UNIT 8 MATRICES - II

Structure

8.1 INTRODUCTION

In Unit 7 we introduced you to a matrix and showed you how a system of linear equations can give us a matrix. An important reason for which linear algebra arose is the theory of simultaneous linear equations. A system of simultaneous linear equations can be translated into a matrix equation, and solved by using matrices.

The study of the rank of a matrix is a natural forerunner to the theory of simultaneous linear equations. Because, it is in terms of rank that we can find out whether a simultaneous system of equations has a solution or not. In this unit we start by studying the rank of a matrix. Then we discuss row operations on a matrix and use them for obtaining the rank and inverse of a matrix. Finally, we apply this knowledge to determine the nature of solutions of a system of linear equations. The method of solving a system of linear equations that we give here is by "successive elimination of variables". It is also called the Gaussian elimination process.

With this unit we finish Block 2. In the next block we will discuss concepts that are intimately related to matrices.

Objectives

After reading this unit, you should be able to

- obtain the rank of a matrix;
- reduce a matrix to the echelon form;
- obtain the inverse of a matrix by row-reduction;
- solve a system of simultaneous linear equations by the method of successive elimination of variables.

8.2 RANK OF A MATRIX

Consider any $m \times n$ matrix A, over a field **F**. We can associate two vector spaces with it, in a very natural way. Let us see what they are. Let $A = [a_{ij}]$. A has m rows, say, $R_1, R_2, \ldots, R_m$, where $R_1 = (a_{11}, a_{12}, \ldots, a_{1n})$, $R_2 = (a_{21}, a_{22}, \ldots, a_{2n}) \ldots, R_m = (a_{m1}, a_{m2}, \ldots, a_{mn})$.

Thus, $R_i \in \mathbf{F}^n \ \forall i$, and $A = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{bmatrix}$

The subspace of $\mathbf{F}^n$ generated by the row vectors $R_1, \ldots, R_m$ of A, is called the **row space** of A, and is denoted by **RS** (A).

Example 1: If $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, does $(0, 0, 1) \in$ RS (A)?

Solution : The row space of A is the subspace of $\mathbf{R}^3$ generated by (1, 0, 0) and (0, 1, 0). Therefore, RS (A) = {(a, b, 0)| a, b ∈ **R**}. Therefore (0, 0, 1) ∉ RS (A).

ρ is the Greek letter 'rho'

The dimension of the row space of A is called the **row rank** of A, and is denoted by ρ_r (A).

Thus, ρ_r (A) = maximum number of linearly independent rows of A.

In Example 1, ρ_r(A) = 2 = number of rows of A. But consider the next example.

Example 2: If $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 0 \end{bmatrix}$, find ρ_r (A).

Solution: The row space of A is the subspace of $\mathbf{R}^2$ generated by (1,0), (0,1) and (2,0). But (2,0) already lies in the vector space generated by (1,0) and (0,1), since (2,0) = 2 (1,0). Therefore, the row space of A is generated by the linearly independent vectors (1, 0) and (0,1). Thus, ρ_r (A) = 2.

So, in Example 2 , ρ_r (A) < number of rows of A.

In general, for any m × n matrix A, RS (A) is generated by m vectors. Therefore, ρ_r (A) ≤ m. Also, RS (A) is a subspace of $\mathbf{F}^n$ and $\dim_{\mathbf{F}} \mathbf{F}^n = n$. Therefore, ρ_r (A) ≤ n.

min (m, n) denotes 'the minimum of the numbers of m and n'.

Thus, **for any m × n matrix A, 0 ≤ ρ_r (A) ≤ min (m, n).**

E E1) Show that A = **0** ⇔ ρ_r (A) = **0**.

Just as we have defined the row space of A, we can define the column space of A. Each column of A is an m-tuple, and hence belongs to $\mathbf{F}^m$. We denote the columns of A by $C_1,...,C_n$. The subspace of $\mathbf{F}^m$ generated by {$C_1,...,C_n$} is called the **column space** of A and is denoted by **CS (A)**. The dimension of CS (A) is called the **column rank** of A, and is denoted of ρ_c (A). Again, since CS (A) is generated by n vectors and is a subspace of $\mathbf{F}^m$, we get 0 ≤ ρ_c (A) ≤ min (m, n).

E E2) Obtain the column rank and row rank of $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}$

In E2 you may have noticed that the row and column ranks of A are equal. In fact, in Theorem 1, we prove that ρ_r (A) = ρ_c (A), for any matrix A. But first, we prove a lemma.

Lemma 1: Let A, B be two matrices over **F** such that AB is defined. Then

a) CS (AB) ⊆ CS (A),

b) RS (AB) ⊆ RS (B).

Thus, ρ_c (AB) ≤ ρ_c (A), ρ_r (AB) ≤ ρ_r (B).

Proof: (a) Suppose A = [a_{ij}] is an m × n matrix and B = [b_{ij}] is an n × p matrix. Then, from Sec. 7.5, you know that the jth column of C = AB will be

$$\begin{bmatrix} c_{1j} \\ c_{2j} \\ \vdots \\ c_{mj} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{n} a_{1k} b_{kj} \\ \sum_{k=1}^{n} a_{2k} b_{kj} \\ \vdots \\ \sum_{k=1}^{n} a_{mk} b_{kj} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} b_{1j} + \ldots + \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} b_{nj}$$

$$= C_1 b_{1j} + \ldots + C_n b_{nj},$$

where $C_1, \ldots, C_n$ are the columns of A.

Thus, the columns of AB are linear combinations of the columns of A. Thus, the columns of $AB \in CS(A)$. So, $CS(AB) \subseteq CS(A)$.
Hence, $\rho_c(AB) \le \rho_c(A)$.

b) By a similar argument as above, we get $RS(AB) \subseteq RS(B)$, and so, $\rho_r(AB) \le \rho_r(B)$.

E E3) Prove (b) of Lemma 1.

We will now use Lemma 1 for proving the following theorem.

Theorem 1: $\rho_c(A) = \rho_r(A)$, for any matrix **A** over **F**.

Proof: Let $A \in \mathbf{M_{m \times n}(F)}$ Suppose $\rho_r(A) = r$ and $\rho_c(A) = t$.

Now, $RS(A) = [\{R_1, R_2, \ldots, R_m\}]$, where $R_1, R_2, \ldots, R_m$ are the rows of A. Let $\{e_1, e_2, \ldots, e_r\}$ be a basis of RS(A). Then R_i is a linear combination of $e_1, \ldots, e_r$, for each $i = 1, \ldots, m$. Let

$$R_i = \sum_{j=1}^{r} b_{ij} e_j, \quad i = 1, 2, \ldots, m, \quad \text{where } b_{ij} \in \mathbf{F} \text{ for } 1 \le i \le m,\ 1 \le j \le r.$$

We can write these equations in matrix form as

$$\begin{bmatrix} R_1 \\ \vdots \\ R_m \end{bmatrix} = \begin{bmatrix} b_{11} \ldots\ldots b_{1r} \\ \ldots\ldots\ldots \\ \ldots\ldots\ldots \\ \ldots\ldots\ldots \\ b_{m1} \ldots\ldots b_{mr} \end{bmatrix} \begin{bmatrix} e_1 \\ \vdots \\ e_r \end{bmatrix}$$

So, $A = BE$, where $B = [b_{ij}]$ is an $m \times r$ matrix and E is the $r \times n$ matrix with rows $e_1, e_2, \ldots, e_r$. (Remember, $e_i \in \mathbf{F}^n$, for each $i = 1, \ldots, r$.)

So, $t = \rho_c(A) = \rho_c(BE) \le \rho_c(B)$, by Lemma 1.

$$\le \min(m, r,)$$

$$\le r$$

Thus, $t \le r$.

Just as we got $A = BE$ above, we get $A = [f_1, \ldots, f_t]D$, where $\{f_1, \ldots, f_t\}$ is a basis of the column space of A and D is a $t \times n$ matrix. Thus, $r = \rho_r(A) \le \rho_r(D) \le t$, by Lemma 1.
So we get $r \le t$ and $t \le r$. This gives us $r = t$.

Theorem 1 allows us to make the following definition.

Definition: The integer $\rho_c(A)$ $(=\rho_r(A))$ is called the **rank** of A, and is denoted by $\rho(A)$.

You will see that Theorem 1 is very helpful if we want to prove any fact about $\rho(A)$. If it is easier to deal with the rows of A we can prove the fact for $\rho_r(A)$. Similarly, if it is easier to deal with the columns of A, we can prove the fact for $\rho_c(A)$. While proving Theorem 3 we have used this facility that Theorem 1 gives us.

Use Theorem 1 to solve the following exercises.

E E4) If A,B are two matrices such that AB is defined then show that $\rho(AB) \leq \min(\rho(A), \rho(B))$.

E E5) Suppose $C \neq 0 \in M_{m\times 1}(F)$, and $R \neq 0 \in M_{1\times n}(F)$, then show that the rank of the $m \times n$ matrix CR is 1. (Hint: Use E4).

Does the term 'rank' seem familiar to you? Do you remember studying about the rank of a linear transformation in Unit 5? We will now see if the rank of a linear transformation is related to the rank of its matrix. The following theorem brings forth the precise relationship. (Go through Sec. 5.3 before going further.)

Theorem 2: Let U,V be vector spaces over F of dimensions n and m, respectively. Let B_1 be a basis of U and B_2 be a basis of V. Let $T \in L(U,V)$.

Then $R(T) \simeq CS([T]_{B_1, B_2})$.

Thus, rank (T) = rank of $[T]_{B_1, B_2}$.

Proof: Let $B_1 = \{e_1, e_2, ..., e_n\}$ and $B_2 = \{f_1, f_2, ..., f_m\}$. As in the proof of Theorem 7 of Unit 7, $\theta: V \rightarrow M_{m\times 1}(F)$: $\theta(v)$ = coordinate vector of v with respect to the basis B_2, is an isomorphism.

Now, $R(T) = [\{T(e_1), T(e_2), ..., T(e_n)\}]$. Let $A = [T]_{B_1, B_2}$ have $C_1, C_2, ..., C_n$ as its columns. Then $CS(A) = [\{C_1, C_2, ..., C_n\}]$. Also, $\theta(T(e_i)) = C_i \; \forall i = 1, ..., n$.

Thus, $\theta: R(T) \rightarrow CS(A)$ is an isomorphism. $\therefore R(T) \simeq CS(A)$.

In particular, dim R(T) = dim CS(A) = $\rho(A)$.

That is, rank (T) = $\rho(A)$.

Theorem 2 leads us to the following corollary. It says that pre-multiplying or post-multiplying a matrix by invertible matrices does not alter its rank.

Corollary 1: Let A be an $m \times n$ matrix. Let P,Q be $m \times m$ and $n \times n$ invertible matrices, respectively.

Then $\rho(PAQ) = \rho(A)$.

Proof: Let $T \in L(U,V)$ be such that $[T]_{B_1, B_2} = A$. We are given matrices Q and P^{-1}. Therefore, by Theorem 8 of Unit 7, $\exists$ bases B_1' and B_2' of U and V, respectively, such that $Q = M_{B_1}^{B_1'}$ and $P^{-1} = M_{B_2}^{B_2'}$

Then, by Theorem 10 of Unit 7,

$[T]_{B_1', B_2'} = M_{B_2'}^{B_2} [T]_{B_1, B_2} M_{B_1}^{B_1'} = PAQ$

In other words, we can change the bases suitably so that the matrix of T with respect to the new bases is PAQ.

So, by Theorem 2, $\rho(PAQ) = \text{rank}(T) = \rho(A)$. Thus, $\rho(PAQ) = \rho(A)$.

E E6) Take $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \end{bmatrix}$, $P = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $Q = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Obtain PAQ and show that $\rho(PAQ) = \rho(A)$.

Now we state and prove another corollary to Theorem 2. This corollary is useful because it transforms any matrix into a very simple matrix, namely, a matrix whose entries are 1 and 0 only.

Corollary 2: Let A be an $m \times n$ matrix with rank r. Then $\exists$ invertible matrices P and Q such that $PAQ = \begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$

Proof: Let $T \in L(U,V)$ be such that $[T]_{B_1, B_2} = A$. Since $\rho(A) = r$, rank $(T) = r$. $\therefore$ nullity $(T) = n - r$ (Unit 5, Theorem 5).

Let $\{u_1, u_2, \ldots, u_{n-r}\}$ be a basis of Ker T. We extend this to form the basis

$B_1' = \{u_1, u_2, \ldots u_{n-r}, u_{n-r+1}, \ldots, u_n\}$ of U. Then $\{T(u_{n-r+1}), \ldots, T(u_n)\}$ is a basis of R(T) (See Unit 5, proof of Theorem 5). Extend this set to form a basis B_2' of V, say $B_2' = \{T(u_{n-r+1}), \ldots, T(u_n), v_1, \ldots, v_{m-r}\}$. Let us reorder the elements of B_1' and write it as $B_1' = \{u_{n-r+1}, \ldots, u_n, u_1, \ldots, u_{n-r}\}$.

Then, by definition, $[T]_{B_1', B_2'} = \begin{bmatrix} I_r & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(m-r) \times r} & \mathbf{0}_{(m-r) \times (n-r)} \end{bmatrix}$

where $\mathbf{0}_{s \times t}$ denotes the zero matrix of size $s \times t$. (Remember that $u_1, \ldots, u_{n-r} \in$ Ker T.)

Hence, $PAQ = \begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$

where $Q = M_{B_1}^{B_1'}$ and $P = M_{B_2'}^{B_2}$, by Theorem 10 of Unit 7.

Note: $\begin{bmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$ is called the **normal form** of the matrix A.

Consider the following example, which is the converse of E5.

Example 3: A is an $m \times n$ matrix of rank 1, show that $\exists\, C \neq \mathbf{0}$ in $M_{m \times 1}(F)$ and $R \neq \mathbf{0}$ in $M_{1 \times n}(F)$ such that $A = CR$.

Solution: By Corollary 2 (above), $\exists$ P, Q such that

$$PAQ = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix}, \text{ since } \rho(A) = 1.$$

$$= \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} [1\ 0 \ldots 0]$$

$$\therefore A = P^{-1}(PAQ)Q^{-1} = P^{-1}\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}[1\ 0 \ldots 0]Q^{-1} = CR,$$

$$\text{where } C = P^{-1}\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \neq \mathbf{0},\ R = [1\ 0 \ldots 0]Q^{-1} \neq \mathbf{0}.$$

E E7) What is the normal form of diag (1, 2, 3)?

The solution of E7 is a particular case of the general phenomenon: the normal form of an $n \times n$ invertible matrix is I_n.

Let us now look at some ways of transforming a matrix by playing around with its rows. The idea is to get more and more entries of the matrix to be zero. This will help us in solving systems of linear equations.

8.3 ELEMENTARY OPERATIONS

Consider the following set of 2 equations in 3 unknowns x, y and z:

$x + y + z = 1$

$2x + 3z = 0$

How can you express this system of equations in matrix form?
One way is

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

In general, if a system of m linear equations in n variables, $x_1, \ldots, x_n$, is

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$$

$$\vdots \qquad \vdots \qquad\qquad \vdots \qquad \vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m$$

where $a_{ij}, b_i \in \mathbf{F}\ \forall\ i = 1, \ldots, m$ and $j = 1, \ldots, n$, then this can be expressed as $AX = B$,

$$\text{where } A = [a_{ij}]_{m\times n},\ X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},\ \mathbf{B} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

In this section we will study methods of changing the matrix A to a very simple form so that we can obtain an immediate solution to the system of linear equations AX = B. For this purpose, we will always be multiplying A on the left or the right by a suitable matrix. In effect, we will be applying elementary row or column operations on A.

8.3.1 Elementary Operations on a Matrix

Let A be an $m \times n$ matrix. As usual, we denote its rows by $R_1, \ldots, R_m$, and columns by $C_1, \ldots, C_n$. We call the following operations **elementary row operations:**

1) Interchanging R_i and R_j, for $i \neq j$.

2) Multiplying R_i by some $a \in \mathbf{F}$, $a \neq 0$.

3) Adding aR_j to R_i, where $i \neq j$ and $a \in \mathbf{F}$.

We denote the operation (1) by $\mathbf{R_{ij}}$, (2) by $\mathbf{R_i(a)}$, (3) by $\mathbf{R_{ij}(a)}$.

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

then $R_{12}(A) = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix}$ (interchanging the two rows).

Also $R_2(3)(A) = \begin{bmatrix} 1 & 2 & 3 \\ 0 \times 3 & 1 \times 3 & 2 \times 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 6 \end{bmatrix}$,

and $R_{12}(2)(A) = \begin{bmatrix} 1 + 0 \times 2 & 2 + 1 \times 2 & 3 + 2 \times 2 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 7 \\ 0 & 1 & 2 \end{bmatrix}$

E

E8) If $A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, what is

a) $R_{21}(A)$ b) $R_{32} \circ R_{21}(A)$ c) $R_{13}(-1)(A)$?

Just as we defined the row operations, we can define the three column operations as follows:

1) Interchanging C_i and C_j for $i \neq j$, denoted by $\mathbf{C_{ij}}$.

2) Multiplying C_i by $a \in \mathbf{F}$, $a \neq 0$, denoted by $\mathbf{C_i(a)}$.

3) Adding aC_j to C_i, where $a \in \mathbf{F}$, denoted by $\mathbf{C_{ij}(a)}$.

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

then $C_{21}(10)(A) = \begin{bmatrix} 1 & 13 \\ 2 & 24 \end{bmatrix}$

and $C_{12}(10)(A) = \begin{bmatrix} 31 & 3 \\ 42 & 4 \end{bmatrix}$.

We will now prove a theorem which we will use in Sec. 8.3.2 for obtaining the rank of a matrix easily.

Theorem 3: Elementary operations on a matrix do not alter its rank.

Proof: The way we will prove the statement is to show that the row space remains unchanged under row operations and the column space remains unchanged under column operations. This means that the row rank and the column rank remain unchanged. This immediately shows, by Theorem 1, that the rank of the matrix remains unchanged.

Now, let us show that the row space remains unaltered. Let $R_1,\ldots,R_m$ be the rows of a matrix A. Then the row space of A is generated by $\{R_1,\ldots,R_i,\ldots,R_j,\ldots,R_m\}$. On applying R_{ij} to A, the rows of A remain the same. Only their order gets changed. Therefore, the row space of $R_{ij}(A)$ is the same as the row space of A.

If we apply $R_i(a)$, for $a \in \mathbf{F}$, $a \neq 0$, then any linear combination of $R_1,\ldots,R_m$ is $a_1R_1 + \ldots + a_mR_m = a_1R_1 + \ldots + \frac{a_i}{a}aR_i + \ldots + a_mR_m$, which is a linear combination of $R_1,\ldots,aR_i,\ldots,R_m$.

Thus, $[\{R_1,\ldots,R_i,\ldots,R_m\}] = [\{R_1,\ldots,aR_i,\ldots,R_m\}]$. That is, the row space of A is the same as the row space of $R_i(a)(A)$.

If we apply $R_{ij}(a)$, for $a \in \mathbf{F}$, then any linear combination

$b_1R_1 + \ldots + b_iR_i + \ldots + b_jR_j + \ldots + b_mR_m = b_1R_1 + \ldots + b_i(R_i + aR_j) + \ldots + (b_j - b_ia)R_j + \ldots + b_mR_m$.

Thus, $[\{R_1,\ldots,R_m\}] = [\{R_1,\ldots,R_i + aR_j,\ldots,R_j,\ldots,R_m\}]$.

Hence, the row space of A remains unaltered under any elementary row operations.

We can similarly show that the column space remains unaltered under elementary column operations.

Elementary operations lead us to the following definition.

Definition: A matrix obtained by subjecting I_n to an elementary row or column operation is called an **elementary matrix**.

For example, $C_{12}(I_3) = C_{12}\left(\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}\right) = \begin{bmatrix}0 & 1 & 0\\1 & 0 & 0\\0 & 0 & 1\end{bmatrix}$ is an elementary matrix.

Since there are six types of elementary operations, we get six types of elementary matrices, but not all of them are different.

E E9) Check that $R_{23}(I_4) = C_{23}(I_4)$, $R_2(2)(I_4) = C_2(2)(I_4)$ and $R_{12}(3)(I_4) = C_{21}(3)(I_4)$

In general, $R_{ij}(I_n) = C_{ij}(I_n)$, $R_i(a)(I_n) = C_i(a)(I_n)$ for $a \neq 0$, and $R_{ij}(a)(I_n) = C_{ji}(a)(I_n)$ for $i \neq j$ and $a \in \mathbf{F}$.

Thus, **there are only three types of elementary matrices.** We denote

$R_{ij}(I) = C_{ij}(I)$ by $\mathbf{E_{ij}}$,

$R_i(a)(I) = C_i(a)(I)$, (if $a \neq 0$) by $\mathbf{E_i(a)}$ and

$R_{ij}(a)(I) = C_{ji}(a)(I)$ by $\mathbf{E_{ij}(a)}$ for $i \neq j$, $a \in \mathbf{F}$.

E_{ij}, $E_i(a)$ and $E_{ij}(a)$ are called the elementary matrices corresponding to the pairs R_{ij} and C_{ij}, $R_i(a)$ and $C_i(a)$, $R_{ij}(a)$ and $C_{ji}(a)$, respectively.

Caution: $E_{ij}(a)$ corresponds to $C_{ji}(a)$, and not $C_{ij}(a)$.

Now, see what happens to the matrix

$A = \begin{bmatrix}0 & 1 & 2\\3 & 0 & 0\\2 & 1 & 0\end{bmatrix}$ if we multiply it on the left by

$E_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ We get

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 2 \\ 2 & 1 & 0 \end{bmatrix} = R_{12}(A)$$

Similarly, $AE_{12} = C_{12}(A)$.

Again, consider $E_3(2)A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix}$

$$= \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 4 & 2 & 0 \end{bmatrix} = R_3(2)(A)$$

Similarly, $AE_3(2) = C_3(2)(A)$

Finally, $E_{13}(5)A = \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 10 & 6 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix}$

$= R_{13}(5)(A)$

But, $AE_{13}(5) = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 0 & 15 \\ 2 & 1 & 10 \end{bmatrix}$

$= C_{31}(5)(A)$.

What you have just seen are examples of a general phenomenon. We will now state this general result formally. (Its proof is slightly technical, and so, we skip it.)

Theorem 4: For any matrix A

a) $R_{ij}(A) = E_{ij}A$

b) $R_i(a)(A) = E_i(a)A$, for $a \neq 0$.

c) $R_{ij}(a)(A) = E_{ij}(a)A$

d) $C_{ij}(A) = AE_{ij}$

e) $C_i(a)(A) = AE_i(a)$, for $a \neq 0$

f) $C_{ij}(a)(A) = AE_{ji}(a)$

In (f) note the change of indices i and j.

An immediate corollary to this theorem shows that all the elementary matrices are invertible (see Sec. 7.6).

Corollary: An elementary matrix is invertible. In fact,

a) $E_{ij}E_{ij} = I$,

b) $E_i(a^{-1})\,E_i(a) = I$, for $a \neq 0$.

c) $E_{ij}(-a)\,E_{ij}(a) = I$.

Proof: We prove (a) only and leave the rest to you (see E10).

Now, from Theorem 4,

$E_{ij}E_{ij} = R_{ij}(E_{ij}) = R_{ij}(R_{ij}(I)) = I$, by definition of R_{ij}.

E E10) Prove (b) and (c) of the corollary above.

The corollary tells us that the **elementary matrices are invertible** and **the inverse of an elementary matrix is also an elementary matrix of the same type.**

E F 11) Actually multiply the two 4 × 4 matrices $E_{13}(-2)$ and $E_{13}(2)$ to get I_4.

And now we will introduce you to a very nice type of matrix, which any matrix can be transformed to by applying elementary operations.

8.3.2 Row-reduced Echelon Matrices

Consider the matrix

$$\begin{bmatrix} 1 & 0 & 9 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

In this matrix the three non-zero rows come before the zero row, and the first non-zero entry in each non-zero row is 1. Also, below this 1, are only zeros. This type of matrix has a special name, which we now give.

An echelon matrix is so-called because of the step-like structure of its non-zero rows.

Definition: An m × n matrix A is called a **row-reduced echelon matrix** if

a) the non-zero rows come before the zero rows,

b) in each non-zero row, the first non-zero entry is 1, and

c) the first non-zero entry in every non-zero row (after the first row) is to the right of the first non-zero entry in the preceding row.

Is $\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \end{bmatrix}$ a row-reduced echelon matrix? Yes. It satisfies all the conditions of the definition. On the other hand, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ are not row-reduced echelon matrices, since they violate conditions (a), (b) and (c), respectively.

The matrix

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

is a 6 × 11 row-reduced echelon matrix. The dotted line in it is to indicate the step-like structure of the non-zero rows.

But, why bring in this type of a matrix? Well the following theorem gives us one good reason.

Theorem 5: The rank of a row-reduced echelon matrix is equal to the number of its non-zero rows.

Proof: Let $R_1, R_2, \ldots, R_r$ be the non-zero rows of an m × n row-reduced echelon matrix, E. Then RS(E) is generated by $R_1, \ldots, R_r$. We want to show that $R_1, \ldots, R_r$ are linearly independent. Suppose R_1 has its first non-zero entry in column k_1, R_2 in column k_2, and so on. Then, for any r scalars $c_1, \ldots, c_r$ such that $c_1R_1 + c_2R_2 + \ldots + c_rR_r = \mathbf{0}$, we immediately get

$$\begin{array}{ll} & \qquad\quad k_1 \qquad\qquad k_2 \qquad\qquad k_r \\ & \qquad\quad \downarrow \qquad\qquad\; \downarrow \qquad\qquad \downarrow \\ c_1 & [0, \ldots\ldots, 0, 1, *, \ldots\ldots\ldots, *, \ldots\ldots\ldots\ldots, *, \quad \ldots\ldots, *] \\ + c_2 & [0, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, 0, 1, \; *, \ldots\ldots\ldots, *, \quad \ldots\ldots *] \\ \vdots & \vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \vdots \\ + c_r & [0, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, 0, 1, \; *, \ldots\ldots *] \\ = & [0, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 0] \end{array}$$

where $*$ denotes various entries that we aren't bothering to calculate.

This equation gives us the following equations (when we equate the k_1th entries, the k_2th entries , ..., the k_rth entries on both sides of the equation):

$c_1 = 0,\ c_1(*) + c_2 = 0, \ldots, c_1(*) + c_2(*) + \ldots + c_{r-1}(*) + c_r = 0.$

On solving these equations we get

$c_1 = 0 = c_2 = \ldots = c_r$. $\therefore R_1, \ldots, R_r$ are linearly independent $\therefore \rho(E) = r$.

Not only is it easy to obtain the rank of an echelon matrix, one can also solve linear equations of the type AX = B more easily if A is in echelon form.

Now, here is some good news!

Every matrix can be transformed to the row echelon form by a series of elementary row operations. We say that the matrix is **reduced** to the row echelon form. Consider the following example.

Example 4: Let $A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & -1 & -1 & 1 \\ 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 & 4 & 1 \\ 0 & 2 & 4 & 1 & 10 & 2 \end{bmatrix}$

Reduce A to the row echelon form.

Solution: The first column of A is zero. The second column is non-zero. The (1,2)th element is 0. We want 1 at this position. We apply R_{12} **to A and get**

$$A_1 = \begin{bmatrix} 0 & 1 & 2 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 & 4 & 1 \\ 0 & 2 & 4 & 1 & 10 & 2 \end{bmatrix}$$

The (1,2)th entry has become 1. Now, we subtract multiples of the first row from other rows so that the (2,2)th, (3,2)th, (4,2)th and (5,2)th entries become zero. So we apply $R_{31}(-1)$, and $R_{51}(-2)$, and get

$$A_2 = \begin{bmatrix} 0 & 1 & 2 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 3 & 12 & 0 \end{bmatrix}$$

Now, beneath the entries of the first row we have zeros in the first 3 columns, and in the fourth column we find non-zero entries. We want 1 at the (2,4)th position, so we interchange the 2nd and 3rd rows. We get

$$A_3 = \begin{bmatrix} 0 & 1 & 2 & -1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 3 & 12 & 0 \end{bmatrix}$$

We now subtract suitable multiples of the 2nd row from the 3rd, 4th and 5th rows so that the (3,4)th, (4,4)th and (5,4)th entries all become zero. $\therefore$,

$$A_3 \overset{R_{42}(-1),\ R_{52}(-3)}{\sim} \begin{bmatrix} 0 & 1 & 2 & -1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = A_4$$

Now we have zeros below the entries of the 2nd row, except for the 6th column. The (3,6)th element is 1. We subtract suitable multiples of the 3rd row from the 4th and 5th rows so that the (4,6)th, (5,6)th elements become zero. $\therefore$,

$A \overset{R}{\sim} B$ means that on applying the operation R to A we get the matrix B.

$$\overset{R_{41}(-1)}{A_4 \sim} \begin{bmatrix} 0 & 1 & 2 & -1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

And now we have achieved a row echelon matrix. Notice that we applied 7 elementary operations to A to obtain this matrix.

In general, we have the following theorem.

Theorem 6: Every matrix can be reduced to a row-reduced echelon matrix by a finite sequence of elementary row operations.

The proof of this result is just a repetition of the process that you went through in Example 4.

For practice, we give you the following exercise.

E E12) Reduce the matrix $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 3 & 1 & 0 \end{bmatrix}$ to echelon form.

Theorem 6 leads us to the following definition.

Definition: If a matrix A is reduced to a row-reduced echelon matrix E by a finite sequence of elementary row operations then E is called a **row-reduced echelon form** (or, the row echelon form) of A. We now give a useful result that immediately follows from Theorems 3 and 5.

Theorem 7: Let E be a row-reduced echelon form of A. Then the rank of A = number of non-zero rows of E.

Proof: We obtain E from A by applying elementary operations. Therefore, by Theorem 3, $\rho(A) = \rho(E)$. Also, $\rho(E)$ = the number of non-zero rows of E, by Theorem 5.

Thus, we have proved the theorem.

Let us look at some examples to actually see how the echelon form of a matrix simplifies matters.

Example 5: Find $\rho(A)$, where

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

by reducing it to its row-reduced echelon form.

Solution: $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 5 & 6 \end{bmatrix} \overset{R_{21}(-1)}{\sim} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 3 \end{bmatrix} \overset{R_2(1/3)}{\sim} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \end{bmatrix}$

which is the desired row-reduced echelon form. This has 2 non-zero rows. Hence, $\rho(A) = 2$.

E E13) Obtain the row-reduced echelon form of the matrix

$$A = \begin{bmatrix} 1 & 2 & 0 & 5 \\ 2 & 1 & 7 & 6 \\ 4 & 5 & 7 & 10 \end{bmatrix}$$

Hence determine the rank of the matrix.

By now you must have got used to obtaining row echelon forms. Let us discuss some ways of applying this reduction.

8.4 APPLICATIONS OF ROW-REDUCTION

In this section we shall see how to utilise row-reduction for obtaining the inverse of a matrix, and for solving a system of linear equations.

8.4.1 Inverse of a Matrix

In Theorem 4 you discovered that applying a row transformation to a matrix A is the same as multiplying it on the left by a suitable elementary matrix. Thus, applying a series of row transformations to A is the same as pre-multiplying A by a series of elementary matrices. This means that, after the nth row transformation we obtain the matrix $E_n E_{n-1} \ldots E_2 E_1 A$, where $E_1, E_2, \ldots, E_n$, are elementary matrices.

Now, how do we use this knowledge for obtaining the inverse of an invertible matrix? Suppose we have an n × n invertible matrix A. We know that A = IA, where $I = I_n$. Now, we apply a series of elementary row operations $E_1, \ldots, E_s$ to A so that A gets transformed to I_n. Thus,

$$I = E_s E_{s-1} \ldots E_2 E_1 A = E_s E_{s-1} \ldots E_2 E_1 (IA)$$
$$= (E_s E_{s-1} \ldots E_2 E_1 I)A = BA$$

where $B = E_s \ldots E_1 I$. Then, B is the inverse of A!

Note that we are reducing A to I, and not only to the row echelon form.

We illustrate this below.

Example 6: Determine if the matrix

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

is invertible. If it is invertible, find its inverse.

Solution: Can we transform A to I? If so, then A will be invertible.

$$\text{Now, } A = IA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$$

To transform A we will be pre-multiplying it by elementary matrices. We will also be pre-multiplying IA by these matrices. Therefore, as A is transformed to I, the same transformations are done to I on the right hand side of the matrix equation given above. Now

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -5 \\ 0 & -5 & -7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} A \text{ (applying } R_{21}(-2) \text{ and } R_{31}(-3) \text{ to A)}$$

$$\Rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 3 & 0 & -1 \end{bmatrix} A \quad \text{(applying } R_2(-1) \text{ and } R_3(-1))$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & -7 \\ 0 & 1 & 5 \\ 0 & 0 & -18 \end{bmatrix} = \begin{bmatrix} -3 & 2 & 0 \\ 2 & -1 & 0 \\ -7 & 5 & -1 \end{bmatrix} A \text{ (applying } R_{12}(-2) \text{ and } R_{32}(-5))$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & -7 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -3 & 2 & 0 \\ 2 & -1 & 0 \\ 7/18 & -5/18 & 1/18 \end{bmatrix} A \text{ (applying } R_3(-1/18))$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -5/18 & 1/18 & 7/18 \\ 1/18 & 7/18 & -5/18 \\ 7/18 & -5/18 & 1/18 \end{bmatrix} A \text{ (applying } R_{13}(7) \text{ and } R_{23}(-5))$$

Hence, A is invertible and its inverse is $B = 1/18 \begin{bmatrix} -5 & 1 & 7 \\ 1 & 7 & -5 \\ 7 & -5 & 1 \end{bmatrix}$

To make sure that we haven't made a careless mistake at any stage, check the answer by multiplying B with A. Your answer should be I.

E E14) Show that $\begin{bmatrix} 0 & 1 & 3 \\ 2 & 3 & 5 \\ 3 & 5 & 7 \end{bmatrix}$ is invertible. Find its inverse.

Let us now look at another application of row-reduction.

8.4.2 Solving a System of Linear Equations

Any system of m linear equations, in n unknowns $x_1, \ldots, x_n$, is

$$\begin{array}{ccc} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \vdots \quad \vdots \quad \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m, \end{array}$$

where all the a_{ij} and b_i are scalars.

This can be written in matrix form as

$$AX = B, \text{ where } A = [a_{ij}], X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, B = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

If B = 0, the system is called **homogeneous.** In this situation we are in a position to say how many linearly independent solutions the system of equations has.

Theorem 8: The number of linearly independent solutions of the matrix equation $AX = \mathbf{0}$ is $n - r$, where A is an $m \times n$ matrix and $r = \rho(A)$.

Proof: In Unit 7 you studied that given the matrix A, we can obtain a linear transformation $T: \mathbf{F}^n \to \mathbf{F}^m$ such that $[T]_{B,B'} = A$, where B and B′ are bases of $\mathbf{F}^n$ and $\mathbf{F}^m$, respectively.

Now, $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is a solution of $AX = \mathbf{0}$ if and only if it lies in Ker T (since $T(X) = AX$).

Thus, the number of linearly independent solutions is dim Ker T = nullity (T) = n − rank (T) (Unit 5, Theorem 5.)

Also, rank (T) = $\rho(A)$ (Theorem 2)

Thus, the number of linearly independent solutions is $n - \rho(A)$.

This theorem is very useful for finding out whether a homogeneous system has any non-trivial solutions or not.

Example 7: Consider the system of 3 equations in 3 unknowns:

$3x - 2y + z = 0$

$x + y \quad\quad = 0$

$x - 3z \quad\quad = 0$

How many solutions does it have which are linearly independent over **R**?

Solution: Here our coefficient matrix, $A = \begin{bmatrix} 3 & -2 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & -3 \end{bmatrix}$

Thus, n = 3. We have to find r. For this, we apply the row-reduction method. We obtain

$A \sim \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$, which is in echelon form and has rank 3.

Thus, $\rho(A) = 3$.

Thus, the number of linearly independent solutions is $3 - 3 = 0$. This means that this system of equation has no non-zero solution.

In Example 7 the number of unknowns was equal to the number of equations, that is, n = m. What happens if n > m?

A system of m homogeneous equations in n unknowns has a non-zero solution if n > m. Why? Well, if n > m, then the rank of the coefficient matrix is less than or equal to m, and hence, less than n. So, $n - r > 0$. Therefore, at least one non-zero solution exists.

Note: If a system $AX = \mathbf{0}$ has one solution, X_0, then it has an infinite number of solutions of the form cX_0, $c \in \mathbf{F}$. This is because $AX_0 = \mathbf{0} \Rightarrow A(cX_0) = \mathbf{0} \; \forall c \in \mathbf{F}$.

E E15) Give a set of linearly independent solutions for the system of equations

$x + 2y + 3z = 0$

$2x + 4y + z = 0$

Now consider the general equation AX = B, where A is an $m \times n$ matrix. We form the **augmented matrix** [A B]. This is an $m \times (n + 1)$ matrix whose last column is the matrix B. Here, we also include the case $B = \mathbf{0}$.

Interchanging equations, multiplying an equation by a non-zero scalar, and adding to any equation a scalar times some other equation does not alter the set of solutions of the system of equations. In other words, if we apply elementary row operations on [A B] then the solution set does not change.

The following result tells us under what conditions the system AX = B has a solution.

Theorem 9: The system of linear equations given by the matrix equation AX = B has a solution if $\rho(A) = \rho([A\ B])$.

Proof: AX = B represents the system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ \vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

This is the same as

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n - b_1 &= 0 \\ \vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \quad\quad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n - b_m &= 0 \end{aligned}$$

which is represented by $[A\ B]\begin{bmatrix} X \\ -1 \end{bmatrix} = \mathbf{0}$. Therefore, any solution of AX = B is also a solution of $[A\ B]\begin{bmatrix} X \\ -1 \end{bmatrix} = \mathbf{0}$, and vice versa. By Theorem 8, this system has a solution if and only if $n + 1 > \rho([A\ B])$.

Now, if the equation $[A\ B]\begin{bmatrix} X \\ -1 \end{bmatrix} = \mathbf{0}$ has a solution, say $\begin{bmatrix} c_1 \\ \vdots \\ c_{n-1} \end{bmatrix}$, then $c_1C_1 + c_2C_2 + \dots + c_nC_n = B$, where $C_1, \dots, C_n$ are the columns of A. That is, B is a linear combination of the C_i's. $\therefore$ RS ([A B]) = RS (A). $\therefore$ $\rho(A) = \rho([A\ B])$.

Conversely, if $\rho(A) = \rho([A\ B])$, then the number of linearly independent columns of A and [A B] are the same. Therefore, B must be a linear combination of the columns $C_1, \dots, C_n$ of A.

Let $B = a_1C_1 + \dots + a_nC_n$, $a_i \in \mathbf{F}\ \forall\ i$.

Then a solution of AX = B is $X = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$

Thus, AX = B has a solution if and only if $\rho(A) = \rho([A\ B])$.

Remark: If A is invertible then the system AX = B has the unique solution $X = A^{-1}B$.

Now, once we know that the system given by AX = B is consistent, how do we find a solution? We utilise the method of successive (or Gaussian) elimination. This method is attributed to the famous German mathematician, Carl Friedrich Gauss (1777–1855) (see Fig. 1). Gauss was called the "prince of mathematicians" by his contemporaries. He did a great amount of work in pure mathematics as well as in the probability theory of errors, geodesy, mechanics, electro-magnetism and optics.

A system of equations is called **consistent** if it has a solution.

To apply the method of Gaussian elimination, we first reduce [A B] to its row echelon form, E. Then, we write out the equations $E\begin{bmatrix} X \\ -1 \end{bmatrix} = \mathbf{0}$ and solve them, which is simple.

Let us illustrate the method.

Example 8: Solve the following system by using the Gaussian elimination process.

$x + 2y + 3z = 1$

$2x + 4y + z = 2$

Solution: The given system is the same as

$$\begin{bmatrix}1 & 2 & 3 & 1\\ 2 & 4 & 1 & 2\end{bmatrix}\begin{bmatrix}x\\ y\\ z\\ -1\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

We first reduce the coefficient matrix to echelon form.

$$\begin{bmatrix}1 & 2 & 3 & 1\\ 2 & 4 & 1 & 2\end{bmatrix}\overset{R_2(-2)}{\sim}\begin{bmatrix}1 & 2 & 3 & 1\\ 0 & 0 & -5 & 0\end{bmatrix}\overset{R(-1/5)}{\sim}\begin{bmatrix}1 & 2 & 3 & 1\\ 0 & 0 & 1 & 0\end{bmatrix}$$

This gives us an equivalent system of equations, namely,

$x + 2y + 3z = 1$ and $z = 0$.

These are, again, equivalent to $x = 1 - 2y$ and $z = 0$.

We get the solution in terms of a parameter. Put $y = \alpha$. Then $x = 1 - 2\alpha$, $y = \alpha$, $z = 0$ is a solution, for any scalar α. Thus, the solution set is $\{(1 - 2\alpha, \alpha, 0) \mid \alpha \in \mathbf{R}\}$.

Fig. 1: Carl Friedrich Gauss

Now let us look at an example where $B = \mathbf{0}$, that is, the system is homogeneous.

Example 9: Obtain a solution set of the simultaneous equations

$$\begin{aligned}x + 2y \quad\quad + 5t &= 0\\ 2x + y + 7z + 6t &= 0\\ 4x + 5y + 7z + 16t &= 0\end{aligned}$$

Solution: The matrix of coefficients is

$$A = \begin{bmatrix}1 & 2 & 0 & 5\\ 2 & 1 & 7 & 6\\ 4 & 5 & 7 & 16\end{bmatrix}$$

The given system is equivalent to $AX = \mathbf{0}$. A row-reduced echelon form of this matrix is

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

Then the given system is equivalent to

$$\left.\begin{aligned}x + y \quad\quad\quad + 5t &= 0\\ y - (7/3)z + (4/3)t &= 0\end{aligned}\right\} \Rightarrow \begin{aligned}x &= (-14/3)z - (7/3)t\\ y &= (7/3)z - (4/3)t\end{aligned}$$

which is the solution in terms of z and t. Thus, the solution set of the given system of equations, in terms of two parameters α and β, is

$\{((-14/3)\alpha - (7/3)\beta, (7/3)\alpha - (4/3)\beta, \alpha, \beta) \mid \alpha, \beta \in \mathbf{R}\}$.

This is a two-dimensional vector subspace of $\mathbf{R}^4$ with basis

$\{(-14/3, 7/3, 1, 0), (-7/3, -4/3, 0, 1)\}$.

For practice we give you the following exercise.

E

E16) Use the Gaussian method to obtain solution sets of the following system of equations.

$$\begin{aligned}4x_1 - 3x_2 + x_3 - 7 &= 0\\ x_1 - 2x_2 - 2x_3 - 3 &= 0\\ 3x_1 - x_2 + 2x_3 + 1 &= 0\end{aligned}$$

And now we are near the end of this unit.

8.5 SUMMARY

In this unit we covered the following points.

1) We defined the row rank, column rank and rank of a matrix, and showed that they are equal.
2) We proved that the rank of a linear transformation is equal to the rank of its matrix.
3) We defined the six elementary row and column operations.
4) We have shown you how to reduce a matrix to the row-reduced echelon form.
5) We have used the echelon form to obtain the inverse of a matrix.
6) We proved that the number of linearly independent solutions of a homogeneous system of equations given by the matrix equation $AX = \mathbf{0}$ is $n-r$, where r = rank of A, n = number of columns of A.
7) We proved that the system of linear equations given by the matrix equation $AX = B$ is consistent if and only if $\rho(A) = \rho([A\ B])$.
8) We have shown you how to solve a system of linear equations by the process of successive elimination of variables, that is, the Gaussian method.

8.6 SOLUTIONS/ANSWERS

E1) A is the $m \times n$ zero matrix $\Leftrightarrow RS(A) = \{\mathbf{0}\}. \Leftrightarrow \rho_r(A) = 0$.

E2) The column space of A is the subspace of $\mathbf{R}^2$ generated by (1,0), (0,2), (1,1). Now $\dim_{\mathbf{R}} CS(A) \le \dim_{\mathbf{R}} \mathbf{R}^2 = 2$. Also (1,0) and (0,2), are linearly independent.
$\therefore \{(1,0), (0,2)\}$ is a basis of CS(A), and $\rho_c(A) = 2$.

The row space of A is the subspace of $\mathbf{R}^3$ generated by (1,0,1) and (0,2,1). These vectors are linearly independent, and hence, form a basis of RS (A). $\therefore \rho_r(A) = 2$.

E3) The ith row of $C = AB$ is

$[c_{i1}\ c_{i2} \dots c_{ip}]$

$$= \left[\sum_{k=1}^{n} a_{ik} b_{k1}\ \ \sum_{k=1}^{n} a_{ik} b_{k2} \dots\ \ \sum_{k=1}^{n} a_{ik} b_{kp} \right]$$

$= a_{i1}[b_{11}\ b_{12} \dots b_{1p}] + a_{i2}[b_{21}\ b_{22} \dots b_{2p}] + \dots + a_{in}[b_{n1}\ b_{n2} \dots b_{np}]$, a linear combination of the rows of B. $\therefore RS(AB) \subseteq RS(B)$. $\therefore \rho_r(AB) \le \rho_r(B)$.

E4) By Lemma 1, $\rho(AB) \le \rho_c(A) = \rho(A)$
Also $\rho(AB) \le \rho_r(B) = \rho(B)$.
$\therefore \rho(AB) \le \min(\rho(A), \rho(B))$.

E5) $\rho(CR) \le \min(\rho(C), \rho(R))$
But $\rho(C) \le \min(m, 1) = 1$. Also $C \ne \mathbf{0}$. $\therefore \rho(C) = 1$. $\therefore \rho(CR) \le 1$.

Now, if $C = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$ $R = [b_1, \dots\dots, b_n]$, then

$$CR = \begin{bmatrix} a_1b_1 & a_1b_2 & \dots & a_1b_n \\ a_2b_1 & a_2b_2 & \dots & a_2b_n \\ \vdots & \vdots & & \vdots \\ a_mb_1 & a_mb_2 & \dots & a_mb_n \end{bmatrix}$$

Since $C \neq \mathbf{0}$, $a_i \neq 0$, for some i. Similarly, $b_j \neq 0$, for some j. $\therefore a_i b_j \neq 0$. $\therefore CR \neq \mathbf{0}$.
$\therefore \rho(CR) \neq 0$. $\therefore \rho(CR) = 1$.

E6) $PAQ = \begin{bmatrix} 0 & -2 & -2 \\ -3 & -4 & -3 \end{bmatrix}$. The rows of PAQ are linearly **independent.** $\therefore \rho(PAQ) = 2$. Also the rows of A are linearly independent. $\therefore \rho(A) = 2$. $\therefore \rho(PAQ) = \rho(A)$.

E7) Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. Then $\rho(A) = 3$. $\therefore$ A's normal form is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

E8) a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

b) $R_{32} \circ R_{21}(A) = R_{32}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 0 + 0\times(-1) & 0 + 1\times(-1) & 1 + 0\times(-1) \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

E9) $R_{23}(I_4) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = C_{23}(I_4)$

$R_2(2)(I_4) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = C_2(2)(I_4)$

$R_{12}(3)(I_4) = \begin{bmatrix} 1 & 3 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = C_{21}(3)(I_4)$

E10) $E_i(a^{-1}) E_i(a) = R_i(a^{-1})(E_i(a)) = R_i(a^{-1}) R_i(a)(I) = I$.
This proves (b).
$E_{ij}(-a) E_{ij}(a) = R_{ij}(-a)(E_{ij}(a)) = R_{ij}(-a)(R_{ij}(a)(I)) = I$, providing (c).

E11) $E_{13}(-2) E_{13}(2) = \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

E12) $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 3 & 1 & 0 \end{bmatrix} \underset{\sim}{R_{31}(-3)} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & -5 & 0 \end{bmatrix} \underset{\sim}{R_{32}(5)} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

E13) $\begin{bmatrix} 1 & 2 & 0 & 5 \\ 2 & 1 & 7 & 6 \\ 4 & 5 & 7 & 10 \end{bmatrix} \underset{\sim}{R_{21}(-2),\ R_{31}(-4)} \begin{bmatrix} 1 & 2 & 0 & 5 \\ 0 & -3 & 7 & -4 \\ 0 & -3 & 7 & -10 \end{bmatrix}$

$R_2(-1/3) \begin{bmatrix} 1 & 2 & 0 & 5 \\ 0 & 1 & -7/13 & 4/3 \\ 0 & -3 & 7 & -10 \end{bmatrix}$

$$\underset{\sim}{R_{32}(3)} \begin{bmatrix} 1 & 2 & 0 & 5 \\ 0 & 1 & -7/3 & 4/3 \\ 0 & 0 & 0 & -6 \end{bmatrix} \quad \underset{\sim}{R_3(-1/6)} \begin{bmatrix} 1 & 2 & 0 & 5 \\ 0 & 1 & -7/3 & 4/3 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

$\therefore \rho(A) = 3$

E14) $$A = \begin{bmatrix} 0 & 1 & 3 \\ 2 & 3 & 5 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A$$

$$\Rightarrow \begin{bmatrix} 2 & 3 & 5 \\ 0 & 1 & 3 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \text{ (applying } R_{12})$$

$$\Rightarrow \begin{bmatrix} 1 & 3/2 & 5/2 \\ 0 & 1 & 3 \\ 0 & 1/2 & -1/2 \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 0 \\ 1 & 0 & 0 \\ 0 & -3/2 & 1 \end{bmatrix} A \text{ (applying } R_1(1/2), R_{31}(-3))$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} -3/2 & 1/2 & 0 \\ 1 & 0 & 0 \\ -1/2 & -3/2 & 1 \end{bmatrix} A \text{ (applying } R_{12}(-3/2), R_{32}(-1/2))$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 2 & -1 \\ 1/4 & -9/4 & 3/2 \\ 1/4 & 3/4 & -1/2 \end{bmatrix} A \text{ (applying } R_3(-1/2), R_{23}(-3) \text{ and } R_{13}(2))$$

$$\therefore A \text{ is invertible, and } A^{-1} = \begin{bmatrix} -1 & 2 & -1 \\ 1/4 & -9/4 & 3/2 \\ 1/4 & 3/4 & -1/2 \end{bmatrix}$$

E15) The given system is equivalent to

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Now, the rank of $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \end{bmatrix}$ is 2. $\therefore$, the number of linearly independent solutions is $3 - 2 = 1$. $\therefore$, any non-zero solution will be a linearly independent solution. Now, the given equations are equivalent to

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

$2x + 4y = -z$ ..... (2)

(–3) times Equation (2) added to Equation (1) gives $-5x - 10y = 0$.

$\therefore x = -2y$. Then (1) gives $z = 0$. Thus, a solution is $(-2, 1, 0)$, $\therefore$, a set of linearly independent solutions is $\{(-2, 1, 0)\}$.

Note that you can get several answers to this exercise. But any solution will be $\alpha(-2, 1, 0)$, for some $\alpha \in \mathbf{R}$.

E16) The augmented matrix is [A B] .

$$= \begin{bmatrix} 4 & -3 & 1 & 7 \\ 1 & -2 & -2 & 3 \\ 3 & -1 & 2 & -1 \end{bmatrix}. \text{ Its row-reduced echelon form is}$$

$$\begin{bmatrix} 1 & -2 & -2 & 3 \\ 0 & 1 & 9/5 & -1 \\ 0 & 0 & 1 & 5 \end{bmatrix}$$

Thus, the given system of equations is equivalent to

$$\begin{aligned} x_1 - 2x_2 - 2x_3 &= 3 \\ x_2 + (9/5)x_3 &= -1 \\ x_3 &= 5. \end{aligned}$$

We can solve this system to get the unique solution $x_1 = -7$, $x_2 = -10$, $x_3 = 5$.