21}(-2), R_{31}(-1)} \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{bmatrix}$.

(b) $\begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{bmatrix} \xrightarrow{R_2(-1), R_{32}(1)} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

From the last matrix in Step 2b, we deduce row-rank$(A) = 2$.

Remark 2.5.4 *Let $A\mathbf{x} = \mathbf{b}$ be a linear system with m equations and n unknowns. Then the row-reduced echelon form of A agrees with the first n columns of $[A \;\; \mathbf{b}]$, and hence*

$$\text{row-rank}(A) \leq \text{row-rank}([A \;\; \mathbf{b}]).$$

The reader is advised to supply a proof.

Remark 2.5.5 *Consider a matrix A. After application of a finite number of elementary column operations (see Definition 2.4.16) to the matrix A, we can have a matrix, say B, which has the following properties:*

1. *The first nonzero entry in each column is 1.*
2. *A column containing only 0's comes after all columns with at least one non-zero entry.*
3. *The first non-zero entry (the leading term) in each non-zero column moves down in successive columns.*

Therefore, we can define **column-rank** of A as the number of non-zero columns in B. It will be proved later that

$$\text{row-rank}(A) = \text{column-rank}(A).$$

Thus we are led to the following definition.

Definition 2.5.6 The number of non-zero rows in the row reduced form of a matrix A is called the **rank** of A, denoted rank (A).

Theorem 2.5.7 Let A be a matrix of rank r. Then there exist elementary matrices $E_1, E_2, \ldots, E_s$ and $F_1, F_2, \ldots, F_\ell$ such that

$$E_1 E_2 \ldots E_s \; A \; F_1 F_2 \ldots F_\ell = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Proof. Let C be the row reduced echelon matrix obtained by applying elementary row operations to the given matrix A. As rank$(A) = r$, the matrix C will have the first r rows as the non-zero rows. So by Remark 2.4.5, C will have r leading columns, say $i_1, i_2, \ldots, i_r$. Note that, for $1 \leq s \leq r$, the i_s^{th} column will have 1 in the s^{th} row and zero elsewhere.

We now apply column operations to the matrix C. Let D be the matrix obtained from C by successively interchanging the s^{th} and i_s^{th} column of C for $1 \leq s \leq r$. Then the matrix D can be written in the form $\begin{bmatrix} I_r & B \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$, where B is a matrix of appropriate size. As the $(1,1)$ block of D is an identity matrix, the block $(1,2)$ can be made the zero matrix by application of column operations to D. This gives the required result. $\square$

Exercise 2.5.8

1. Determine the ranks of the coefficient and the augmented matrices that appear in Part 1 and Part 2 of Exercise 2.4.12.
2. For any matrix A, prove that rank$(A) = $ rank(A^t).
3. Let A be an $n \times n$ matrix with rank$(A) = n$. Then prove that A is row-equivalent to I_n.

2.6 Existence of Solution of $A\mathbf{x} = \mathbf{b}$

We try to understand the properties of the set of solutions of a linear system through an example, using the Gauss-Jordan method. Based on this observation, we arrive at the existence and uniqueness results for the linear system $A\mathbf{x} = \mathbf{b}$. This example is more or less a motivation.

2.6.1 Example

Consider a linear system $A\mathbf{x} = \mathbf{b}$ which after the application of the Gauss-Jordan method reduces to a matrix $[C \;\; \mathbf{d}]$ with

$$[C \;\; \mathbf{d}] = \begin{bmatrix} 1 & 0 & 2 & -1 & 0 & 0 & 2 & 8 \\ 0 & 1 & 1 & 3 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

For this particular matrix $[C \;\; \mathbf{d}]$, we want to see the set of solutions. We start with some observations.
Observations:

1. The number of non-zero rows in C is 4. This number is also equal to the number of non-zero rows in $[C \;\; \mathbf{d}]$.

2. The first non-zero entry in the non-zero rows appear in columns $1, 2, 5$ and 6.

3. Thus, the respective variables x_1, x_2, x_5 and x_6 are the basic variables.

4. The remaining variables, x_3, x_4 and x_7 are free variables.

5. We assign arbitrary constants k_1, k_2 and k_3 to the free variables x_3, x_4 and x_7, respectively.

Hence, we have the set of solutions as

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} &= \begin{bmatrix} 8 - 2k_1 + k_2 - 2k_3 \\ 1 - k_1 - 3k_2 - 5k_3 \\ k_1 \\ k_2 \\ 2 + k_3 \\ 4 - k_3 \\ k_3 \end{bmatrix} \\ &= \begin{bmatrix} 8 \\ 1 \\ 0 \\ 0 \\ 2 \\ 4 \\ 0 \end{bmatrix} + k_1 \begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} 1 \\ -3 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + k_3 \begin{bmatrix} -2 \\ -5 \\ 0 \\ 0 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \end{aligned}$$

where k_1, k_2 and k_3 are arbitrary.

Let $\mathbf{u}_0 = \begin{bmatrix} 8 \\ 1 \\ 0 \\ 0 \\ 2 \\ 4 \\ 0 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ -3 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ and $\mathbf{u}_3 = \begin{bmatrix} -2 \\ -5 \\ 0 \\ 0 \\ 1 \\ -1 \\ 1 \end{bmatrix}$.

Then it can easily be verified that $C\mathbf{u}_0 = \mathbf{d}$, and for $1 \le i \le 3$, $C\mathbf{u}_i = \mathbf{0}$.

A similar idea is used in the proof of the next theorem and is omitted. The interested readers can read the proof in Appendix 14.1.

2.6.2 Main Theorem

Theorem 2.6.1 [Existence and Non-existence] Consider a linear system $A\mathbf{x} = \mathbf{b}$, where A is a $m \times n$ matrix, and $\mathbf{x}$, $\mathbf{b}$ are vectors with orders $n \times 1$, and $m \times 1$, respectively. Suppose rank $(A) = r$ and rank$([A \ \ \mathbf{b}]) = r_a$. Then exactly one of the following statement holds:

1. if $r_a = r < n$, the set of solutions of the linear system is an infinite set and has the form
$$\{\mathbf{u}_0 + k_1\mathbf{u}_1 + k_2\mathbf{u}_2 + \cdots + k_{n-r}\mathbf{u}_{n-r} \ : \ k_i \in \mathbb{R}, \ 1 \le i \le n-r\},$$
where $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-r}$ are $n \times 1$ vectors satisfying $A\mathbf{u}_0 = \mathbf{b}$ and $A\mathbf{u}_i = \mathbf{0}$ for $1 \le i \le n-r$.
2. if $r_a = r = n$, the solution set of the linear system has a unique $n \times 1$ vector $\mathbf{x}_0$ satisfying $A\mathbf{x}_0 = \mathbf{b}$.
3. If $r < r_a$, the linear system has no solution.

Remark 2.6.2 *Let A be an $m \times n$ matrix and consider the linear system $A\mathbf{x} = \mathbf{b}$. Then by Theorem 2.6.1, we see that the linear system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if*
$$\text{rank } (A) = \text{rank}([A \ \ \mathbf{b}]).$$

The following corollary of Theorem 2.6.1 is a very important result about the homogeneous linear system $A\mathbf{x} = \mathbf{0}$.

Corollary 2.6.3 Let A be an $m \times n$ matrix. Then the homogeneous system $A\mathbf{x} = \mathbf{0}$ has a non-trivial solution if and only if rank$(A) < n$.

PROOF. Suppose the system $A\mathbf{x} = \mathbf{0}$ has a non-trivial solution, $\mathbf{x}_0$. That is, $A\mathbf{x}_0 = \mathbf{0}$ and $\mathbf{x}_0 \neq \mathbf{0}$. Under this assumption, we need to show that rank$(A) < n$. On the contrary, assume that rank$(A) = n$. So,
$$n = \text{rank}(A) = \text{rank}\big([A \ \ \mathbf{0}]\big) = r_a.$$
Also $A\mathbf{0} = \mathbf{0}$ implies that $\mathbf{0}$ is a solution of the linear system $A\mathbf{x} = \mathbf{0}$. Hence, by the uniqueness of the solution under the condition $r = r_a = n$ (see Theorem 2.6.1), we get $\mathbf{x}_0 = \mathbf{0}$. A contradiction to the fact that $\mathbf{x}_0$ was a given non-trivial solution.

Now, let us assume that rank$(A) < n$. Then
$$r_a = \text{rank}\big([A \ \ \mathbf{0}]\big) = \text{rank}(A) < n.$$
So, by Theorem 2.6.1, the solution set of the linear system $A\mathbf{x} = \mathbf{0}$ has infinite number of vectors $\mathbf{x}$ satisfying $A\mathbf{x} = \mathbf{0}$. From this infinite set, we can choose any vector $\mathbf{x}_0$ that is different from $\mathbf{0}$. Thus, we have a solution $\mathbf{x}_0 \neq \mathbf{0}$. That is, we have obtained a non-trivial solution $\mathbf{x}_0$.