General Introduction and Objectives

We shall see that there is a one-to-one correspondence between the set of matrices $M_{m,n}(F)$ and the set of linear transformations $\mathcal{L}(V_n(F), V_m(F))$, between vector spaces. We shall therefore study in details the algebra of matrices together with their properties of determinant and rank. The correspondence will then be applied to linear transformations in the study of their kernel, image, rank, nullity, eigenvalues and eigenvectors. We shall then apply the theory of matrices and linear transformations to solve any system of m linear equations in n unknowns.

To objectives of the course are as follows. The reader should be able to

(i) state and prove the properties of the algebra of matrices, their determinants and ranks.

(ii) state and prove properties of vector spaces, and linear transformations between vector spaces, and

(iii) solve any system of m linear equations in n unknowns.

LECTURE ONE

Algebra of Matrices

Introduction

We shall consider a generalisation of a vector to an array, called a matrix consisting of m rows and n columns of numbers from a commutative ring R with identity or a field F. If $M = M_{m,n}(F)$ denotes the set of all such matrices, we define a binary operation of addition $+$ and show that $(M, +)$ is an Abelian group. If $m = n$, we shall show that multiplication $\cdot$ can be defined and then $(M, +, \cdot)$ becomes a non-commutative ring with identity.

Objectives

The reader should be able to

(i) show that the set of square matrices of a given order forms a non-commutative ring with identity,

(ii) do computations of addition and multiplication on matrices.

Pre-Test

1. Calculate $f(A)$ if

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \quad \text{and} \quad f(x) = 1 + 3x + x^2.$$

2. Calculate A^2 and A^3 for the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 3 & 1 & 1 \end{pmatrix}$$

3. Given that

$$\begin{pmatrix} 4 & 2a & 3 \\ \\ -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 5 \\ 3 & 0 \\ -6 & 3b \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 5 & -5 \end{pmatrix}$$

find a and b.

4. (a) Let

$$A = \begin{pmatrix} 0 & p & q \\ 0 & 0 & r \\ 0 & 0 & 0 \end{pmatrix}$$

Find A^3 and A^{29}.
(b) If

$$B = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix},$$

prove that

$$B^n = \begin{pmatrix} 2n+1 & 2n \\ -2n & -2n+1 \end{pmatrix} \quad \text{for all } \; n = 1, 2, 3, \ldots$$

5. If $A = (a_{ij})_{m,n}$ is an $(m \times n)$-matrix over a field F and k is a scalar in F, show that

(i) $(kA)^T = k \cdot A^T$
(ii) $(A + B)^T = A^T + B^T$
(iii) $(A \cdot B)^T = B^T \cdot A^T$

6. Show that the following set of 2×2 matrices forms a group under matrix multiplication. Find all its subgroups

(a) $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$

(b) $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}, \right\}$

7. (a) Show that the set of 2×2 matrices

$$\left\{ \begin{bmatrix} x & x \\ x & x \end{bmatrix} : x \in \mathbb{R}, \quad x \neq 0 \right\}$$

forms a group under matrix multiplication.
(b) Determine whether the set of 2×2 matrices

$$\left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} : x \in \mathbb{R} \right\}$$

forms a group under matrix multiplication.

8. Show that each of the following sets of 2×2 matrices forms an Abelian group under multiplication.

(a) $\left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} : a, b \in \mathbb{R}^* \right\}$

(b) $\left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} : a, b \in \mathbb{R}, \ a^2 + b^2 = 1 \right\}$

9. (a) Show that the set

$$S = \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \middle| a, b \in \mathbb{Z} \right\}$$

is a subring and a right ideal but not a left ideal in the ring $(M_2(\mathbb{Z}), +, \cdot)$
(b) Show that the set

$$W = \left\{ \begin{pmatrix} x & y \\ -y & x \end{pmatrix} \middle| x, y \in \mathbb{R} \right\}$$

is a subring (indeed a field) of the ring $(M_2(\mathbb{R}), +, \cdot)$.

10. Verify whether the mapping is a ring homomorphism

$$\phi : \mathbb{Z} \to M_2(\mathbb{Z}), \quad \phi(a) = \begin{pmatrix} a & a \\ a & a \end{pmatrix}$$

Definitions

A matrix is a rectangular array of numbers from a commutative ring with identity (e.g. a field) enclosed within curved or square brackets. A general matrix is of the form.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{or} \quad \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

We use capital letters of the alphabet to denote matrices. If A is a matrix having m rows and n columns, we say that A is an $(m \times n)$-matrix (read as 'm by n matrix'). The *order* of matrix A is then $m \times n$. Each number in a matrix is called an *entry*. One can express an entry by specifying in which row and column it is located. Thus a_{ij} is the entry located in the ith row and jth column.

A is sometimes written in a shortened form as

$$A = (a_{ij})_{m,n} \text{ or } [a_{ij}]_{m,n}$$

or as $A = (a_{ij})$ or $[a_{ij}]$ when the order of A is understood.

Examples of matrices are:

(a) When $m = 1$, i.e. when the number of rows is 1, a $1 \times n$ matrix is of the form.

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \quad \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$$

and is called a *row matrix* or *row vector*.

(b) When $n = 1$, i.e. when the number of columns is 1, and $m \times 1$ matrix is of the form.

$$\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \quad \text{or} \quad \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$$

and is called a *column matrix* or *column vector.*

(c) When $m = n$, i.e. when the number of rows = the number of columns $= n$, then an $n \times n$ matrix is called a *square matrix of order n.*

(d) A $m \times n$ matrix whose entries are all equal to zero is called the *zero matrix of order* $m \times n$ or *the null matrix of order* $m \times n$.

(e) Let A be an $m \times n$ matrix, then the *transpose* of A denoted by A' or A^T is an $n \times m$ matrix such that the entry a_{ij} of A^T is equal to the entry a_{ji} of A for all i and j. For example, if

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

then

$$A^T = A' = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{pmatrix}$$

(f) If $A = A^T$, then we say that matrix A is *symmetric.*
If $A = -A^T$, i.e. $a_{ij} = -a_{ji}$ for all i, j, then A is said to be *skew-symmetric.*

(g) A square matrix of order n such that all the entries outside the leading diagonal are all zero is called a *diagonal matrix* of order n. For example,

$$\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}$$

is a diagonal matrix of order 3.

(h) An *identity matrix* is a diagonal matrix such that all the entries in the leading diagonal are all equal to 1 and it is denoted by I or I_n.

Thus,

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Equality of Matrices
Two matrices $A = (a_{ij})$ and $B = (b_{ij})$ are said to be equal $A = B$ if they are of the same order (i.e. they have equal rows and equal columns) and if all the corresponding entries are equal (i.e. if $a_{ij} = b_{ij}$ for all i, j).

Addition of Matrices
Let $M_{m,n}(R)$ denote the set of all $m \times n$ matrices with entries from a ring R. Then if $A = (a_{ij})$ and $B = (b_{ij})$ are members of $M_{m,n}(R)$, we define the *addition* or *sum* of A and B as a matrix $C = (c_{ij})$ such that

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

Thus, $C = A + B \in M_{m,n}(R)$.

Proposition 1
$(M_{m,n}(R), +)$ is an Abelian group.

Proof: Put $M = M_{m,n}(R)$

Closure: $A, B \in M \Rightarrow A + B \in M$ (by definition).

Commutativity: $A = (a_{ij})$, $B = (b_{ij})$ in M.

$$\Rightarrow \quad A + B = (a_{ij} + b_{ij})$$

and $B + A = (b_{ij} + a_{ij})$. Since in R

$$a_{ij} + b_{ij} = b_{ij} + a_{ij}$$

for all i, j it follows that M is commutative under $+$.

Associativity: Let $A = (a_{ij})$, $B = (b_{ij})$ and $C = (c_{ij})$ be in M, then

$$\begin{aligned} (A+B)+C &= (a_{ij}+b_{ij})+(c_{ij}) = ((a_{ij}+b_{ij})+c_{ij}) \\ A+(B+C) &= (a_{ij})+(b_{ij}+c_{ij}) = (a_{ij}+(b_{ij}+c_{ij})) \end{aligned}$$

Since in R, $(a_{ij}+b_{ij})+c_{ij} = a_{ij}+(b_{ij}+c_{ij})$ for all i, j and so $(A+B)+C = A+(B+C)$. Hence the set M is associative under $+$.

Identity Element: Let $A = (a_{ij})$ be in M.
Consider 0, the zero matrix of order $m \times n$, i.e. $0 = (0)$.
Then

$$A + 0 = (a_{ij}) + (0) = (a_{ij} + 0) = (a_{ij}) = A.$$

Hence the zero matrix of order $m \times n$ is an identity matrix of $(M, +)$.

Inverse law: Let $A = (a_{ij})$ be in M. Consider $B = (-a_{ij})$. B is in M and

$$A + B = (a_{ij}) + (-a_{ij}) = (a_{ij} - a_{ij}) = (0) = 0$$

Hence every element A in M has an additive inverse in M which we denote by $-A$.

Hence $(M_{m,n}(R), +)$ is an Abelian group.

Multiplication by a Scalar
If $A = (a_{ij}) \in M_{m,n}(F)$ and k is a scalar (i.e. a member of F), we define multiplication of A by k as a matrix whose entries are equal to k times the corresponding entries of A and denote it by kA, i.e.

$$kA = k(a_{ij}) = (ka_{ij})$$

It is easy to check that the following properties are satisfied. If A and B are matrices of the same order and if k and l are scalars, then

$$\begin{aligned} k(A+B) &= kA + kB \\ (k+l)A &= kA + lA \\ (kl)A &= k(lA) = l(kA) \end{aligned}$$

Multiplication of Matrices

If A and B are two matrices, we shall define a multiplication AB of A by B only if the number of columns of A is equal to the number of rows of B. Therefore, if A is an $l \times m$ matrix, then B must be an $m \times n$ matrix. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \\ a_{i1} & a_{i2} & & a_{im} \\ \vdots & & & \\ a_{l1} & a_{l2} & \cdots & a_{lm} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{1j} & \cdots & b_{1n} \\ \\ b_{21} & b_{2j} & & a_{2n} \\ \vdots & & & \\ b_{m1} & b_{mj} & \cdots & b_{mn} \end{bmatrix}$$

then we define the product AB as an $l \times n$ matrix (c_{ij}) such that c_{ij} is the dot (or scalar) product of the column vectors.

$$\begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{im} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{mj} \end{pmatrix} \begin{aligned} &= a_{i1}b_{1j} + b_{i2}b_{2j} + \cdots + a_{im}b_{mj} \\ &= \sum_{k=1}^{m} a_{ik}b_{kj} \end{aligned}$$

Thus in order to obtain the entry c_{ij} in the product AB, we look at the i-th row of A and the j-th column of B, consider them as column vectors and then take their dot or scalar product. If the product AB of matrices A and B is defined, we say that A and B are *comformable* (or *compatible*) matrices for multiplication. We they say that A is post-multiplied by B or that B is pre-multiplied by A.

Associative law for multiplication

Proposition 2

Let $A = (a_{ij})_{p,q}$, $B = (b_{jk})_{q,r}$ and $C = (c_{kl})_{r,s}$ be matrices such that the product $(AB)C$ and $A(BC)$ are defined and are $p \times s$ matrices. Then the associative law of matrix multiplication holds.

$$\text{i.e.} \quad (AB)C = A(BC).$$

Proof:

$$AB = \left(\sum_{j=1}^{q} a_{ij}b_{jk} \right)_{p,r}$$

$$
\begin{aligned}
\Rightarrow \quad (AB)C &= \left[\sum_{k=1}^{r}\left(\sum_{j=1}^{q} a_{ij}b_{jk}\right)c_{kl}\right]_{p,s} \\
&= \left(\sum_{k=1}^{r}\sum_{j=1}^{q} a_{ij}b_{jk}c_{kl}\right)_{p,s} \qquad (1) \\
BC &= \left(\sum_{k=1}^{r} b_{jk}c_{kl}\right)_{q,s} \\
\Rightarrow \quad A(BC) &= \left[\sum_{j=1}^{q} a_{ij}\left(\sum_{k=1}^{r} b_{jk}c_{kl}\right)\right]_{p,s} \\
&= \left(\sum_{j=1}^{q}\sum_{k=1}^{r} a_{ij}b_{jk}c_{kl}\right)_{p,s} \qquad (2)
\end{aligned}
$$

Since they are finite sums, the expressions (1) and (2) are equal and we have that

$$(AB)C = A(BC)$$

Distributive laws

Let A and B be matrices of the same order $q \times r$ (so that addition of A and B is well-defined). Also let C and D be matrices of orders $p \times q$ and $r \times s$, respectively (so that the products CA, CB, AD and BD are well-defined). Then we wish to show that

$$C(A+B) = CA + CB \;\text{ and }\; (A+B)D = AD + BD$$

The above equations show that multiplication of matrices satisfies the distributive laws over matrix addition whenever the products and sums are defined.

Proposition 3:

Let $A = (a_{jk})_{q,r}$, $B = (b_{jk})_{q,r}$ and $C = (c_{ij})_{p,q}$. Then $C(A+B) = CA+CB$. Similarly if $D = (d_{kp})_{r,s}$, then $(A+B)D = AD + BD$.

Proof:

$$A + B = (a_{jk} + b_{jk})_{q,r}$$

$$
\begin{aligned}
\Rightarrow \quad C(A+B) &= \left[\sum_{j=1}^{q} c_{ij}(a_{jk}+b_{jk})\right]_{p,r} \\
&= \left[\sum_{j=1}^{q} c_{ij}a_{jk} + \sum_{j=1}^{q} c_{ij}b_{jk}\right]_{p,r} \\
&= \left(\sum_{j=1}^{q} c_{ij}a_{jk}\right)_{p,r} + \left(\sum_{j=1}^{q} c_{ij}b_{jk}\right)_{p,r} \\
&= CA + CB.
\end{aligned}
$$

This proves the first distributive law. The proof of the second distributive law is similar and is left as an exercise.

Commutative law for Multiplication

The commutative law does NOT hold for matrices with respect to multiplication. First if A is an $l \times m$ matrix and B is an $m \times n$ matrix (so that the product AB is defined) it does not follow that BA is defined except $n = l$. Thus if $n \neq l$, we cannot even compare AB and BA to show whether they are equal or not since one of then BA is not defined. If $n = l$ and both AB and BA are well-defined, the two products AB and BA may not be equal because if A is an $n \times m$ matrix and B is an $m \times n$ matrix, then the product AB is an $n \times n$ matrix and the product BA is an $m \times m$ matrix. So we cannot compare AB and BA to show whether they are equal or not since they are not of the same order except if $m = n$. Even when $m = n = l$ and we have square matrices, the two products AB and BA may not be equal as shown by the following counter-example. Let

$$
A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}; \quad B = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix},
$$

then

$$
AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix} = \begin{bmatrix} 1.2 + 2.4 & 1.3 + 2.6 \\ 3.2 + 4.4 & 3.3 + 4.6 \end{bmatrix} = \begin{bmatrix} 10 & 15 \\ 22 & 33 \end{bmatrix}
$$

$$
BA = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2.1 + 3.3 & 2.2 + 3.4 \\ 4.1 + 6.3 & 4.2 + 6.4 \end{bmatrix} = \begin{bmatrix} 11 & 16 \\ 22 & 32 \end{bmatrix}
$$

Thus $AB \neq BA$ and the commutative law does not hold for matrices with respect to multiplication. Therefore, the order in which we multiply two matrices is very important. If for two matrices we have $AB = BA$, then we say that A and B commute.

Square Matrices
We shall denote by M_n, the set of all square matrices of order n.

Proposition 4
$(M_n, +, \cdot)$ is a non-commutative ring with identity.

Proof: Proposition 1 shows that $(M_n, +)$ is an Abelian group. Consider $(M_n, \cdot)$. If $A \in M_n$, $B \in M_n$, then $AB \in M_n$. Hence $(M_n, \cdot)$ is closed. By proposition 2, $(M_n, \cdot)$ is associative and so a semigroup. Since the distributive laws hold from Proposition 3, it follows that $(M_n, +, \cdot)$ is a ring.

Identity Element: First define the Kronecker delta, δ_{ij} as follows:

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

Using this symbol we define an $n \times n$ matrix 1_n as $I_n = (\delta_{ij})_{n,n}$. This matrix has only 1's on its leading diagonal entries and 0's off the leading diagonal. We show that if $A = (a_{ij}) \in M_n$, then $AI_n = A = I_nA$ as follows:

$$\begin{aligned} AI_n &= \left(\sum_{j=1}^{n} a_{ij}\delta_{jk}\right)_{n,n} = (a_{ij})_{n,n} = A, \quad \text{and} \\ I_nA &= \left(\sum_{j=1}^{n} \delta_{ij}a_{jk}\right)_{n,n} = (a_{ik})_{n,n} = A. \end{aligned}$$

Therefore I_n is the identity matrix in $(M_n, \cdot)$. Hence $(M_n, +, \cdot)$ is a non-commutative ring with identity.

Example 1
If $A = \begin{pmatrix} 2 & 4 \\ -2 & 2 \end{pmatrix}$ such that $A^2 + sA + tI = 0$, find s and t.

$$A^2 = \begin{pmatrix} 2 & 4 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 2 & 4 \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} -4 & 16 \\ -8 & -4 \end{pmatrix}$$

$$\begin{aligned} \Rightarrow \quad A^2 + sA + tI &= \begin{pmatrix} -4 & 16 \\ -8 & -4 \end{pmatrix} + \begin{pmatrix} 2s & 4s \\ -2s & 2s \end{pmatrix} + \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} \\ &= \begin{pmatrix} -4+2s+t & 16+4s \\ -8-2s & -4+2s+t \end{pmatrix} \end{aligned}$$

$A^2 + sA + tI = 0$ implies

$$\begin{matrix} -4+2s+t=0 & 16+4s=0 \\ -8-2s=0 & -4+2s+t=0 \end{matrix}$$

$\Rightarrow \quad s = -4,\ t = 12.$

Example 2
If A is any square matrix of order 2, show that the matrix AA' is symmetric.
Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Then

$$A' = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

and

$$\begin{aligned} AA' &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix} \\ (AA')' &= \begin{pmatrix} a^2+b^2 & ac+bb \\ ac+bd & c^2+d^2 \end{pmatrix} = AA' \end{aligned}$$

$\Rightarrow \quad AA'$ is symmetric.

Practice Exercises 1

1. Decide whether the products AB and BA are defined. If the answer is yes, compute the products.

   (i) $A = \begin{pmatrix} 3 & 7 & -10 \\ 4 & 5 & 8 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 3 \\ 0 & 8 \end{pmatrix}$

   (ii) $A = \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix}^T$, $B = (5 \quad -4 \quad -3)$

   (iii) $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}$

   (iv) $A = (10 \quad -5 \quad 2)$, $B = (0 \quad 2 \quad 7)^T$.

2. Multiply

   (a) $$\begin{pmatrix} 2 & 5 & 1 & 6 \\ 8 & 1 & -2 & 3 \\ 8 & 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & -6 \\ 1 & 2 \\ 8 & 1 \\ 6 & 0 \end{pmatrix}$$

   (b) $$\begin{pmatrix} 5 & 2 & 1 & 3 \\ 8 & 6 & 1 & 2 \\ 0 & 1 & 5 & 3 \end{pmatrix} \begin{pmatrix} 1 & 4 & 8 \\ 2 & 0 & 1 \\ -1 & 1 & 0 \\ 5 & 8 & 6 \end{pmatrix}$$

   (c) $$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 15 & 2 & 1 & 3 \\ 8 & 6 & 1 & 2 \\ 0 & 1 & 5 & 8 \end{pmatrix}$$

   (d) $$\begin{pmatrix} 5 & 2 & 1 & 3 \\ 8 & 6 & 1 & 2 \\ 0 & 1 & 5 & 8 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Summary
Different types of matrices are defined. These are row and column matrices (or vectors), square matrices, null matrix, transpose of a matrix, symmetric and skew-symmetric matrices, diagonal and identity matrices.

Addition and multiplication are then defined and their properties examined. In particular we show that $(M_{m,n}(F), +)$ is an Abelian group while $(M_n(F), +, \cdot)$ is a non-commutative ring with identity

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.

2. Ayres, F. *Matrices*, Schaum's Outline Series.

3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.

4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

5. Kuku, A.O. *Abstract Algebra,* Ibadan University Press 1980. Chapter 5.

LECTURE TWO

Determinants

Introduction

Given a square matrix, we shall define a scalar called the determinant of the matrix. We shall study the relationship between determinants and products, transpose and scalar multiplication of matrices. We shall then consider other properties of determinants which enable us to evaluate them easily.

Objectives

The reader should be able to

(i) state and prove the properties of determinants; and

(ii) use the properties of determinants to facilitate their evaluation.

Pre-Test

1. If A and B are any square matrices in $M_n(F)$ such that adj = adjoint, show that

   (i) $adj(AB) = adj(B) \cdot adj(A)$

   (ii) $adj(adjA) = |A|^{n-2}A$

2. Evaluate the determinant

$$\begin{vmatrix} 17 & 19 & 18 \\ 5 & 6 & 6 \\ 2 & 2 & 2 \end{vmatrix}$$

3. Factorise

$$\begin{vmatrix} 1 & 1 & 1 \\ ab & bc & ca \\ a^2b^2 & b^2c^2 & c^2a^2 \end{vmatrix}$$

4. Show that

$$\begin{vmatrix} n! & (n+1)! & (n+2)! \\ (n+1)! & (n+2)! & (n+3)! \\ (n+2)! & (n+3)! & (n+4)! \end{vmatrix} = n!(n+1)!(n+2)!2$$

5. Show by column operations that $(a+b+c)$ is factor of

$$\begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix}$$

and find the other factors.

6. By considering the matrix product

$$\begin{bmatrix} 1 & 0 & a \\ 0 & 1 & 0 \\ a & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & b & 0 \\ b & 1 & b \\ 0 & b & 1 \end{bmatrix}$$

show that

$$\begin{vmatrix} 1 & (a+1)b & a \\ b & 1 & b \\ a & (a+1)b & 1 \end{vmatrix} = (1-a^2)(1-2b)$$

7. Show that

$$\begin{vmatrix} 0 & x & y & z \\ -x & 0 & u & v \\ -y & -u & 0 & w \\ -z & -v & -w & 0 \end{vmatrix} = (xy - yv + uz)^2$$

8. Solve for x

$$\begin{vmatrix} x-1 & 4 & -1 \\ 1 & x+2 & 1 \\ 2x-4 & 4 & x-4 \end{vmatrix} = 0$$

9. Factorise

$$\left[\begin{array}{ccc} a^3 & b^3 & c^2 \\ a^2 & b^2 & c^2 \\ b+c & c+a & a+b \end{array} \right|$$

10. Show that

$$\begin{vmatrix} \log x & \log y & \log z \\ \log 2x & \log 2y & \log 2z \\ \log 3x & \log 3y & log 3z \end{vmatrix} = 0$$

where x, y, z are positive real numbers.

Definition of the determinant
Let S_n be the symmetric group on n letters.
Then

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix} \in S_n$$

is *even* or *odd* depending on whether there exists an even or odd number of pairs $(\sigma(k), \sigma(l))$ where $\sigma(k) > \sigma(l)$ but $\sigma(k)$ precedes $\sigma(l)$ in the second row of σ.
The *sign or parity* of σ is then defined by

$$sgn\sigma = \begin{cases} 1, & \text{if } \sigma \text{ is even} \\ \\ -1, & \text{if } \sigma \text{ is odd} \end{cases}$$

Let $M_n(F)$ denote the set of all square matrices of order n with entries from a field F. We define a function, called the *determinant*

$$\det : M_n(F) \to F \;\text{ or }\; \triangle : M_n(F) \to F \;\text{ or }\; |\cdot| : M_n(F) \to F$$

as

$$\det A = \sum_{\sigma \in S_n} (sgn\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

where $A = (a_{ij}) \in M_n(F)$.

Example 1

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{array}{l} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \\ -a_{13}a_{22}a_{31} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \end{array}$$

with each term corresponding to the following members of S_3:
$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$,
$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$.

Properties of the determinant

1. $\det A = \det(A^T)$ where A^T is the transpose of A.

   Proof: $A = (a_{ij}) \in M_n \Rightarrow |A| = \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$
   Since S_n is a group, it follows that

   $$a_{1\sigma(1)} \cdots a_{n\sigma(n)} = a_{\sigma^{-1}(1)1} \cdots a_{\sigma^{-1}(n)n}$$

   $$\Rightarrow |A| = \sum_{\sigma \in S_n} sgn(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} = |A^T|$$

   since $sgn(\sigma) = sgn(\sigma^{-1})$.

2. $|AB| = |A| \cdot |B|$.

   Proof: Denote the columns of A and B by

   $$A^{(1)}, \cdots, A^{(n)} \text{ and } B^{(1)}, \cdots, B^{(n)}$$

   respectively. Then

   $$\det(AB) = \det(AB^{(1)}, \cdots, AB^{(n)})$$

$$
\begin{aligned}
&= \det\left(\sum_{\sigma\in S_n} A^{(\sigma(1))} b_{\sigma(1)1}, \ldots, \sum_{\sigma\in S_n} A^{(\sigma(n))} b_{\sigma(n)n}\right) \\
&= \det A \sum_{\sigma\in S_n} (sgn\sigma) b_{\sigma(1)1} \cdots b_{\sigma(n)n} \\
&= \det A \cdot \det B.
\end{aligned}
$$

3. *The Laplace expansion of the determinant*
   If $A = (a_{ij}) \in M_n$, then we define for each entry a_{ij}, the *minor* for a_{ij} as the determinant of the $(n-1)\times(n-1)$-matrix M_{ij} obtained from A by deleting the i-th row and j-th column of A. We then define the cofactor of a_{ij} as $A_{ij} = (-1)^{i+j}|M_{ij}|$. Then the Laplace expansion of $\det A$ is given y

   $$|A| = \sum_{j=1}^{n} a_{ij} A_{ij} = \sum_{i=1}^{n} a_{ij} A_{ij}$$

   In other words the determinant can be obtained as a linear combination of the cofactors of either any row with coefficients from the same row or any column with coefficients from the same column.

   Proof: From the definition of $\det A$, we can write

   $$|A| = \sum_{j=1}^{n} a_{ij} \bar{A}_{ij}$$

   where $\bar{A}_{ij}$ is a sum of terms not involving any element of the i-th row of A. We must show that $\bar{A}_{ij} = A_{ij}$. Now $a_{nn}\bar{A}_{nn} = a_{nn}\sum(sgn\sigma)a_{1\sigma(n)}\cdots a_{n-1\sigma(n-1)}$ is the sum of terms of $|A|$ containing a_{nn}, where the summation is over all σ in S_n such that $\sigma(n) = n$. This implies that

   $$\bar{A}_{nn} = |M_{nn}| = A_{nn}$$

   Now let A' be the matrix obtained from A by interchanging the i-th and $(i+1)$-th rows. Then

   $$|A| = \sum_{\sigma\in S_n} (sgn\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

$$
\begin{aligned}
&= -\sum_{\sigma\in S_n}(sgn\sigma)a_{1\sigma(1)}\cdots a_{i-1\sigma(i-1)}a_{i+1\sigma(i+1)}a_{i\sigma(i)}\cdots a_{n\sigma(n)} \\
&\quad \text{(since either } \sigma(i-1)<\sigma(i) \text{ or } \sigma(i-1)>\sigma(i)) \\
&= -|A'|
\end{aligned}
$$

Similarly, $|A| = -|A''|$ where $|A''|$ is the matrix obtained from A by interchanging the j-th and $(j+1)$-the columns.

Next consider $\bar{A}_{ij}$ for some fixed i and$j$. interchange in A the i-th row and the $(i+1)$-th row, then interchange the new $(i+1)$-th row and the $(i+2)$-th row, etc. until the n-th row is reached. Similarly, interchange the j-th column and the $(j+1)$-th column, etc. until the n-th column is reached. Then

$$\bar{A}_{ij} = (-1)^{n-i+n-j}|M_{ij}| = (-1)^{i+j}|M_{i+j}| = A_{ij}$$

where $(-1)^{n-i}$ represents the change in sign by interchanging the i-th and n-th rows and $(-1)^{n-j}$ represents the change in sign by interchanging the j-th and n-th columns. Hence

$$|A| = \sum_{j\in 1}^{n} a_{ij}A_{ij}$$

From property (1) above, we have

$$|A| = \sum_{i=1}^{n} a_{ij}A_{ij}.$$

4. *Interchanging of two rows*
   Suppose two row of a determinant are interchanged. The effect on the determinant is that the sign of the determinant is changed.

   Proof: First we show that the property is true when succeeding rows are interchanged. Let $A = (a_{ik}) \in M_n$ such that A' is the matrix obtained by interchanging the i-th and $(i+1)$-th rows. Then

$$|A| = \sum_{\sigma\in S_n}(sgn\sigma)a_{1\sigma(1)}\cdots a_{n\sigma(n)}$$

$$
\begin{aligned}
&= -\sum_{\sigma\in S_n}(sgn\sigma)a_{1\sigma(1)}\cdots a_{i-1\sigma(i-1)}a_{i+1\sigma(i+1)}a_{i\sigma(i)}\cdots a_{n\sigma(n)} \\
&= -|A'|
\end{aligned}
$$

Since either $\sigma(i-1) < \sigma(i)$ or $\sigma(i-1) > \sigma(i)$.
Now assume that the i-th and j-th rows are interchanged. Let $A^{(i)}$ denote the i-th row of A. Then assume $i < j$ and put

$$
\begin{aligned}
A &= (A^{(1)}, \ldots A^{(i)}, \ldots, A^{(j)}, \ldots, A^{(n)})^T \\
A' &= (A^{(1)}, \ldots A^{(j)}, \ldots, A^{(i)}, \ldots, A^{(n)})^T \\
A'' &= (A^{(1)}, \ldots \hat{A}^{(i)}, A^{(j)}, A^{(i)}, \ldots, A^{(n)})^T
\end{aligned}
$$

where A'' is obtained from A' by successively interchanging the row $A^{(j)}$ in the i-th row of A' and its succeeding row until the row $A^{(j-1)}$ is reached. Then from the first part of this proof

$$
\begin{aligned}
|A'' &= (-1)^{j-1-i}|A^i|. \text{ Also} \\
|A''| &= \sum_i (-1)^{j+k} a_{ik}|M_{ik}|, \quad \text{(Laplace expansion)} \\
\Rightarrow \quad |A'| &= \sum_i (-1)^{j+k-j+1+i} a_{ik}|M_{ik}| \\
&= \sum_i (-1)^{k+i+1} a_{ik}|M_{ik}| \\
&= -\sum_i (-1)^{k+i}|M_{ik}| \\
&= -|A|
\end{aligned}
$$

5. *Two identical rows*
If two rows are identical in a matrix, the determinant is equal to 0.

Proof: Interchange the identical rows to obtain matrices 4 and A'.

$$A = A' \Rightarrow |A| = |A'|$$

Also by property 4 above

$$|A'| = -|A| \Rightarrow 2|A| = 0 \Rightarrow |A| = 0$$

6. *Addition to one row of a multiple of another row*
   The determinant remains unchanged when we add to one row, a multiple of another row.

   Proof: Let $A = (a_{ik}) \in M_n$ be such that
$$A = (A^{(1)}, \dots, A^{(n)})^T.$$
   Consider the matrix
$$\begin{aligned}
A' &= (A^{(1)}, \dots, A^{(i-1)}, A^{(i)} + \lambda A^{(j)}, A^{(i+1)}, \dots, A^{(n)})^T \\
|A'| &= \sum_k (a_{ik} + \lambda a_{jk}) A_{ik} \\
&= \sum_k a_{ik} A_{ik} + \lambda \sum_k a_{ij} A_{ik} \\
&= |A| + \lambda |(A^{(1)}, \dots, A^{(i-1)}, A^{(j)}, A^{(i+1)}, \dots, A^{(j)}, \dots, A^{(n)})^T| \\
&= |A| + \lambda \cdot 0 \qquad \text{(by property (5))} \\
&= |A|.
\end{aligned}$$

7. *Multiplication of the elements of one row by a non-zero factor λ*
   If we multiply the elements of one row by a non-zero factor λ the determinant is also multiplied by the factor.

   Proof: Let $A = (A_{ik}) \in M_n$ be such that
$$A = (A^{(1)}, \dots, A^{(n)})^T.$$
   Consider the matrix
$$A' = (A^{(1)}, \dots, A^{(i-1)}, \lambda A^{(i)}, A^{(i+1)}, \dots, A^{(n)})^T.$$
   Then
$$|A'| = \sum_j \lambda a_{ij} A_{ij} = \lambda \sum a_{ij} A_{ij} = \lambda |A|.$$

8. *Scalar multiplication of a matrix*

$$|\alpha A| = \alpha^n |A|$$

Proof: By property (7), we have if

$$A = (A^{(1)}, \ldots, A^{(n)})^T$$

$$\begin{aligned}
\alpha^n |A| &= \alpha^{n-1} |(\alpha A^{(1)}, A^{(2)}, \ldots, A^{(n)})^T| \\
&= \alpha^{n-2} |(\alpha A^{(1)}, \alpha A^{(2)}, A^{(3)}, \ldots, A^{(n)})^T| \\
&\vdots \\
&= |(\alpha A^{(1)}, \alpha A^{(2)}, \ldots, \alpha A^{(n)})^T| \\
&= |\alpha A|.
\end{aligned}$$

9. *Determinant of a triangular matrix*
The determinant of a triangular matrix is the product of all the elements in the leading diagonal. That is if $A = (a_{ij}) \in M_n$ such that either $a_{ij} = 0$ for all $i > j$ (i.e. an upper triangular matrix) or $a_{ij} = 0$ for all $i < j$ (i.e. a lower triangular matrix), then $|A| = a_{11} \cdot a_{22} \ldots a_{nn}$.

Proof: Use induction on n. When $n = 2$.

$$\begin{vmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} = \begin{vmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{vmatrix}$$

We now assume that if $A \in M_r$, $r < n$, where A is a triangular matrix then $|A| = a_{11} a_{22} \ldots a_{rr}$. Now consider a triangular matrix $A \in M_n$. Then

$$\begin{aligned}
|A| &= \sum a_{ij} A_{ij} = a_{11} A_{11} \quad (a_{ij} = 0, j = 1) \\
&= a_{11} |M_{11}|
\end{aligned}$$

where M_{11} is a triangular matrix in M_{n-1}. By the inductive hypothesis,

$$\begin{aligned}
|M_{11}| &= a_{22} a_{33}, \ldots, a_{nn} \\
\Rightarrow \quad |A| &= a_{11} a_{22}, \ldots, a_{nn}.
\end{aligned}$$

Corollary:

1. If $D = diag(a_1, \ldots, a_n)$ is a diagonal matrix in M_n, i.e. $a_{ij} = 0$, $i \neq j$, then
$$|D| = a_1 a_2 \ldots a_n.$$
2. If $I \in M_n$ is the identity matrix, then $|I| = 1$.

10. *A matrix with a zero row*
The determinant of a matrix with a zero row is equal to 0.

Proof: Let $A = (a_{ij}) \in M_n$ such that $a_{Ij} = 0$ for a fixed I and $1 \leq j \leq n$.
Then
$$|A| = \sum_j a_{Ij} A_{Ij} = \sum_j 0 \cdot A_{Ij} = 0$$

11. $AA^* = A^*A = |A|I$ where A^* is the transpose of the matrix of cofactors of A (called the adjoint of A).

Proof: $A = (a_{ij}) \in M_n$, $A^* = (A_{jk})^T$. Then
$$AA^* = \left(\sum_j a_{ij} A_{kj} \right) = (b_{ik})$$

Now
$$b_{ii} = \sum_j a_{ij} A_{ij} = |A|$$

$$\begin{aligned} i \neq k,\ b_{ik} &= |(A^{(1)}, \ldots, A^{(i-1)}, A^{(k)}, A^{(i+1)}, \ldots, A^{(k)}, \ldots, A^{(n)})^T| \\ &= 0 \text{ (by property 5:two identical rows)} \\ \Rightarrow\ AA^* &= diag(|A|, \ldots, |A|) = |A| \cdot I \end{aligned}$$

Similarly $A^*A = |A| \cdot I$.

12. *Determinants of* 3×3 *matrices*
An interesting way of expanding a 3×3 determinant is as follows:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{array}{l} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{12}a_{23}a_{31} \\ -a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{12}a_{21}a_{33} \end{array}$$

The three positive terms are obtained by multiplying the entries in the array

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{matrix} a_{11} & a_{12} \\ a_{12} & a_{22} \\ a_{31} & a_{33} \end{matrix}$$

The three negative terms are obtained by multiplying the entries in the following array

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{matrix} a_{11} & a_{12} \\ a_{12} & a_{22} \\ a_{31} & a_{33} \end{matrix}$$

One can combine the two sets of arrows in the two arrays to obtain

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

so that one can write down the expansion straight away without resorting to writing down the first two arrays above.

13. *Column Operations*
All the properties which have been enumerated concerning rows of a determinant are also true if we replace rows by columns since by property 1, $\det(A^T) = \det(A)$.

Notations
We shall be using the following notation

R_{ij}	-	interchanging the i-th and j-th rows
$R_i(\lambda)$	=	multiplying the i-th row by a non-zero scalar λ
$R_{ij}(\lambda)$	-	adding the j-th row multiplied by a scalar λ to the i-th row.
C_{ij}	-	interchanging the ith and j-th columns
$C_i(\lambda)$	-	multiplying the i-th column by a non-zero scalar λ
$C_{ij}(\lambda)$	-	adding the j-th column multiplied by a scalar λ to the i-th column

Example 2: Evaluate the following determinant using

(i) the first row as the row of coefficients,

(ii) the second column as the column of coefficients,

(iii) the direct method.

(i) $\begin{vmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{vmatrix} = 2 \begin{vmatrix} 6 & 7 \\ 9 & 1 \end{vmatrix} - 3 \begin{vmatrix} 5 & 7 \\ 8 & 1 \end{vmatrix} + 4 \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix}$
$= 2(6 - 23) - 3(5 - 56) + 4(45 - 48) = 27$

(ii) $\begin{vmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{vmatrix} = -3 \begin{vmatrix} 5 & 7 \\ 8 & 1 \end{vmatrix} + 6 \begin{vmatrix} 2 & 4 \\ 8 & 1 \end{vmatrix} - 9 \begin{vmatrix} 2 & 4 \\ 5 & 7 \end{vmatrix}$
$= -3(5 - 56) + 6(2 - 32) - 9(14 - 20) = 27$

(iii) $\begin{vmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{vmatrix} = 2(6)1 + 3(7)8 + 5(9)4 - 4(6)8 - 3(5)1 - 7(9)2 = 27$

Example 3
Factorise the determinant

$$\begin{vmatrix} 1 & 1 & 1 \\ \lambda & \mu & \nu \\ \lambda^2 & \mu^2 & \nu^2 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 & 0 \\ \lambda & \mu-\lambda & \nu-\lambda \\ \lambda^2 & \mu^2-\lambda^2 & \nu^2-\lambda^2 \end{vmatrix} C_{21}(-1), C_{31}(-1)$$

$$= \begin{vmatrix} \mu-\lambda & \nu-\lambda \\ \mu^2-\lambda^2 & \nu^2-\lambda^2 \end{vmatrix} \quad \text{expanding using the first row.}$$

$$= (\mu-\lambda)(\nu-\lambda) \begin{vmatrix} 1 & 1 \\ \mu+\lambda & \nu+\lambda \end{vmatrix}, \; C_1\left(\frac{1}{\mu-\nu}\right), \; C_2\left(\frac{1}{\nu-\lambda}\right)$$

$$= (\mu-\lambda)(\nu-\lambda)(\nu-\mu)$$

Alternatively
Put $\lambda = \mu$. Then two columns of the determinant become identical, and the determinant becomes zero. Therefore $\lambda - \mu$ is a factor of the determinant.

Similarly, by putting $\lambda = \mu$ or $\mu = \nu$ we find that $\lambda - \nu$ and $\mu - \nu$ are also factors of the determinant.
Since the determinant is of total degree 3, we have that
$\triangle = k(\lambda-\mu)(\lambda-\nu)(\mu-\nu)$.
We then determine the constant k by considering a particular term:

Coefficient of $\mu\nu^2$: $1 = -k \Rightarrow k = -1$
Therefore

$$\begin{aligned} \triangle &= -(\lambda-\mu)(\lambda-\nu)(\mu-\nu) \\ &= (\lambda-\mu)(\mu-\nu)(\nu-\lambda) \end{aligned}$$

Example 4
Show that

$$\begin{vmatrix} bc & ca & ab \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{vmatrix}$$

$$\begin{aligned} L.H.S. &= \frac{1}{abc} \begin{vmatrix} abc & abc & abc \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{vmatrix}, \; C_1(a), C_2(b), C_3(c) \\ &= \begin{vmatrix} 1 & 1 & 1 \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{vmatrix}; \; R_1\left(\frac{1}{abc}\right) \end{aligned}$$

Practice Exercise

1. Evaluate the determinant

$$\begin{vmatrix} 18 & 16 & 2 \\ 7 & 5 & 1 \\ 2 & 1 & -1 \end{vmatrix}$$

2. Factorise

$$\begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^3 & y^3 & z^3 \end{vmatrix}$$

3. Show that

$$\begin{vmatrix} 1 & \alpha & \beta\gamma \\ 1 & \beta & \gamma\alpha \\ 1 & \gamma & \alpha\beta \end{vmatrix} = \begin{vmatrix} 1 & \alpha & \alpha^2 \\ 1 & \beta & \beta^2 \\ 1 & \gamma & \gamma^2 \end{vmatrix}$$

4. Show that

$$\begin{vmatrix} 1 & a & a^2 & a^3+bcd \\ 1 & b & b^2 & b^3+cda \\ 1 & c & c^2 & c^3+abd \\ 1 & d & d^2 & d^3+abc \end{vmatrix} = 0$$

5. Solve for x

$$\begin{vmatrix} x^3 & x & 1 \\ 8 & 2 & 1 \\ 27 & 3 & 1 \end{vmatrix} = 0$$

Summary
The determinant is defined on square matrices as a scalar using the parity of members of the symmetric group. Properties of determinants are then considered which include the Laplace expansion of the determinant and the effect on the elementary row and column operations on determinants. These properties are then applied to the evaluation and the factorisation of determinants.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.

2. Ayres, F. *Matrices*, Schaum's Outline Series.

3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.

4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

5. Kuku, A.O. *Abstract Algebra,* Ibadan University Press 1980. Chapter 5.

LECTURE THREE

Matrix Inverses and Systems of Linear Equations

Introduction

We shall see that some square matrices have multiplicative inverses while others do not have. We shall examine some elementary properties of matrix inverses and study two methods, the adjoint method, and the method of elementary operations, to determine the inverse of a non-singular matrix. We shall then apply matrix inverses to the solution of systems of linear equations which have unique solutions. The application will include Cramer's rule which is an elegant short method not involving an explicit computation of matrix inverses.

Objectives

The reader should be able to

(i) distinguish non-singular and singular matrices,

(ii) prove elementary properties of matrix inverses

(iii) compute matrix inverses, using the adjoint method or the method of elementary operations,

(iv) apply matrix inverses and Cramer's rule to the solution of systems of linear equations which have unique solutions.

Pre-Test

1. Calculate A^{-1} over $\mathbb{R}$ if

$$A = \begin{pmatrix} 4 & 3 & 2 \\ 3 & 5 & 2 \\ 2 & 2 & 1 \end{pmatrix}$$

2. (a) If $D =$ diagonal $(d_1, d_2, \ldots, d_n)$ and $d_1 d_2 \ldots d_n \neq 0$, show that D^{-1} $=$ diagonal $(d_1^{-1}, d_2^{-1}, \ldots, d_n^{-1})$.
   (b) Show that $GL_n(\mathbb{R})$ is a group under multiplication.

3. If

$$B = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{pmatrix}, \quad \text{prove that} 16B^{-2} + 6B^{-1} - 9I + B = 0$$

4. By finding the inverse of the coefficient matrix, solve over $\mathbb{R}$

$$\begin{aligned} x + y + z &= 4 \\ 2x + 5y + 2z &= 3 \\ 3x + 7y + z &= 5 \end{aligned}$$

5. Determine the values of λ such that the matrix

$$\begin{pmatrix} 4 & -2 & 5 \\ 5 & 2 & 4 \\ 2 & 8 & \lambda \end{pmatrix}$$

   has an inverse. Find the inverse when $\lambda = 1$ and hence solve the system of equations.

$$\begin{aligned} 4x - 2y + 5x &= 2 \\ 5x + 2y + 4x &= 8 \\ 2x + 8y + z &= 1 \end{aligned}$$

6. For which values of k does the system of linear equations over $\mathbb{R}$ have a unique solution?

$$\begin{aligned} 3x - 2y + kz &= 1 \\ x + 7y + z &= 2k \\ x + 4y + 2z &= 0 \end{aligned}$$

Find the solution (in terms of k) in the cases in which it is unique.

7. If A and B in $M_n(\mathbb{R})$ satisfy the equation $A + B = k \cdot AB$ where k is a non-zero constant, show that if A has an inverse, then B also has an inverse. State the inverse of B. If

$$A = \begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix},$$

determine B such that $A + B = 2 \cdot AB$.

8. If

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 3 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & -8 \\ -4 & 1 & 1 \end{pmatrix}$$

calculate BA and hence solve the system of linear equations.

$$\begin{aligned} x - 2x &= 1 \\ 3x + y + 2z &= 2 \\ x - y &= 3 \end{aligned}$$

9. Use Cramer's rule to solve the system of linear equations over $\mathbb{R}$.

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

10. Use Cramer's rule to solve the system of linear equations in x, y, z simplifying your answers as far as possible, given that a, b, c are different, $a + b + c \neq 0$ and $ab + bc + ca \neq 0$.

$$\begin{aligned} ax + by + cz &= 0 \\ (b+c)x + (c+a)y + (a+b)z &= a+b+c \\ bcx + cay + abz &= abc \end{aligned}$$

Non-singular matrices
A square matrix in $M_n(F)$ which has a multiplicative inverse is called an *invertible* or a *non-singular* matrix. A square matrix which does not have a multiplicative inverse is called a *non-invertible* or *singular* matrix. We denote by $GL_n(F)$ the set of all invertible or non-singular matrices of order n, i.e. $GL_n(F) \subset M_n(F)$. One can show that $GL_n(F)$ is a group under multiplication, called the *general linear group of order n*.

To show that not all matrices have inverses, consider in $M_2(\mathbb{R})$, the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

Suppose that it has an inverse of the form

$$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Then

$$AB = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

i.e.

$$\begin{bmatrix} a+2c & b+2d \\ 2a+4c & 2b+4d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Therefore
$a + 2c = 1$ (i)
$2a + 4c = 0$ (ii)
$b + 2d = 0$ (iii)

$2b + 4d = 1$ (iv)

Equations (i) and (ii) represent two parallel lines which are not coincident. Similarly, equations (iii) and (iv) represent two non-coincident parallel lines. Therefore, the equations (i) to (iv) have no solution. Therefore, there do not exist any values of a, b, c, d such that $AB = I_2$ and so that matrix A above has no inverse.

Proposition

If A is non-singular, then its inverse is unique.

Proof:

Let B and C be two inverses of A.

We want to show that $B = C$.

Now $AB = 1$ (since B is an inverse of A)

$\Rightarrow$ $C(AB) = CI$ (premultiply both sides by C)

$\Rightarrow$ $(CA)B = C$ (associative law)

$\Rightarrow$ $IB = C$ (since C is an inverse of A)

$\Rightarrow$ $B = C$.

Proposition

If A and B are non-singular matrices in $M_n(F)$, then $(AB)^{-1} = B^{-1} \cdot A^{-1}$.

Proof: To show that $B^{-1}A^{-1}$ is the inverse of AB, we shall prove that

$$\begin{aligned}
(AB)(B^{-1}A^{-1}) &= I \quad \text{and} \quad (B^{-1}A^{-1})(AB) = I \\
(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} \ \text{(associative law)} \\
&= AIA^{-1} = AA^{-1} = I \\
(B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B \ \text{(associative law)} \\
&= B^{-1}IB = B^{-1}B = I
\end{aligned}$$

Note the order in which the inverse of AB has been expressed.

The Inverse of a Non-Singular Matrix

The Adjoint Method

If $A = (a_{ij}) \in M_n$, then as in Unit 2, we define for each entry a_{ij}, the *minor* for a_{ij} as the determinant of the $(n-1) \times (n-1)$-matrix M_{ij} obtained from A by deleting the i-th row and the j-th column of A. We then, define the cofactor of a_{ij} as $A_{ioj} = (-1)^{i+j}|M_{ij}|$. If A^* is the transpose of the matrix of cofators of A (called the adjoint of A), then by Property 11 of determinants in Unit 1, we have

$$AA^* = A^*A = |A| \cdot I$$

If $|A| \neq 0$, then

$$A^{-1} = \frac{1}{|A|} \cdot A^*$$

Remarks:

1. To compute the inverse of a non-singular matrix using the adjoint method, one needs only to compute the determinant and the adjoint of the matrix.

2. Since $A^{-1} = \frac{1}{|A|}$. A^*, it follows that the determinant $|A|$ gives an indication of whether or not a given matrix is invertible or not. It is invertible if and only if its determinant is non-zero.

Example 1:
Find, using the adjoint method, the inverse of

$$\begin{aligned} B &= \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & 2 \\ 1 & -1 & 0 \end{bmatrix} \\ |B| &= \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & 2 \\ 1 & -1 & 0 \end{bmatrix} = 0+0+6+2-0+2 = 10 \end{aligned}$$

Therefore, the matrix B is non-singular and so is invertible. The matrix of its cofactors

$$N = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = \begin{bmatrix} 2 & 2 & -4 \\ -2 & 2 & 1 \\ 2 & -8 & 1 \end{bmatrix}$$

The adjoint of B is

$$B^* = N^T = \begin{bmatrix} 2 & -2 & 2 \\ 2 & 2 & -8 \\ -4 & 1 & 1 \end{bmatrix}$$

Therefore

$$B^{-1} = \frac{1}{|B|} \cdot B^* = \frac{1}{10} \begin{bmatrix} 2 & -2 & 2 \\ 2 & 2 & -8 \\ -4 & 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{5} & -\frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & -\frac{4}{5} \\ -\frac{2}{5} & \frac{1}{10} & \frac{1}{10} \end{bmatrix}$$

The Inverse of a non-singular matrix using elementary operations

Another method of finding inverses of non-singular matrices is as follows. Suppose we wish to find the inverse of a non-singular matrix $A \in M_n(F)$. Then we form the matrix equation $AX = Y$ where $X = [x_1, \ldots, x_n]^T$ and $Y = [y_1, \cdots, y_n]^T$. We then solve for X by means of elimination to obtain $X = BY$. Then it follows that

$$\begin{aligned} AX = Y \text{ and } X = BY &\Rightarrow (AB)Y = Y \\ &\Rightarrow AB = I_n \Rightarrow A^{-1} = B \end{aligned}$$

In other words, the inverse A^{-1} of A is the matrix B obtained in the process of solving the matrix equation $AX = Y$ in the form $X = BY$.

To illustrate the operations involved in the solution of a matrix equation $AX = Y$ by means of elimination, we consider the following matrix equation

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 4 & 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

The matrix equation can be written in a system of 3 linear equations as follows:

$$\begin{array}{rcll} x_2 + 2x_3 & = & y_1 & (1) \\ x_1 + 2x_2 + 3x_3 & = & y_2 & (2) \\ 4x_1 + 3x_2 + 4x_3 & = & y_3 & (3) \end{array}$$

$$\begin{array}{lrcll} & x_2 + 2x_3 & = & y_1 & (1) \\ & x_1 + 2x_2 + 3x_3 & = & y_2 & (2) \\ (3) - 4 \times (2) & 5x_2 - 8x_3 & = & y_3 - 4y_2 & (3) \\ & 2 + 2x_3 & = & y_1 & (4) \\ (2) - 2 \times (1) & x_1 - x_3 & = & y_2 - 2y_1 & (5) \\ (4) + 5 \times (1) & 2x_3 & = & y_3 - 4y_2 + 5y_1 & (6) \end{array}$$

$$\begin{array}{lll} (1) - (6) & x_2 + -4y_1 + 4y_2 - y_3 & (7) \\ (5) + \frac{1}{2} \times (6) & x_1 = \frac{1}{2}y_1 - y_2 + \frac{1}{2}y_3 & (8) \\ \frac{1}{2} \times (6) & x_3 = \frac{1}{2}y_3 - 2y_2 + \frac{5}{2}y_1 & (9) \end{array}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -1 & \frac{1}{2} \\ -4 & 4 & -1 \\ \frac{5}{2} & -2 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ x_3 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 4 & 3 & 4 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -1 & \frac{1}{2} \\ -4 & 4 & -1 \\ \frac{5}{2} & -2 & \frac{1}{2} \end{pmatrix}$$

The above example shows that to obtain the inverse of a non-singular matrix, we only need to apply the following operations called elementary operations, on the matrix as well as on the identity matrix of the same order until the given matrix becomes the identity matrix. The elementary operations are:

(i) Interchanging the i-th and j-th rows, denoted by R_{ij}

(ii) Multiplying the i-th row by a non-zero scalar λ_1 denoted by $R_i(\lambda)$

(iii) Adding the j-th row multiplied by a scalar λ to the i-th row, denoted by $R_{ij}(\lambda)$.

Similarly elementary column operations are denoted by

$$C_{ij}, C_i(\lambda),\ C_{ij}(\lambda).$$

Example 2: Find, by means of elementary operations, the inverse of the matrix.

$$A = \begin{pmatrix} 4 & 8 & -1 \\ 5 & -1 & 4 \\ 6 & 8 & 2 \end{pmatrix}$$

$$\left(\begin{array}{ccc|ccc} 4 & 8 & -1 & 1 & 0 & 0 \\ 5 & -1 & 4 & 0 & 1 & 0 \\ 6 & 8 & 2 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{R_1(\frac{1}{4})} \left(\begin{array}{ccc|ccc} 1 & 2 & -\frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 5 & -1 & 4 & 0 & 1 & 0 \\ 6 & 8 & 2 & 0 & 0 & 1 \end{array}\right)$$

$$\underset{R_{31}(-6)}{\xrightarrow{R_{21}(-5)}} \left(\begin{array}{ccc|ccc} 1 & 2 & -\frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & -11 & \frac{21}{4} & -\frac{5}{4} & 1 & 0 \\ 0 & -4 & \frac{7}{2} & -\frac{3}{2} & 0 & 1 \end{array}\right)$$

$$\xrightarrow{R_2(-\frac{1}{11})} \left(\begin{array}{ccc|ccc} 1 & 2 & -\frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 1 & -\frac{21}{44} & \frac{5}{44} & -\frac{1}{11} & 0 \\ 0 & -4 & \frac{7}{2} & -\frac{3}{2} & 0 & 1 \end{array}\right)$$

$$\xrightarrow[R_{32}(4)]{R_{12}(-2)} \left(\begin{array}{ccc|ccc} 1 & 0 & -\frac{31}{44} & \frac{1}{44} & \frac{2}{11} & 0 \\ 0 & 1 & -\frac{21}{44} & \frac{5}{44} & -\frac{1}{11} & 0 \\ 0 & 0 & \frac{35}{22} & -\frac{23}{22} & -\frac{4}{11} & 1 \end{array}\right)$$

$$R_3(\frac{22}{35}) \left(\begin{array}{ccc|ccc} 1 & 0 & -\frac{21}{44} & \frac{1}{44} & \frac{2}{11} & 0 \\ 0 & 1 & -\frac{21}{44} & \frac{5}{44} & -\frac{1}{11} & 0 \\ 0 & 0 & 1 & \frac{22}{35} & -\frac{22}{35} & \frac{22}{35} \end{array}\right)$$

$$\xrightarrow[R_{23}(\frac{21}{44})]{R_{13}(-\frac{31}{44})} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{17}{35} & \frac{12}{35} & \frac{-31}{75} \\ 0 & 1 & 0 & -\frac{1}{5} & -\frac{1}{5} & \frac{3}{10} \\ 0 & 0 & 1 & \frac{23}{35} & -\frac{8}{35} & \frac{22}{35} \end{array}\right)$$

Hence

$$A^{-1} = \begin{pmatrix} \frac{17}{35} & \frac{12}{35} & \frac{-31}{70} \\ -\frac{1}{5} & -\frac{1}{5} & \frac{3}{10} \\ -\frac{23}{35} & -\frac{8}{35} & \frac{22}{35} \end{pmatrix}$$

Systems of Linear equations
Consider the system of linear equations:

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ &\vdots \\ &a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = y_n \end{aligned}$$

We may express this system of equations in a more compact form either as

$$\sum_{j=1}^{n} a_{ij}x_j = y_i, \quad i = 1, \ldots, m$$

or using a matrix equation as

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

So if we put $A = (a_{ij})_{m,n}$, $X = (x_1, \ldots, x_n)^T$ and $Y = (y_1, \ldots, y_m)^T$, we can write the system of linear equations in matrix form as $AX = Y$, A is called the matrix of coefficients, and X the matrix of variables. In this unit we shall consider the case when $m = n$ and the matrix of coefficients A is non-singular.

Proposition: The system of equations $AX = Y$, where $A \in GL_n(F)$ has a unique solution given by $X = A^{-1}Y$.

Proof: Premultiply both sides of $AX = Y$ by A^{-1}

$$A^{-1}(AX) = A^{-1}Y \Rightarrow X = A^{-1}Y$$

is a solution. To see that the solution is unique, suppose X_1 and X_2 are two solutions of the system, i.e.

$$AX_1 = Y \text{ and } AX_2 = Y$$

Then $AX_1 = AX_2$. Premultiply both sides by A^{-1}

$$A^{-1}(AX_1) = A^{-1}(AX_2) \Rightarrow X_1 = X_2$$

However, to obtain this unique solution, one may not actually calculate A^{-1} as can be shown in the following theorem, known as Cramer's Rule.

Theorem: Let $AX = Y$ be a system of linear equations over a field F, where A is a non-singular $(n \times n)$-matrix. Suppose B_i is the matrix obtained from A by replacing the i-th column of A by Y for each $i = 1, \ldots, n$. Then the solutions of the system $AX = Y$ are given by

$$x_i = \frac{|B_i|}{|A|}, \quad i = 1, \ldots, n.$$

Proof: Since A^{-1} exists, it follows that $AX = Y$ implies

$$X = A^{-1}Y = \left(\frac{1}{|A|}A^*\right)Y$$

where A^* is the adjoint of A. Therefore

$$\begin{aligned}
\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} &= \frac{1}{|A|}\begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ \vdots & & & \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \\
&= \frac{1}{|A|}\begin{bmatrix} y_1A_{11} + y_2A_{21} + \cdots + y_nA_{n1} \\ \vdots \\ y_1A_{1n} + y_2A_{2n} + \cdots + y_nA_{nn} \end{bmatrix}
\end{aligned}$$

i.e.

$$x_i = \frac{1}{|A|}(y_1A_{1i} + y_2A_{2i} + \cdots + y_nA_{ni}, \quad i = 1, \ldots, n$$

Since $a_{1i}A_{1i} + A_{2i} + \cdots + a_{ni}A_{nio} = |A|$, by definition, it follows that if we replace the i-th column $(a_{ii} \dots a_{ni})^T$ of A by $Y = (y_1, \dots, y_n)^T$ to obtain matrix B_i, then

$$y_1A_{1i} + y_2A_{2i} + \cdots + y_nA_{ni} = |B_i|$$

and

$$x_i = \frac{|B_i|}{|A|}, \quad i = 1, \dots, n.$$

Example 3: By finding the inverse of the coefficient matrix, solve the system of linear equations over the field of real numbers.

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

In matrix notation, the system of linear equations is

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

If A denotes the coefficient matrix, then

$$|A| = \begin{vmatrix} 1 & 3 & -1 \\ 1 & 1 & 1 \\ -1 & 2 & -1 \end{vmatrix} = -1 - 3 - 2 - 1 + 3 - 2 = -6$$

Hence the coefficient matrix is non-singular and so the system has a unique solution.

The cofators of the matrix A are

$$\begin{array}{lll} A_{11} = -3 & A_{12} = 0 & A_{13} = 3 \\ A_{21} = 1 & A_{22} = -2 & A_{23} = -5 \\ A_{31} = 4 & A_{32} = -2 & A_{33} = -2 \end{array}$$

Hence

$$A^{-1} = \frac{1}{|A|}\begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{23} \\ A_{13} & A_{23} & A_{33} \end{pmatrix} = -\frac{1}{6}\begin{pmatrix} -3 & 1 & 4 \\ 0 & -2 & -2 \\ 3 & -5 & 2 \end{pmatrix}$$

Therefore the solution of the system is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = -\frac{1}{6}\begin{pmatrix} -3 & 1 & 4 \\ 0 & -2 & -2 \\ 3 & -5 & 2 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = -\frac{1}{6}\begin{pmatrix} 5 \\ -4 \\ -7 \end{pmatrix}$$

$\therefore \quad x = -\frac{5}{6},\ y = \frac{2}{3},\ z = \frac{7}{6}.$

Example 4. Use Cramer's rule to solve the system of equations over the field of real numbers.

$$\begin{aligned} 3x + 2y + z &= 0 \\ 5x + 2y + z &= -2 \\ 7x + 5y + 2z &= 1 \end{aligned}$$

In matrix notation, the system of linear equations

$$\begin{pmatrix} 3 & 2 & 1 \\ 5 & 2 & 1 \\ 7 & 5 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \quad \text{i.e.} \quad AX = Y$$

$$|A| = 12 + 14 + 24 - 14 - 20 - 15 = 2$$

Hence, using Cramer's rule, the solution of the system is

$$x = \frac{1}{2}\begin{vmatrix} 0 & 1 & 1 \\ -2 & 2 & 1 \\ 1 & 5 & 2 \end{vmatrix} = \frac{1}{2}(2 - 10 - 2 + 8) = -1$$

$$y = \frac{1}{2}\begin{vmatrix} 3 & 0 & 1 \\ 5 & -2 & 1 \\ 7 & 1 & 2 \end{vmatrix} = \frac{1}{2}(-12 + 5 + 14 - 3) = 2$$

$$z = \frac{1}{2}\begin{vmatrix} 3 & 2 & 0 \\ 5 & 2 & -2 \\ 7 & 5 & 1 \end{vmatrix} = \frac{1}{2}(6 - 28 - 10 + 30) = -1$$

Hence $x = -1$, $y = 2$, $z = -1$ is the unique solution.

Practice Exercise

1. Using the adjoint method, calculate A^{-1} if
   (i) $A = \begin{pmatrix} 3 & -2 \\ 6 & -5 \end{pmatrix}$ (ii) $A = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 4 & 1 \\ 1 & 3 & 0 \end{pmatrix}$

2. Use elementary operations to compute A^{-1} if
   (i) $A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 4 & 3 & 4 \end{pmatrix}$ (ii) $A = \begin{pmatrix} 4 & 8 & -1 \\ 5 & -1 & 4 \\ 6 & 8 & 2 \end{pmatrix}$

3. By calculating the inverse of the matrix of coefficients, solve the system of linear equations over the field of real numbers:

$$\begin{array}{rrrr} (i) & 2x - y + 3z = 7 & (ii) & y + 2z = -2 \\ & x + 2y - z = 1 & & x + 2y + 3z = -1 \\ & x + y + z = 3 & & 4x + 3y + 4z = 2 \end{array}$$

4. Use Cramer's rule to solve over the field of real numbers the system of linear equations.

$$\begin{array}{rrrr} (i) & 4x + 8y - z = 5 & (ii) & x + y = -1 \\ & 5x - y + 4z = 1 & & 2x - 3z = -1 \\ & 6x + 8y + 2z = 4 & & y + 4z = 4 \end{array}$$

Summary
Non-singular and singular matrices are defined. It is shown that if A and B are non-singular, then their inverses are unique and $(AB)^{-1} = B^{-1}A^{-1}$. Two methods, the adjoint method and the method of elementary operations, are discussed for computing the inverse of a non-singular matrix. Systems of linear equations are then considered, where the matrix of coefficients is a square non-singular matrix. Such systems can be solved by either obtaining the inverse of the coefficient matrix or using Cramer's rule.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.

2. Ayres, F. *Matrices*, Schaum's Outline Series.

3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.

4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

5. Kuku, A.O. *Abstract Algebra,* Ibadan University Press 1980. Chapter 5.

LECTURE FOUR

Vector Spaces

Introduction

We shall study linearly dependent and linearly independent subsets, generators and bases of a vector space over a field, and apply them to some familiar examples.

Objectives

The reader should be able to

(i) define a vector space over a field together with linearly dependent and independent subsets, generators and bases.

(ii) determine whether a subset of certain familiar vector spaces is linearly dependent or linearly independent, or forms a generating set or a basis.

Pre-Test

1. Show that the vectors $(1, 1, 0)$, $(2, -1, -1)$, $(3, -2, 1)$ form a basis for $\mathbb{R}^3$.

2. For what values of p, q are the vectors $(p, q, 3)$ and $(2, p - q, 1)$ in $\mathbb{R}^3$ linearly dependent?

3. Show that the vectors $\underline{v}_1 = (1, 2, 3)$, $\underline{v}_2 = (1, 2, 4)$, $\underline{v}_3 = (1, 2, 5)$ are not linearly independent over Q.

4. Show that the vectors
   $\underline{\alpha}_1 = (1, 1, 0, 0)$, $\underline{\alpha}_2 = (0, 0, 1, 1)$
   $\underline{\alpha}_3 = (1, 0, 0, 4)$, $\underline{\alpha}_4 = (0, 0, 0, 2)$
   form a basis for Q^4 over Q. Determine the coordinates of the standard basis vectors.
   $\underline{e}_1 = (1, 0, 0, 0)$, $\underline{e}_2 = (0, 1, 0, 0)$
   $\underline{e}_3 = (0, 0, 1, 0)$, $\underline{e}_4 = (0, 0, 0, 1)$
   with respect to this basis.

5. Show that the vectors in the set
   $\{(1, 2, -3),\ (1, -3, 2),\ (2, -1, 5)\}$
   are linearly independent and span $\mathbb{R}^3$.

6. Find all values of t which makes the vectors
   $(t, -1, 0, 1), (1, 3, t, 0), (0, -1, -1, 0), (-2, 1, t, 1)$ linear dependent.

7. Show that the vectors
   $(-1, 1, 1), (0, 1, 0), (-1, 4, 1)$ are linearly dependent.

8. Let $M_{2,3}(\mathbb{R})$ be the vector space of 2×3 matrices over field $\mathbb{R}$. Find whether the following matrices are linearly independent or not.

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

9. Show that the set of polynomials

$$(2x + 1,\ \ x^2 - 2,\ \ x^3 - x,\ \ 3x^2)$$

   is a basis for the vector space of polynomials over $\mathbb{R}$ having degree at most 3.

10. Let $P_n(t)$ be the vector space of polynomials in t of degree $\leq n$. Determine whether or not the following set forms a basis for $P_n(t)$.

$$(1,\ 1 + t,\ 1 + t + t^2,\ 1 + t + t^2 + t^3, \ldots, 1 + t + \cdots + t^n)$$

Definition of a vector space
A *vector space* (or *linear space*) V over a field F is an Abelian group under addition such that for every $\underline{v} \in V$ and $\lambda \in F$, we have $\lambda\underline{v} \in V$ and also if $\lambda, \mu \in F$ and $\underline{u}, \underline{v} \in V$ we have

$$\left.\begin{array}{ll} (a) & (\lambda+\mu)\underline{v} = \lambda\underline{v} + \mu\underline{v} \\ (b) & \lambda(\underline{u}+\underline{v}) = \lambda u + \lambda\underline{v} \end{array}\right\} \text{Distributive laws}$$

$$(c) \quad (\lambda\mu)\underline{v} = \lambda(\mu\underline{v}) = \mu(\lambda\underline{v}), \text{ and}$$

$$(d) \quad 1 \cdot \underline{v} = \underline{v}$$

In other words, V is closed with respect to external multiplication with F (in addition to being an additive Abelian group).

The elements of the field F are called *scalars* and the members of the vector space V are called *vectors*.

Example 1: It can be easily verified that

$$\begin{array}{rcl} Q^n & = & Q \times \cdots \times Q \;\; (n \text{ factors}) \\ \mathbb{R}^n & = & \mathbb{R} \times \cdots \times \mathbb{R} \;\; (n \text{ factors}), \text{ and} \\ \mathbb{C}^n & = & \mathbb{C} \times \cdots \times \mathbb{C} \;\; (n \text{ factors}) \end{array}$$

are vector spaces over field $Q_1\mathbb{R}$ and $\mathbb{C}$, respectively.
For example if $\underline{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is a vector and $\lambda \in \mathbb{R}$, we have

$$\lambda\underline{x} = (\lambda x_1, \ldots, \lambda x_n) \in \mathbb{R}^n$$

Two vectors $(\underline{a}, \ldots, a_n)$ and $\underline{b} = (b_1, \ldots, b_n)$ are equal if and only if $a_i = b_i$, $i = 1, \ldots, n$.

Note that addition in Q^n,$\mathbb{R}^n$ and $\mathbb{C}^n$ is defined coordinate by coordinate. For instance, if $\underline{x} = (x_1, \ldots, x_n)$ and $\underline{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ then $\underline{x} + \underline{y} = (x_1 + y_1, \ldots, x_n + y_n) \in \mathbb{R}^n$.

Example 2: Let F be any field. Then

$$F^n = F \times \cdots \times F \quad (n \text{ factors})$$

is a vector space over F if we define addition and scalar multiplication as

$$\begin{aligned} \underline{x} + \underline{y} &= (x_1 + y_1, \ldots, x_n + y_n) \\ \lambda \underline{x} &= (\lambda x_1, \ldots, \lambda x_n) \end{aligned}$$

where $\underline{x} = (x_1, \ldots, x_n)$, $\underline{y} = (y_1, \ldots, y_n) \in F^n$ and $\lambda \in F$.
Note that $Q^n, \mathbb{R}^n$ and $\mathbb{C}^n$ in Example 1 are special cases of this example.

Example 3: $M_{n,m}(F)$, the set of all $(n \times m)$-matrices with entries from the field F, is a vector space over F.

We have seen that $(M_{n,m}(F), +)$ is an Abelian group. If $A = (a_{ij}) \in M_{n,m}(F)$ and $\lambda \in F$, define $\lambda A = (\lambda a_{ij})$.

Example 4. $F[x]$, the set of all polynomials in x with coefficients from a field F, is a vector space over F.

We have seen that $(F[x], +)$ is an Abelian group. If $a(x) = a_0 + a_1 x + \cdots + a_n x^n$ and $\lambda \in F$, define $\lambda a(x) = \lambda a_0 + (\lambda a_1)x + \cdots + (\lambda a_n)x^n$.

Linear dependence and independence
The vectors $\underline{v}_1, \ldots, \underline{v}_m$ in a vector space V over a field F are said to be *linearly independent over F if for scalars* $\lambda_1, \ldots, \lambda_m$ in F.

$$\lambda_1 \underline{v}_1 + \cdots + \lambda_m \underline{v}_m = \underline{0}$$

implies that $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0$. Vectors which are not linearly independent are said to be *linearly dependent.*

In other words, in a linearly dependent set of vectors, one can express one of the vectors as a linear combination of the others, while this is not possible in a linearly independent set of vectors.

Vectors $\underline{v}_1, \cdots, \underline{v}_k$ in a vector space V over a field F are said to *generate* or *span* V if every member of V can be expressed as a linear combination of $\underline{v}_1, \cdots, \underline{v}_k$, i.e. if $\underline{v}$ is any member of V, then there exist $\lambda_1, \ldots, \lambda_k \in F$ such that

$$\underline{v} = \lambda_1 \underline{v}_1 + \cdots + \lambda_k \underline{v}_k$$

A *basis for a vector space* V over a field F is a linearly independent subset of V which generates or spans V.

Example 5
Show that in the vector space $\mathbb{R}^n$, the following set of vectors

$$\{\underline{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0), \quad i = 1, \ldots, n\}$$

(with 1 in the i-th entry) is a basis for $\mathbb{R}^n$.

Assume $\lambda_1 \underline{e}_1 + \cdots + \lambda_n \underline{e}_n = \underline{0}$
with $\lambda_i \in \mathbb{R}$, $i = 1, \ldots n$, then

$$(\lambda_1, \ldots, \lambda_n) = (0, \ldots, 0)$$

and by the definition of equality of vectors, we have that

$$\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$$

This implies that $(e_1, \ldots, e_n)$ is a linearly independent subset of $\mathbb{R}^n$. To show that $\{\underline{e}_1, \ldots, \underline{e}_n\}$ generates or spans $\mathbb{R}^n$, let $\underline{a} = (a_1, \ldots, a_n\}$ be any vector in $\mathbb{R}^n$. Then one can easily show that $\underline{a} = a_1 \underline{e}_1 + \cdots + a_n \underline{e}_n$. i.e. any vector in $\mathbb{R}^n$ can be written as a linear combination of the vectors $\underline{e}_1, \ldots, \underline{e}_n$. Hence $\{\underline{e}_1, \ldots, \underline{e}_n\}$ generates or spans $\mathbb{R}^n$. Therefore $\{\underline{e}_1, \ldots, \underline{e}_n\}$ is a linearly independent subset of $\mathbb{R}^n$ which generates or spans $\mathbb{R}^n$ and so is a basis for $\mathbb{R}^n$.

Example 6: Show that vectors of the form

$$\underline{v}_i = (0, 0,\ a_{ii}, a_{ii+1}, \ldots, a_{in}), \quad i = 1, \ldots, k$$

(with $i - 1$ zeroes and $a_{ii} \neq 0$) are linearly independent in $\mathbb{R}^n$.

Let $\lambda_1 \underline{v}_1 + \cdots + \lambda_k v_k = 0$ with $\lambda_i \in \mathbb{R},\ i = 1, \ldots, k$ then

$$(\lambda_1 a_{11}, \lambda_1 a_{12} + \lambda_2 a_{22}, \ldots, \lambda_1 a_{1k} + \cdots + \lambda_k a_{kk}) = (0, \ldots, 0)$$

and by the definition of equality of vectors, we have that
$\lambda_1 a_{11} = 0$ (1)
$\lambda_1 a_{12} + \lambda_2 a_{22} = 0$ (2)
...
$\lambda_1 a_{1k} + \cdots + \lambda_k a_{kk} = 0$ (k)

Solving the above equations successively, starting with eqution (1) one finds that

$$\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0, \quad (\text{since } a_{ii} \neq 0,\ i = 1, \ldots, k)$$

Hence $(\underline{v}_1, \ldots, \underline{v}_k)$ is a linearly independent subset of $\mathbb{R}^n$.

Example 7
Show that the vectors

$$\underline{v}_1 = (1, 1, 1),\ \underline{v}_2 = (0, 1, 1),\ \ \underline{v}_3 = (0, 0, 1)$$

form a basis for $\mathbb{R}^3$.

From example 6 above $\{\underline{v}_1, \underline{v}_2, \underline{v}_3\}$ is a linearly independent subset of $\mathbb{R}^3$. To show that the vectors $\underline{v}_1, \underline{v}_2, \underline{v}_3$ generate $\mathbb{R}^3$, let $\underline{v} = (a_1, a_2, a_3)$ be any vector in $\mathbb{R}^3$, then one can easily show that

$$a_1\underline{v}_1 + (a_2 - a_1)\underline{v}_2 + (a_3 - a_2)\underline{v}_3 = (a_1, a_2, a_3) = \underline{v}.$$

Hence $\{\underline{v}_1, \underline{v}_2, \underline{v}_3\}$ generates $\mathbb{R}^3$ and so is a basis for $\mathbb{R}^3$.

Remarks

1. The vectors $\underline{e}_1, \ldots, \underline{e}_n$ in Example 5 above are called the *usual* or *standard* basis for $\mathbb{R}^n$.

2. In the special case $n = 3$, we recognise this standard basis for $\mathbb{R}^3$ as $\underline{i} = \underline{e}_1 = (1, 0, 0),\ \underline{j} = \underline{e}_2 = (0, 1, 0),\ \ \underline{k} = \underline{e}_3 = (0, 0, 1)$

3. Example 5 gives a basis for $\mathbb{R}^3$ while Example 7 gives another basis for $\mathbb{R}^3$. This shows that there exist several bases for a vector space.

Practice Exercise
Comment on the linear independence or otherwise of the following sets of vectors in $\mathbb{R}^4$.

1. $\{(1, 0, 1, 0),\ (0, 1, 0, 1)\}$

2. $\{(1, 1, 0, 0),\ \ (0, 0, 1, 1),\ \ (2, 2, -1, -1)\}$

3. $\{(1,1,3,4),(2,1,3,4),(0,0,0,0),(3,0,0,0)\}$

4. $\{(1,0,0,1),(0,1,0,0),(2,0,0,0),(1,0,3,0)\}$

5. $\{(2,4,8,4,(1,2,4,2)\}$.

Summary
A vector space over a field is defined giving examples from Cartesian products of a field, matrices and polynomials.
Conditions are then given for a set of vectors in a vector space:
(i) to be linearly dependent or linearly independent
(ii) to generate or span, the vector space
(iii) to form a basis for the vector space
Examples are then given to illustrate these ideas.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.

2. Ayres, F. *Matrices*, Schaum's Outline Series.

3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.

4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

5. Kuku, A.O. *Abstract Algebra,* Ibadan University Press 1980. Chapter 5.

LECTURE FIVE

Subspaces of Vector Spaces

Introduction

We shall give conditions under which a subset of a vector space V becomes a vector space (called a subspace) with the same operations of V. We shall see that any set of vectors in a vector space generate or span a subspace. Given subspaces of a vector space we shall consider how to form new subspaces from them. We shall then investigate some properties of finitely - generated vector spaces.

Objectives

The reader should be able to:

(i) prove whether a given subset of a vector space is a subspace;

(ii) prove when new subspaces can be formed from given ones;

(iii) give and prove properties of finitely-generated vector spaces; and

(iv) give and prove the relationship among the dimensions of subspaces and their sum.

Pre-Test

1. Prove that a set of vectors is linearly independent if and only if it contains a proper subset generating the same subspace.

2. Show that any finite set of vectors (containing at least one non-zero vector) contains a linearly independent subset which spans the same subspace.

3. Prove that any $n+1$ elements of an n-dimensional vector space are linearly dependent and no set of $n-1$ elements can span an n-dimensional vector space.

4. Show that any linearly independent set of elements of a finite-dimensional vector space V can be extended to a basis for V.

5. Let U and W be subspaces of a vector space V over a field F. Show that $U + W$ is a subspace of V where

$$U + W = (\underline{u} + \underline{w} | \underline{u} \in U, \underline{w} \in W)$$

6. If U and W are subspaces of a vector space V, prove that $U \cap W$ is a subspace of V.

7. If U and W are subspaces of a vector space V, prove that $U \cup W$ is a subspace of V if and only if either $U \subseteq W$ or $W \subseteq U$.

8. Show that the set

$$W = \left\{ \begin{pmatrix} x & y \\ -y & x \end{pmatrix} \in M_2(\mathbb{R}) \right\}$$

is a subspace of $M_2(\mathbb{R})$.

9. Show that the subset

$$W = \{\text{Symmetric matrices in } M_n(F)\}$$

is a subspace of $M_n(F)$.

10. Prove that the subset

$$W = \{A \in M_n(F) | AX = XA \text{ for some fixed matrix } X \in M_n(F)\}$$

is a subspace of $M_n(F)$.

Definition of a subspace
A *subspace* of a vector space V is a subset of V which is itself a vector space with respect to the same operations of addition and scalar multiplication in V.

Proposition 1
U is a subspace of a vector space V over a field F if and only if for every $\underline{u}, \underline{v} \in U$ and any scalar $k \in F$, then $\underline{u} - \underline{v} \in U$ and $k\underline{u} \in U$.

Proof:
$\Rightarrow$: Assume U is a subspace of V. Then if $\underline{u}, \underline{v} \in U$ and $k \in F$, we have $-\underline{v} \in U$ and so $\underline{u} - \underline{v} \in U$ and $k\underline{u} \in U$.
$\Leftarrow$: Assume that if $\underline{u}, \underline{v} \in U$ and $k \in F$, then

$$\underline{u} - \underline{v} \in U \text{ and } k\underline{u} \in F.$$

We wish to show that U is a vector space over F. We have to show that $(U, +)$ is an Abelian group. Put $\underline{u} = \underline{v}$ and so

$$\underline{u} - \underline{u} = \underline{0} \in U \text{ (Identity element)}$$

$$\underline{0}, \underline{v} \in U \Rightarrow \underline{0} - \underline{v} = -\underline{v} \in U \text{ (Inverse of } \underline{v})$$

$$\underline{u}, \underline{v} \in U \Rightarrow \underline{u}, -\underline{v} \in U \Rightarrow \underline{u} - (-\underline{v}) = \underline{u} + \underline{v} \in U \text{ (Closure).}$$

Associativity and commutativity in U follow from asociativity and commutativity in V. Hence $(U, +)$ is an Abelian group. Since $k\underline{u} \in F$ whenever $u \in U$ and $k \in F$, then U becomes a vector space, and so a subspace of V.

Corollary: U is a subspace of V if and only if for every $\alpha, \beta \in F$ and $\underline{u}, \underline{v} \in U$, then $\alpha\underline{u} + \beta\underline{v} \in U$.

Remark: Proposition 1 and its Corollary give two ways of verifying whether a subset of a vector space is a subspace.

Proposition 2. If S is a non-empty subset of a vector space V over a field F, then the set

$$\langle S\rangle = \{\alpha_1\underline{v}_1 + \cdots + \alpha_k\underline{v}_k | \alpha_i \in F, \underline{v}_i \in S\}$$

is a subspace of V. $\langle S\rangle$ consists of finite sums which are linear combinations of members of S with coefficients from the field F.

Proof. Let

$$\underline{u} = \beta_1\underline{u}_1 + \cdots + \beta_r\underline{u}_r \in \langle S\rangle$$

and

$$\alpha_1\underline{v}_1 + \cdots + \alpha_s\underline{v}_s \in \langle S\rangle$$

Then,

$$\begin{aligned}\underline{u} - \underline{v} &= \beta_1\underline{u}_1 + \cdots + \beta_r\underline{u}_r - (\alpha_a\underline{v}_1 + \cdots + \alpha_s\underline{v}_s)\\ &= \beta_1\underline{u}_1 + \cdots + \beta_r\underline{u}_r + (-\alpha_1)\underline{v}_1 + \cdots + (-\alpha_s)\underline{v}_s \in \langle S\rangle\end{aligned}$$

being a finite sum which is a linear combination of members of S with coefficients from the field F. Similarly if $k \in F$, then

$$\begin{aligned}k\underline{u} &= k(\beta_1\underline{u}_1 \cdots + \beta_r\underline{u}_r)\\ &= (k\beta_1)\underline{u}_1) + \cdots + (k\beta_r)\underline{u}_r \in \langle S\rangle\end{aligned}$$

Hence $\langle S\rangle$ is a subspace of V.

Corollary. $\langle S\rangle$ is the intersection of all subspaces containing S.

Definition. If S is a non-empty subset of a vector space V over a field F, then the set $\langle S\rangle$, consisting of finite linear combinations of members of S with coefficients from the field F is a subspace of V, called the *span* of S, or the subspace of V *generated* by S.

A vector space V over a field F is said to be *finitely-generated* if it is generated by a finite subset.

Example 1. $\mathbb{R}^n$ is a finitely generated vector space, since by Example 5 of Unit 4, it is generated by a finite subset $(e_1, \ldots, e_n)$.

Proposition 3.
The non-zero vectors $\underline{v}_1, \ldots, \underline{v}_m$ in a vector space V over a field F are linearly dependent if and only if one of the vectors is a linear combination of the preceding ones.

Proof:
$\Leftarrow$: We shall assume that $\underline{v}_k$ is a linear combination of the preceding ones, i.e.

$$\underline{v}_k = \lambda_1 \underline{v}_1 + \cdots + \lambda_{k-1} \underline{v}_{k-1}$$

Then

$$\lambda_1 \underline{v}_1 + \cdots + \lambda_{k-1} \underline{v}_{k-1} + (-1)\underline{v}_k + 0 \cdot \underline{v}_{k+1} + \cdot + 0 \cdot \underline{v}_m = \underline{0}$$

with at least one coefficient, (-1), not equal to zero. Hence $\underline{v}_1, \ldots, \underline{v}_m$ are linearly dependent.

$\Rightarrow$: We shall assume that $\{\underline{v}_1, \ldots, \underline{v}_m\}$ is a linearly dependent set of vectors, i.e. there exist $\lambda_1, \ldots, \lambda_m$ not all zero such that

$$\lambda_1 \underline{v}_1 + \cdots + \lambda_m \underline{v}_m = \underline{0}$$

Choose the last subscript k for which $\lambda_k \neq 0$. Then solve for $\underline{v}_k$ as the following linear combination.

$$\underline{v}_k = (-\lambda_k^{-1} \lambda_1)\underline{v}_1 + \cdots + (-\lambda_k^{-1}\ \lambda_{k-1})\underline{v}_{k-1}.$$

Thus $\underline{v}_k$ is a linear combinatin of the preceding vectors, except in the case $k = 1$. In this case, $\lambda_1 \underline{v}_1$, with $\lambda_1 \neq 0 \Rightarrow \underline{v}_1 = \underline{0}$ contrary to the hypothesis that none of the given vectors is equal to zero.

Corollary 1. A set of vectors is linearly dependent if and only if it contains a proper subset generating the same subspace.

Corollary 2. Any finite set of vectors (containing at least one non-zero vector) contains a linearly independent subset which spans or generates the same subspace.

Proposition 4.
If n vectors span a vector space V and if V contains r linearly independent vectors, then $n \geq r$.

Proof. Let $S = \{\underline{v}_1, \ldots, \underline{v}_n\}$ be a set of n vectors which span V and let $X = \{\underline{u}_1, \ldots, \underline{u}_r\}$ be a set of r linearly independent vectors of V. Since S spans V, $\underline{u}_1$ is a linear combination of the $\underline{v}_1$, so that the set $T = \{\underline{u}_1, \underline{v}_1, \ldots, \underline{v}_n\}$ spans V and is linearly dependent. Then by Proposition 3, some vector of T must be dependent on its predecessors. This cannot be $\underline{u}_1$ since $\underline{u}_1$ belongs to a set of linearly independent vectors. Hence some vector $\underline{v}_i$ is dependent on its predecessors $\underline{u}_1, \underline{v}_1, \ldots, \underline{v}_{i-1}$ in T. Delete this term and we obtain as in Corollary 1 to Proposition 3 a subset $S' = \{\underline{u}_1, \underline{v}_1, \ldots, \underline{v}_{i-1}, \underline{v}_{i+1}, \ldots, \underline{v}_n\}$ of T which still spans V.

Repeat the above argument r times until the elements of S are exhausted. Each time, an element of S is thrown out. Hence it must be that S originally contained at leat r elements, thus proving $n \geq r$.

Corollary 1. Any two bases for a finitely generated vector space have the same number of elements.

Proof. Let $S = \{\underline{v}_1, \ldots, \underline{v}_n\}$, $T = \{\underline{u}_1, \ldots, \underline{u}_r\}$ be two bases containing n and r vectors of a vector space, respectively. Then by Proposition 4, it follows that

$$n \geq r \ \text{ and } \ r \geq n \Rightarrow r = n$$

Remark. By Corollary 1 of Proposition 4 above, it follows that the number of elements in a basis for a finitely-generated vector space is an invariant.

Definition.
The *dimension* of a finitely-generated vector space V is the number of ele-

ments in any basis for V. If the number of elements in any basis for V is n, we then say that V is an n-dimensional vector space.

Example: $Q^n, \mathbb{R}^n$ and $\mathbb{C}^n$ are each an n-dimensional vector space. We therefore, sometimes denote an n-dimensional vector space V over a field F by $V_n(F)$.

Corollary 2. Any $n+1$ elements of an n-dimensional vector space are linearly dependent and no set of $n-1$ elements can span an n-dimensional vector space.

Proposition 5.
Any linearly independent set of elements of a finite-dimensional vector space V can be extended to a basis for V.

The next Proposition gives two ways of forming new subspaces of a vector space with two given subspaces.

Proposition 6.
Let U and W be subspaces of a vector space V over a field F. Then

(a) $U + W$ is a subspace of V, where

$$U + W = (\underline{u} + \underline{w} | \underline{u} \in U, \ \ \underline{w} \in W)$$

called the linear sum of U and W.

(b) $U \cap W$ is a subspace of V.

We shall now prove a formula connecting the dimensions of the given and the new subspaces stated in Proposition 6.

Proposition 7.
Let U and W be subspaces of a finite-dimensional vector space V over a field F. Then $\dim U + \dim W = \dim(U + W) + \dim(U \cap W)$.

Proof. Let $\underline{v}_1, \ldots, \underline{v}_n$ be a basis for $U \cap W$.

Then since $U \cap W \subseteq U$ and $U \cap W \subseteq W$, it follows from Proposition 5 that the linearly independent vectors $\underline{v}_1, \ldots, \underline{v}_n$ can be extended to bases for U and W. Then let U and W have basis.

$$\underline{v}_1, \ldots, \underline{v}_n,\ \underline{u}_1, \ldots, \underline{u}_r$$

and

$$\underline{v}_1, \ldots, \underline{v}_n,\ \underline{w}_1, \ldots, \underline{w}_s,$$

respectively.
Then it follows clearly that the set

$$T = \{\underline{v}_1, \ldots, \underline{v}_n,\ \ \underline{u}_1, \ldots, \underline{u}_r,\ \underline{w}_1, \ldots, \underline{w}_s\}$$

spans $U + W$. To show that T is a basis for $U + W$ we must show that T is a linearly independent set. To show this, let

$$\lambda_1\underline{v}_1 + \cdots + \lambda_n\underline{v}_n + \mu_1\underline{u}_1 + \cdots + \mu_r\underline{u}_r + \nu_1\underline{w}_1 + \cdots + \nu_s\underline{w}_s = \underline{0}$$

$$(*) \qquad \sum_{i=1}^{r} \mu_i\underline{u}_i = -\sum_{j=1}^{n} \lambda_j\underline{v}_j - \sum_{k=1}^{s} \nu_k\underline{w}_k$$

The left hand side of the last equation is a member of U, being a linear combination of the basis elements for U above. Similarly, the right hand side of the last equation is a member of W. Thus each side must be a member of $U \cap W$, i.e.

$$\sum_{i=1}^{r} \mu_i\underline{u}_i \in U \cap W$$

But from the above basis for $U \cap W$, it follows that

$$\sum_{i=1}^{r} \mu_i\underline{u}_i = \sum_{j=1}^{n} a_j\underline{v}_j$$

Since the set $\{\underline{u}_1, \ldots, \underline{u}_r,\ \underline{v}_1, \ldots, \underline{v}_n\}$, being a basis for U, is a linearly independent set in U and V, it follows that a linear combination of this basis as in the last equation must satisfy

$$\mu_i = 0,\ 1 \leq i \leq r \ \text{ and } \ a_j = 0,\ 1 \leq j \leq n.$$

Equation $(*)$ then reduces to a linear combination of the basis $\{\underline{v}_1, \ldots, \underline{v}_n,\ \underline{w}_1, \ldots, \underline{w}_s\}$ of W. Such an equation can only hold if and only if

$$\lambda_j = 0 (1 \leq j \leq n) \text{ and } \nu_k = 0 (1 \leq k \leq s)$$

Hence T is a linearly independent set and so is a basis for $U + W$. Now dim $(U \cap W) = n$, $\dim U = n + r$, $\dim W = n + s$, $\dim(U + W) = n + r + s$. Hence $\dim U + \dim W = \dim(U + W) + \dim(U \cap W)$.

Remark. We shall in Unit 6 apply Proposition 7 to the calculation of the dimensions of the subspaces $U + W$ and $U \cap W$, given generators of U and W.

Practice Exercise

1. Prove that U is a subspace of V over a field F if and only if for every $\alpha, \beta \in F$ and $\underline{u}, \underline{v} \in U$, then $\alpha\underline{u} + \beta\underline{v} \in U$.
2. Show that for $k \leq n$, the subset $(a_1, \ldots, a_k, 0, \ldots, 0) | a_i \in F,\ i = 1, \ldots, k)$ is a subspace of F^n.
3. Prove that if S is a non-empty subset of vector space V over a field F, then $\langle S \rangle$ is the intersection of all subspaces containing S.
4. Show that the subset
$$W = \left\{ \begin{pmatrix} x & x \\ x & x \end{pmatrix} \in M_2(F) \right\}$$
is a subspace of $M_2(F)$.

Summary
A subspace of a vector space is defined. Conditions are given for a subset of a vector space to be a subspace. A non-empty subset of a vector space generates a subspace called the span of the subset. Finitely-generated vector spaces are defined, as well as their dimensions. It is then shown that the sum and intersection of two subspaces are each also a subspace. A relationship is then established among the dimensions of two subspaces of a finitely generated vector space, their sum and intersection.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.

2. Ayres, F. *Matrices*, Schaum's Outline Series.

3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.

4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

5. Kuku, A.O. *Abstract Algebra,* Ibadan University Press 1980. Chapter 5.

LECTURE SIX

Rank of a Matrix

Introduction

Using the idea of subspaces generated by vectors in a vector space, we shall define the row rank and the column rank of a matrix as the dimension of the subspace generated by the corresponding rows and columns, regarded as vectors, of the matrix. We shall then prove that the row rank and the column rank are always equal and are called simply the rank of the matrix.

This is then applied to testing whether a set of vectors is linearly independent or not, for calculating the dimension of the subspace generated by such vectors and for calculating the dimension of the sum and intersection of two vector subspaces.

Objectives

The reader should be able to

(i) compute the row rank, column rank and rank of a matrix; and

(ii) use the rank of a matrix to test the linear independence of given vectors and calculate the dimension of the subspace generated by such vectors as well as of the sum and intersection of two vector subspaces.

Pre-Test

1. Find (a) the row rank; (b) the column rank, and (c) the rank, of the matrices

$$\text{(i)} \begin{pmatrix} 11 & 2 & -1 & 8 \\ 6 & 2 & 4 & 3 \\ 1 & 0 & 1 & 4 \end{pmatrix} \quad \text{(ii)} \begin{pmatrix} 2 & 2 & -1 & 3 \\ 6 & 0 & 2 & 5 \\ 2 & 4 & -4 & 1 \end{pmatrix}$$

2. *Determine whether or not the vectors,* $\underset{\sim}{v}_1 = (0,0,1,1)$, $\underset{\sim}{v}_2 = (2,2-1,-1)$, $\underset{\sim}{v}_3 = (1,1,0,0)$, are linearly independent in $I\!R^4$. Find the dimension and a basis for the subspace V spanned by $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \underset{\sim}{v}_3$.
Which of the vectors $\underset{\sim}{u}_1 = (1,2,-1,2)$, $\underset{\sim}{u}_2 = (3,3,1,1)$ lies in V?

3. Find the dimension and a basis for the subspace of R^4 spanned by the vectors.
$(1,0,2,-1), (3,-1,-2,0), (1,-1,-6,2), (0,1,5,-3)$.

4. Find the dimension and a basis for the subspace of $I\!R^3$ spanned by the vectors

$$(-1,1,1), (0,1,0), (-1,4,1)$$

5. Let $P_n[t]$ be the vector space of polynomials in t of degree $\leq n$. Determine whether or not the following set forms a basis for $P_n(t)$

$$(1, 1+t, 1+t+t^2, 1+t+t^2+t^3, \cdots, 1+t+\cdots+t^n).$$

6. Let $P_3[t] = \{\text{polynomials in } I\!R[t] \text{ of degree} \leq 3\}$. Find the dimension and a basis for the subspace generated by

$$t^3+2t^2-2t+1, t^3+3t^2-t+4, 2t^3+t^2-7t-7.$$

7. Find the dimension and a basis for the subspace of R^3 spanned by the vectors.

$$(2,1,-1), (3,2,1), (1,0,-3).$$

8. Let U be the subspace of R^5 generated by $[(1,3,-3,-1,4),(1,4,-1,-2,-2),(2,9,0,-5,-2)]$ and W the subspace generated by $\{(1,6,2,-2,3),(2,8,-1,-6,-5),(1,3,-1,-5,-6)\}$.
Find the dimension and a basis for
(i) $U+W$ (ii) $U\cap W$.

9. Let $M_{2,3}(R)$ be the vector space of 2×3 matrices, over field R. Find the dimension and a basis for the subspace of $M_{2,3}(R)$ generated by

$$\begin{pmatrix} -1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix} \quad \begin{pmatrix} 2 & -1 & 3 \\ 1 & 0 & -2 \end{pmatrix} \quad \begin{pmatrix} 3 & 0 & 1 \\ 2 & -3 & 4 \end{pmatrix}$$

obtained from A by elementary operations (row or column). The matrices obtained by applying the elementary row operations, R_{ij}, $R_i(\lambda)$ and $R_{ij}(\lambda)$ respectively to I_n, the identity matrix of order n, are represented by E_{ij}, $E_i(\lambda)$ and $E_{ij}(\lambda)$ and are called *elementary matrices.*

Practice Exercise

1. Show that the relation of equivalence among matrices is an equivalence relation.

2. Write out the elementary matrices for 3×3 matrices, as well as their inverses.

3(a) If $A\in M_{m,n}(F)$, show that the matrices obtained by applying R_{ij}, $R_i(\lambda)$ and $R_{ij}(\lambda)$ to A are the matrices. $E_{ij}\cdot A, E_i(\lambda)\cdot A$ and $E_{ij}(\lambda)\cdot A$ respectively, where E_{ij}, $E_i(\lambda)$ and $E_{ij}(\lambda)$ are elementary $(m\times m)$-matrices.

(b) If $A\in M_{m,n}(F)$ show that the matrices obtained by applying C_{ij}, $C_i(\lambda)$ and $C_{ij}(\lambda)$ to A are the matrices $A\cdot E_{ij}$, $A\cdot E_i(\lambda)$ and $A\cdot E_{ij}(\lambda)$ respectively where E_{ij}, $E_i(\lambda)$ and $E_{ij}(\lambda)$ are elementary $(n\times n)$-matrices.

4. Let $A\in M_n(F)$, show that

   (i) $|E_{ij}\cdot A| = |A\cdot E_{ij}| = -|A|$

(ii) $|E_i(\lambda) \cdot A| = A \cdot E_i(\lambda)| = \lambda \cdot |A|$

(iii) $|E_{ij}(\lambda) \cdot A| = |A \cdot E_{ij}(\lambda)| = |A|$.

Proposition 1
If two matrices A, $B \in M_{m,n}(F)$ are equivalent, then $B = PAQ$ for some invertible matrices $P \in GL_m(P)$, $Q \in GL_n(F)$.
[Note that $GL_q(F)$ is the multiplicative group of all invertible matrices of order q over F called the *general linear group of order q over F*].

Proof
Since equivalent matrices use both row and column operations, it follows that B can be obtained from A by pre-multiplying and post-multiplying A by elementary matrices. Thus if A and B are equivalent, there exists elementary matrices.
$P_1, \ldots, P_r \in GL_m(F)$ and $Q_1, \ldots, Q_s \in GL_n(F)$ such that
$P_r P_{r-1}, \ldots, P_1 A Q_1, \ldots, Q_s = B$.

Since elementary matrices are invertible, then A and B are equivalent implies there exist $P = P_r P_{r-1}, \ldots, P_1 \in GL_m(F)$ and $Q = Q_1, \ldots, Q_s \in GL_m(F)$ such that $PAQ = B$.

Rank of a matrix
Definition. The rows R_i of a matrix in $M_{m,n}(F)$ can be considered as vectors in the vector space F^n while the column C_i of A can be considered as vectors in F^m. Thus the rows R_i of A generate a subspace R_A of F^n while the columns F_i of A generate a subspace C_A of F^m. R_A is called the *row-space* of A and C_A, *the column-space* of A. The dimension of R_A is called the *row-rank* of A and is denoted by r_A while the dimension of C_A is called the *column rank of A and is denoted by* c_A. We shall show later that c_A = column rank = rank row = r_A.

2. The rank of a matrix A is defined as

$$\text{rank}(A) = r_A = c_A$$

Proposition 2
Row-equivalent matrices over a field have the same row space. Also column-

equivalent matrices over a field have the same column space. In other words, pre-multiplying or post-multiplying a matrix over a field by an elementary matrix does not alter the rank of a matrix.

Proof.
Consider the three elementary row operations on a matrix A with rows $A_1, \ldots, A_m$. If $B = E_{ij}A$, then obviously $R_A = R_B$, since
$\langle A_1, \ldots, A_m\rangle = \langle B_1, \ldots, B_m\rangle$. If $B = E_i(\lambda)A$, then $B_i = \lambda A_i$ and $B_j = A_j$, $j \neq i$. Now any $X \in R_A$ is of the form
$X = a_1A_t + \cdots + a_iA_i + \cdots + a_mA_m$ which implies that
$X = a_1B_1 + \cdots + \underline{a}_iB_i + \cdots + a_mB_m \in R_B$. Hence $R_A \subseteq R_B$. Also any $Y \in R_B$ is of the form

$$Y = a_1B_1 + \cdots + a_iB_i + \cdots + a_mB_m$$

which implies that

$$Y = a_1A_1 + \cdots + a_i\lambda A_i + \cdots + a_mA_m \in R_A.$$

Hence $R_B \subseteq R_A$ and so $R_A = R_B$. Finally, if $B = E_{ij}(\lambda)A$, then $B_i = A_i + \lambda A_j$ and $B_k = A_k$, $k \neq i$. Any $X \in R_A$ is of the form $X = a_1A_1 + \cdots + a_mA_m$ which implies that

$$X = a_1B_1 + \cdots + a_iB_i + \cdots + (a_j - \lambda a_i)B_j + \cdots + a_mB_m \in R_B.$$

Hence $R_A \subseteq R_B$.

Also any $Y \in R_B$ is of the form $Y = a_1B_1 + \cdots + a_iB_i + \cdots + a_mB_m$, which implies that

$$Y = a_1A_1 + \cdots + a_iA_i + \cdots + (a_j + \lambda a_i)A_j + \cdots + a_mA_m \in R_A.$$

Hence $R_B \subseteq R_A$ and so $R_A = R_B$.

The proof for column equivalent matrices is similar.

Remark
We shall apply Proposition 2 to determine the rank of any matrix by reducing the given matrix to an equivalent matrix using elementary row operations.

The reduced equivalent matrix will be such that its rank can be easily found.

Row-reduced echelon matrix
Let $A \in M_{m,n}(F)$. The first non-zero entry of a non-zero row R_i of A is called the *leading entry* of R_i. A is said to be a row-reduced echelon matrix or *in row-reduced echelon form* if

(i) the leading entry in any non-zero is 1

(ii) if C_i is a column containing a leading entry 1 of a row, then the entries in C_i below 1 are zeros.

(iii) each zero row in A is below all the non-zero rows of A; and

(iv) if there are r non-zero rows and the leading entry in row i appears in column l_i for $i = 1, 2, \ldots, r$ then $l_1 < l_2 < \ldots < l_r$.

Remarks 1. Any matrix $A \in M_{m,n}(F)$ is row-equivalent to a row-reduced echelon matrix, and the two matrices then have the same rank, by Proposition 2.

2. The rank of a row-reduced echelon matrix containing r non-zero rows is equal to r since the non-zero rows are linearly independent by Example 6 of §4. An invertible matrix cannot be equivalent to a matrix containing a zero row.

3. We can, in a similar manner, reduce any matrix to a column-reduced echelon form by using elementary column operations and the column rank of the matrix will be equal to the number of non-zero columns in the column-reduced echelon matrix.

4. We can also prove that a set of vectors is linearly independent or not by computing the rank of the matrix whose rows are the given vectors. If the rank is equal to the number of given vectors, then the vectors are linearly independent and if not the vectors are linearly dependent. Also the rank of the matrix gives the dimension of the subspace generated by the given vectors.

5. An echelon matrix is similar to a row-reduced echelon matrix except

that the leading entry in any non-zero row can be any non-zero number.

Canonical or normal form of a matrix
A row-reduced matrix is said to be in the canonical or normal form if

(i) the leading entry 1 in any non-zero is also the only non-zero entry in the column containing it; and

(ii) the leading entry in the first row is in the (1,1)-position.

Proposition 3. Any invertible matrix $A \in GL_n(F)$ is a product of elementary matrices.

Proof. If A is reduced to a canonical form, then A is equivalent to I_n. By Proposition 2, there exists invertible matrices P, Q such that $PAQ = I_n$, where P, Q are products of elementary matrices. Thus $A = P^{-1}Q^{-1}$ which is still a product of elementary matrices.

Corollary. If $B = PAQ$ for some non-singular matrices P and Q, then A and B are equivalent.

Proof. From Proposition 3, P and Q can be expressed as a product of elementary matrices. Hence A and B are equivalent.

Example 1
Find (i) the row rank, and (ii) the column rank, of the matrix over the

$$A = \begin{pmatrix} 8 & -1 & 6 & 3 & 7 \\ 3 & 0 & -1 & 4 & 3 \\ -1 & -1 & 9 & -9 & -2 \end{pmatrix}$$

field of real numbers.

Hence deduce the rank of the matrix.

(i) We shall reduce the given matrix to a row-reduced echelon form using elementary row operations.

$$\overset{R_1(\frac{1}{8})}{\sim}\begin{pmatrix} 1 & -\frac{1}{8} & \frac{3}{4} & \frac{3}{8} & \frac{7}{8} \\ 3 & 0 & -1 & 4 & 3 \\ -1 & -1 & 9 & -9 & -2 \end{pmatrix} \begin{matrix} R_{21}(-3) \\ \sim \\ R_{31}(1) \end{matrix} \begin{pmatrix} 1 & -\frac{1}{8} & \frac{3}{4} & \frac{3}{8} & \frac{7}{8} \\ 0 & \frac{3}{8} & -\frac{13}{4} & \frac{23}{8} & \frac{3}{8} \\ 0 & -\frac{9}{8} & \frac{39}{4} & -\frac{69}{8} & -\frac{9}{8} \end{pmatrix}$$

$$\overset{R_{32}(3)}{\sim}\begin{pmatrix} 1 & -\frac{1}{8} & \frac{3}{4} & \frac{3}{8} & \frac{7}{8} \\ 0 & \frac{3}{8} & -\frac{13}{4} & \frac{23}{8} & \frac{3}{8} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \overset{R_2(\frac{8}{3})}{\sim}\begin{pmatrix} 1 & -\frac{1}{8} & \frac{3}{4} & \frac{3}{8} & \frac{7}{8} \\ 0 & \frac{1}{3} & -\frac{26}{3} & \frac{23}{3} & \frac{1}{8} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Since there are only 2 non-zero rows in the row-reduced echelon form of matrix A, it follows that

$$r_A = \text{row rank of } A = 2.$$

(ii) We shall reduce the given matrix to a column-reduced echelon form using elementary column operations.

$$\overset{C_1(\frac{1}{8})}{\sim}\begin{pmatrix} 1 & -1 & 6 & 3 & 0 \\ \frac{3}{8} & 0 & -1 & 4 & 3 \\ \frac{1}{8} & -1 & 9 & -9 & -2 \end{pmatrix} \begin{matrix} C_{21}(1) \\ \sim \\ C_{31}(-6) \\ C_{41}(-3) \\ C_{51}(-7) \end{matrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{3}{8} & \frac{3}{8} & -\frac{13}{4} & \frac{23}{8} & \frac{3}{8} \\ -\frac{1}{8} & -\frac{9}{8} & \frac{39}{4} & -\frac{69}{8} & -\frac{9}{8} \end{pmatrix}$$

$$\overset{C_2(\frac{8}{3})}{\sim}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{3}{8} & 1 & -\frac{13}{4} & \frac{23}{8} & \frac{3}{8} \\ \frac{1}{8} & -3 & \frac{39}{4} & -\frac{69}{8} & -\frac{9}{8} \end{pmatrix} \begin{matrix} C_{32}(-\frac{13}{4}) \\ \sim \\ C_{42}(-\frac{23}{8}) \\ C_{52}(-\frac{3}{8}) \end{matrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ -\frac{1}{8} & -3 & 0 & 0 & 0 \end{pmatrix}$$

Since there are only 2 non-zero columns in the column-reduced echelon form of the matrix A, it follows that

$$c_A = \text{ column rank of } A = 2.$$

Hence since rank $A = r_A = c_A$, it follows that rank of $A = 2$.

Example 2.

(i) Establish the linear dependence or independence of the vectors in R^4.

$$\underline{v}_1 = (1, 0, -1, 2), \underline{v}_2 = (1, 3, 1, 6), \underline{v}_3 = (1, 5, -1, 16), \underline{v}_4 = (4, 1, 0, 2)$$

(ii) Find the dimension of the space V spanned by these vectors.

(iii) Determine whether the vector $(-5, 0, 10, -21)$ lies on V.

(i) We shall find the rank of the matrix whose rows are the given vectors

$$\begin{pmatrix} 1 & 0 & -1 & 2 \\ 1 & 3 & 1 & 6 \\ 1 & 5 & -1 & 16 \\ 4 & 1 & 0 & 2 \end{pmatrix} \begin{matrix} R_{21}(-1) \\ R_{31}(-1) \\ \sim \\ R_{41}(-4) \end{matrix} \begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 3 & 2 & 4 \\ 0 & 5 & 0 & 14 \\ 0 & 1 & 4 & -6 \end{pmatrix} \begin{matrix} R_2(\frac{1}{3}) \\ \sim \end{matrix} \begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & \frac{2}{3} & \frac{4}{3} \\ 0 & 5 & 0 & 14 \\ 0 & 1 & 4 & -6 \end{pmatrix}$$

$$\begin{matrix} R_{32}(-5) \\ \sim \\ R_{42}(-1) \end{matrix} \begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & \frac{2}{3} & \frac{4}{3} \\ 0 & 0 & -\frac{10}{3} & \frac{22}{3} \\ 0 & 0 & \frac{10}{3} & -\frac{22}{3} \end{pmatrix} \begin{matrix} R_3(-\frac{3}{10}) \\ \sim \\ R_{43}(1) \end{matrix} \begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & \frac{2}{3} & \frac{4}{3} \\ 0 & 0 & 1 & -\frac{11}{5} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Since there are only 3 non-zero rows in the row reduced echelon form of the matrix, it follows that the rank of the matrix is 3. Since the rank is not

equal to the number of given vectors, it follows that the given vectors are linearly dependent.
(ii) Hence it also follows that the dimension of the space V spanned by these vectors is 3.
(iii) To determine whether the vector $(-5, 0, 10, -21)$ lines on V, we must show that $(-5, 0, 10, -21)$ is a linear combination of the given vectors. One way to do this is to show that the basis vectors for V obtained from the row-reduced echelon form and $(-5, 0, 10, -21)$ are linearly dependent. We do this by calculating the rank of the matrix whose rows are the basis vectors for V and$(-5, 0, 10, -21)$.

$$\begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & \frac{2}{3} & \frac{4}{3} \\ 0 & 0 & 1 & -\frac{11}{5} \\ -5 & 0 & 10 & -21 \end{pmatrix} \overset{R_{41}(5)}{\sim} \begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & \frac{2}{3} & \frac{4}{3} \\ 0 & 0 & 1 & -\frac{11}{5} \\ 0 & 0 & 5 & -11 \end{pmatrix}$$

$$\overset{R_{43}(-5)}{\sim} \begin{pmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & \frac{2}{3} & \frac{4}{3} \\ 0 & 0 & 1 & -\frac{11}{5} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Since the rank of the matrix is 3, it follows that the basis vectors and $(-5, 0, 10, -21)$ are linearly dependent and so $(-5, 0, 10, -21)$ lies on V.

Example 3
Let $R[t]$ be the vector space of polynomials in t over the field R of real numbers. Find a basis and the dimension for a subspace of $R[t]$ generated by

$$[t^3 + 4t^2 - t + 3,\ t^3 + 5t^2 + 5,\ 3t^3 + 10t^2 - 5t + 5].$$

We shall write the generators in vector form using the convention that the constant term occupies the first coordinate, the coefficient of t the second coordinate, etc. Thus the generators become $(3,-1,4,1)$, $(5,0,5,1)$, $(5,-5,10,3)$.

We can then proceed as in Example 2 above. First reduce the matrix whose rows are the generators to a row-reduced echelon form.

$$\begin{pmatrix} 3 & -1 & 4 & 1 \\ 5 & 0 & 5 & 1 \\ 5 & -5 & 10 & 3 \end{pmatrix} \overset{R_1(\frac{1}{3})}{\sim} \begin{pmatrix} 1 & -\frac{1}{3} & \frac{4}{3} & \frac{1}{3} \\ 5 & 0 & 5 & 1 \\ 5 & -5 & 10 & 3 \end{pmatrix}$$

$$\begin{matrix} R_{21}(-5) \\ \sim \\ \\ R_{31}(-5) \end{matrix} \begin{pmatrix} 1 & -\frac{1}{3} & \frac{4}{3} & \frac{1}{3} \\ 0 & \frac{5}{3} & -\frac{5}{3} & -\frac{2}{3} \\ 0 & -\frac{10}{3} & \frac{10}{3} & \frac{4}{3} \end{pmatrix} \begin{matrix} R_2(\frac{3}{5}) \\ \sim \\ \\ R_{32}(2) \end{matrix} \begin{pmatrix} 1 & -\frac{1}{3} & \frac{4}{3} & \frac{1}{3} \\ 0 & 1 & -1 & -\frac{2}{5} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Since there are 2 non-zero rows in the row-reduced echelon form, it follows that the rank of the matrix is 2 and so the dimension of the subspace is 2.

A basis for the subspace is, therefore,

$$\left[\frac{1}{3}t^3 + \frac{4}{3}t^2 - \frac{1}{3}t + 1, -\frac{2}{5}t^3 - t^2 + t\right]$$

or

$$\left[t^3 + 4t^2 - t + 3,\ 2t^3 + 5t^2 - 5t\right]$$

or

$$\left[t^3 + 4t^2 - t + 3,\ t^3 + 5t^2 + 5\right] \ \text{– the original polynomials.}$$

Example 4

Find a basis and the dimension for a subspace of $M_{2,3}(F)$ generated by

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

We can represent the generators of the subspace of $M_{2,3}(F)$ as vectors in F^6 as

$$(1,2,2,0,1,3), (3,-1,1,1,0,2),\ (4,0,3,2,-3,1)$$

Form the matrix containing these generators as rows and obtain a basis and the dimension of the subspace by reducing the matrix to a row-reduced echelon form.

$$\begin{pmatrix} 1 & 2 & 2 & 0 & 1 & 3 \\ \\ 3 & -1 & 1 & 1 & 0 & 2 \\ \\ 4 & 0 & 3 & 2 & -3 & 1 \end{pmatrix} \begin{matrix} R_{21}(\frac{3}{5}) \\ \\ \sim \\ \\ R_{31}(-4) \end{matrix} \begin{pmatrix} 1 & 2 & 2 & 0 & 1 & 3 \\ \\ 0 & -7 & -5 & 1 & -3 & -7 \\ \\ 0 & -8 & -5 & 2 & -7 & -11 \end{pmatrix}$$

$$\begin{matrix} R_2(-\frac{1}{7}) \\ \sim \end{matrix} \begin{pmatrix} 1 & 2 & 2 & 0 & 1 & 3 \\ \\ 1 & 1 & \frac{5}{7} & -\frac{1}{7} & \frac{3}{7} & 1 \\ \\ 0 & -8 & -5 & 2 & -7 & -11 \end{pmatrix} \begin{matrix} R_{32}(8) \\ \\ \sim \end{matrix} \begin{pmatrix} 1 & 2 & 2 & 0 & 1 & 3 \\ \\ 0 & 1 & \dfrac{5}{7} & -\dfrac{1}{7} & \dfrac{3}{7} & 1 \\ \\ 0 & 0 & \dfrac{5}{7} & \dfrac{6}{7} & -\dfrac{27}{7} & -3 \end{pmatrix}$$

$$\begin{matrix} R_3(\frac{7}{5}) \\ \sim \end{matrix} \begin{pmatrix} 1 & 2 & 2 & 0 & 1 & 3 \\ \\ 0 & 1 & \dfrac{5}{7} & -\dfrac{1}{7} & \dfrac{3}{7} & 1 \\ \\ 0 & 0 & 1 & \dfrac{6}{5} & -\dfrac{27}{5} & -\dfrac{21}{5} \end{pmatrix}$$

Since there are 3 non-zero rows in the row-reduced echelon form, it follows that the rank of the matrix is 3 and so the dimension of the subspace is 3.

A basis for the subspace is, therefore

$$\left\{ \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 0 & 1 & \frac{5}{7} \\ -\frac{1}{7} & \frac{3}{7} & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ \frac{6}{5} & -\frac{27}{5} & -\frac{21}{5} \end{pmatrix} \right\}$$

or

$$\left\{ \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 0 & 7 & 5 \\ -1 & 3 & 7 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 5 \\ 6 & -27 & -21 \end{pmatrix} \right\}$$

or

$$\left\{ \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 3 & -1 & 1 \\ 1 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 4 & 0 & 3 \\ 2 & -3 & 1 \end{pmatrix} \right\}$$

Example 5
Let U be the subspace of R^5 generated by

$$\{(1,-1,-1,-2,0),(1,-2,-2,0,-3),(1,-1,-2,-2,1)\}$$

and W the subspace generated by

$$\{(1,-2,-3,0,-2),)1,-1,-3,2,-4),(1,-1,-2,2,-5)\}$$

Find a basis and dimension for
(i) $U+W$ (ii) $U \cap W$

(i) The vector subspace $U+W$ is generated by the union of the generators of U and W. We shall first reduce the matrix, whose rows are the generators of $U+W$, to a row-reduced echelon form

$$\begin{pmatrix} 1 & -1 & -1 & -2 & 0 \\ 1 & -2 & -2 & 0 & -3 \\ 1 & -1 & -2 & -2 & 1 \\ 1 & -2 & -3 & 0 & -2 \\ 1 & -1 & -3 & 2 & -4 \\ 1 & -1 & -2 & 2 & -5 \end{pmatrix} \begin{matrix} R_{21}(-1) \\ \sim \\ R_{31}(-1) \\ R_{41}(-1) \\ R_{51}(-1) \\ R_{61}(-1) \end{matrix} \begin{pmatrix} 1 & -1 & -1 & -2 & 0 \\ 0 & -1 & -1 & 2 & -3 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & -1 & -2 & 2 & -2 \\ 0 & 0 & -2 & 4 & -4 \\ 0 & 0 & -1 & 4 & -5 \end{pmatrix}$$

$$\begin{matrix} R_2(-1) \\ \sim \\ R_3(-1) \\ R_{42}(-1) \end{matrix} \begin{pmatrix} 1 & -1 & -1 & -2 & 0 \\ 0 & 1 & 1 & -2 & 3 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 02 & -1 & 0 & 1 \\ 0 & 0 & -2 & 4 & -4 \\ 0 & 0 & -1 & 4 & -5 \end{pmatrix} \begin{matrix} R_{43}(1) \\ \sim \\ R_{53}(2) \\ R_{63}(1) \end{matrix} \begin{pmatrix} 1 & -1 & -1 & -2 & 0 \\ 0 & 1 & 1 & -2 & 3 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 4 & -6 \\ 0 & 0 & 0 & 4 & -6 \end{pmatrix}$$

$$\begin{matrix} R_{65}(-1) \\ \sim \\ R_5(\frac{1}{4}) \end{matrix} \begin{pmatrix} 1 & -1 & -1 & -2 & 0 \\ 0 & 1 & 1 & -2 & 3 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -\frac{3}{2} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} R_{45} \\ \sim \end{matrix} \begin{pmatrix} 1 & -1 & -1 & -2 & 0 \\ 0 & 1 & 1 & -2 & 3 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -\frac{3}{2} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Since there are only 4 non-zero rows in the row-reduced echelon form of the matrix, it follows that the rank of the matrix is 4 and so $\dim(U + W) = 4$.

It also follows from Proposition 2 that a basis for the subspace $U + W$ is the set of non-zero vectors in the row-reduced echelon form of the matrix, i.e.

$$\{(1, -1, -1, -2, 0), (0, 1, 1, -2, 3), (0, 0, 1, 0, -1), (0, 0, 0, 1, -\frac{3}{2})\}$$

or the rows in the original matrix which correspond to the non-zero rows in the row-reduced echelon form, i.e.

$$\{(1, -1, -1, -2, 0), (1, -2, -2, 0, -3), (1, -1, -2, -2, 1), (1, -1, -3, 2, -4)\}$$

(ii) To determine the dimension of $U \cap W$, we shall employ Proposition 7 in §5. But first, we determine $\dim U$ and $\dim W$ as follows:

From (i) above, $\dim U = 3$. Consider subspace W,

$$\begin{pmatrix} 1 & -2 & -3 & 0 & -2 \\ 1 & -1 & -3 & 2 & -4 \\ 1 & -1 & -2 & 2 & -5 \end{pmatrix} \begin{matrix} R_{21}(-1) \\ \sim \\ R_{31}(-1) \end{matrix} \begin{pmatrix} 1 & -2 & -3 & 0 & -2 \\ 0 & 1 & 0 & 2 & -2 \\ 0 & 1 & 1 & 2 & -3 \end{pmatrix}$$

$$R_{32}(-1) \atop \sim \quad \begin{pmatrix} 1 & -2 & -3 & 0 & -2 \\ 0 & 1 & 0 & 2 & -2 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix}$$

Hence $\dim W = 3$. Using the formula

$$\dim U + \dim W = \dim(U + W) + \dim(U \cap W)$$

we obtain

$$\dim(U \cap W) = 3 + 3 - 4 = 2.$$

To determine a basis for $U \cap W$, we solve the following vector equation:

$$a_1(1, -1, -1, -2, 0) + a_2(0, 1, 1, -2, 3) + a_3(0, 0, 1, 0, -1)$$

$$= b_1(1, -2, -3, 0, -2) + b_2(0, 1, 0, 2, -2) + b_3(0, 0, 1, 0, -1)$$

for constants $a_1, a_2, a_3, b_1, b_2, b_3$ where the left hand side is a linear combination of a basis for U and the right hand side is a linear combination of a basis for W. Therefore

$$a_1 = b_1$$

$$-a_1 + a_2 = -b_1 + b_2$$

$$-a_1 + a_2 + a_3 = -3b_1 + b_3$$

$$-2a_1 - 2a_4 = 2b_2$$

$$3a_2 - a_3 = -2b_1 - 2b_2 - b_3$$

We then obtain, by putting $b_1 = k_1$, $b_3 = k_2$ that
$a_1 = k_1, a_2 = -k_1, a_3 = k_2 - k_1$
$b_1 = k_1, b_2 = 0, b_3 = k_2$.
Since

$$\begin{aligned} U \cap W &= \{b_1(1, -2, -3, 0, -2) + b_2(0, 1, 0, 2, -2) + b_3(0, 0, 1, 0, -1)\} \\ &= \{k_1(1, -2, -3, 0, -2) + k_2(0, 0, 1, 0, -1)|(k_1, k_2 \in R\} \end{aligned}$$

Hence a basis for $U \cap W$ is

$$\{(1, -2, -3, 0, -2), (0, 0, 1, 0, -1)\}$$

Practice Exercise

1. Reduce the matrix
$$\begin{pmatrix} 1 & 6 & -2 & 5 \\ 4 & 0 & 4 & -2 \\ -6 & 3 & -3 & 3 \end{pmatrix}$$
   (a) to a row-equivalent echelon form,
   (b) to a column equivalent form
   (c) to a canonical on normal form.

2. Find (a) the row rank; (b) the column rank; and (c) the rank, of the matrices.
$$\text{(i)} \begin{pmatrix} 3 & 1 & 2 & 1 \\ 2 & 0 & 3 & 2 \\ 1 & 2 & 3 & 4 \end{pmatrix} \text{(ii)} \begin{pmatrix} 1 & 2 & 4 & 3 \\ 3 & 0 & -1 & 4 \\ -1 & -1 & 9 & -9 \end{pmatrix}$$

3. Find the dimension and a basis for the subspace for R^4 generated by $\{(1, 1, 0, 0), (1, 2, 4, 0), (0, 0, 1, 1), (1, 0, 0, 4)\}$.

4. Let $\underline{v}_1 = (1, 3, 1, 0)$, $\underline{v}_2 = (4, 1, 0, 1)$, $\underline{v}_3 = (1, -1, 2, 0)$ be vectors in R^4, where R is the field of real numbers.
   (a) Show that the vectors are linearly independent and so span a subspace V of R^4 of dimension 3.
   (b) Determine whether the vector $(1, 0, 5, -1)$ lies on V
   (c) Find a vector $\underline{v}_4$ in R^4 such that $\underline{v}_1$, $\underline{v}_2$,$\underline{v}_3$, $\underline{v}_4$, span$\mathbb{R}^4$
   Justify your answer.

5. Find the dimension and a basis for the subspace of $M_{2,3}(F)$ generated by

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

6. Construct four vectors in R^3 such that any three of them form a linearly independent subset.

7. Show that the set of polynomials

$$\{2x+1, x^2-2, x^3-x, 3x^2\}$$

is a basis for the vector space of polynomials over R having degree at most 3.

8. Let V be the subspace of R^4 generated by

$$\{(3,7,10,4),\ (1,1,0,0),\ (0,0,2,2),\ (0,2,3,0)\}$$

and S the subspace generated by

$$\{(0,2,9,6),(0,-4,-6,0)\}$$

Find the dimension and a basis for
(a) $V+S$ (b) $V \cap S$.

9. Suppose U and W are subspaces of $R[t]$ generated by

$$\{(t^3+4t^2-t+3, t^3+5t^2+5, 3t^3+10t^2-5t+5)\}$$

and

$$\{t^3+4t^2+6, t^3+2t^2-t+5, 2t^3+2t^2-2t+9\}$$

respectively. Find a basis and dimension for
(i) $U+W$ (ii) $U \cap W$.

Summary
Using elementary row and column operations on matrices, we define elementary matrices, row and column ranks of any matrix, and equivalent matrices. It is shown that equivalent matrices have the same rank. We then use this idea to calculate the rank of a matrix by reducing such a matrix to a row-reduced or column-reduced echelon form. Applications are then given in testing linear independence or not of given vectors, calculating the dimension and finding bases for subspaces generated by given vectors, as well as for sum and intersection of given subspaces R^n, $R[t]$ and $M_{m,n}(R)$.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.

2. Ayres, F. *Matrices*, Schaum's Outline Series.

3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.

4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

5. Kuku, A.O. *Abstract Algebra,* Ibadan University Press 1980. Chapter 5.

LECTURE SEVEN

Linear Transformations and Matrices

Introduction

The mappings between vector spaces which preserve the operations of addition and scalar multiplication of vectors are called linear transformations. We shall show that the set of all linear transformation from a vector space to another forms a vector space. Linear transformations which are one-to-one, onto and bijective will be considered, as well as their images and kernels.

We shall then show that there is a one-to-one correspondence between matrices and linear transformations. Composition of linear transformations correspond to multiplication in matrices while the sum of linear transformations correspond to addition in matrices.

Objectives

The reader should be able to

(i) give properties of a linear transformation,

(ii) show that linear transformations from a vector space to another form a vector space,

(iii) give the connection between matrices and linear transformations.

Pre-Test

1. Let $M_{2,3}(F)$ denote the vector space of 2×3 matrices over a field F. Prove that $M_{2,3}(F)$ is isomorphic to the vector space F^6 over F.

2. Let V and W be vector spaces over the same field. Show that if $\{\underline{\beta}_1, \ldots, \underline{\beta}_m\}$ is any basis for V and $\underline{\alpha}_1, \ldots, \underline{\alpha}_m$ are any m vectors in W, then there exists one linear transformation.

$$T : V \to W \text{ with } T(\underline{\beta}_i) = \underline{\alpha}_i, \quad i = 1, \ldots, m$$

3. Let $T : R^2 \to R^2$ be defined by

$$T(a_1, a_2) = (a_2 - 3a_1, 11a_1 + 4a_2)$$

Find the matrix of T with respect to basis $\{(1, 0), (0, 1)\}$

4. Let V be the vector space of differentiable functions $f : R \to R$. Suppose D is the differential operator (linear transformation) i.e. $D = \dfrac{d}{dt}$. Find the matrix of D relative to the basis $[1, t, \sin 3t, \cos 3t]$ which generates a subspace of V.

5. Let S, T, U be linear transformations defined by the following matrices

$$A_S = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad A_T = \begin{pmatrix} 3 & -1 \\ 1 & 0 \end{pmatrix}, \quad A_U = \begin{pmatrix} 0 & 3 & 2 \\ -3 & 1 & 3 \\ 1 & 2 & 1 \end{pmatrix}$$

Calculate matrices associated with the following linear transformations.
(a) $S \circ T$ (b) $T \circ S$ (c) U^2 (d) U^{-1} (e) $S^2 - 2S + I$.

6. Find a linear transformation from R^3 to R^2 such that $(-1, 2, 1) \to (1, 0)$ and $(4, 1, 2) \to (0, 1)$.
Show that there are more than one such transformations, but that for all of them, the image of (1,1,1) is the same.

7. Verify whether the mapping is linear

$$T : M_n(F) \to M_n(F),\ T(A) = AM + MA$$

for some fixed matrix M.

8. Let $T : R^3 \to R^3$ be a mapping defined by

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y \\ 2x+y+z \\ x+3y+2z \end{bmatrix}$$

with respect to the basis $\underline{i}, \underline{j}$ and $\underline{k}$ of R^3.

(i) Write down the matrix of the transformation with respect to the given basis.

(ii) Show that T is bijective.

(iii) Define the inverse mapping T^{-1}.

(iv) Write down the matrix of transformation for the mapping T with respect to the basis

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{of } R^3$$

9. Let $T : R^2 \to R^2$ be a mapping defined by

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x - 5y \\ \\ -4x + 10y \end{bmatrix}$$

Find the kernel and the image of T.

10. Let $T : R^2 \to R^2$ be a transformation defined by the matrix

$$A = \begin{pmatrix} 1+a & 1 \\ \\ ab & 1+b \end{pmatrix}$$

where $a, b \in R$, with respect to basis vector $\underline{i}$, $\underline{j}$.

(i) Find the relation between a and b for T to be bijective.

(ii) When T is not bijective, find its image and its kernel.

(iii) Find the set of invariant points of the mapping T.

Definitions.

1. Let $V_n(F)$ and $V_m(F)$ be two vectors spaces. A mapping $T : V_n(F) \to V_m(F)$ is called a *linear transformation* (or a linear mapping or a linear operator) if for every pair of vectors $\underline{u}, \underline{v} \in V_n(F)$ and every scalar $\lambda \in F$,
$$T(\underline{u} + \underline{v}) = T(\underline{u}) + T(\underline{v}) \text{ and } T(\lambda\underline{v}) = \lambda \cdot T(\underline{v}).$$

2. A linear transformation $T : V_n(F) \to V_m(F)$ is a *monomorphism* if T is one-to-one T is an *epimorphism* if T is onto.
If T is one-to-one and onto, then T is called a vector space *isomorphism* or an $F - isomorphism$.

3. The kernel of a linear transformation $T : V_n(F) \to V_m(F)$ is defined as the set.
$$ker(T) = [\underline{v} : T(\underline{v}) = 0]$$
and the image of T is defined as the set
$$Im(T) = [T(\underline{v}) | v \in V_n(F)].$$

Proposition 1. Let $\mathcal{L} = \mathcal{L}(V_n(F), V_m(F))$ denote the set of all linear transformations from $V_n(F)$ to $V_m(F)$. Then $\mathcal{L}$ is a vector space over F under suitable addition and scalar multiplication.

Proof: If $T_1, T_2 \in \mathcal{L}$, define a mapping
$$T = T_1 + T_2 : V_n(F) \to V_m(F)$$
called the addition of T_1 and T_2 by the rule
$$T(\underline{v}) = (T_1 + T_2)(\underline{v}) = T_1(\underline{v}) + T_2(\underline{v}) \text{ for every } \underline{v} \in V_n(F)$$

Also define scalar multiplication in $\mathcal{L}$ with respect to the field F as follows. If $T \in \mathcal{L}$ and $\lambda \in F$, define $\lambda T : V_n(F) \to V_m(F)$ by the rule $(\lambda T)\underline{v} = \lambda \cdot T(\underline{v})$, for every $\underline{v} \in V_n(F)$.
Show $T_1, T_2 \in \mathcal{L} \Rightarrow T_1 + T_2 \in \mathcal{L}$:- If $\underline{u}, \underline{v} \in V_n(F)$ and $\lambda \in F$, then

$$\begin{aligned} (T_1 + T_2)(\underline{u} + \underline{v}) &= T_1(\underline{u} + \underline{v}) + T_2(\underline{u} + \underline{v}) \\ &= T_1(\underline{u}) + T_1(\underline{v}) + T_2(\underline{u}) + T_2(\underline{v}), \quad \text{since } T_1, T_2 \in \mathcal{L} \\ &= (T_1 + T_2)(\underline{u}) + (T_1 + T_2)(\underline{v}) \end{aligned}$$

Also

$$\begin{aligned} & (T_1 + T_2)(\lambda(\underline{v}) = T_1(\lambda \underline{v}) + T_2(\lambda \underline{v}) \\ = \;& \lambda T_1(\underline{v}) + \lambda T_2(\underline{v}), \text{ since } \; T_1, T_2 \in \mathcal{L} \\ = \;& \lambda(T_1(\underline{v}) + T_2(\underline{v}) = \lambda \cdot (T_1 + T_2)(\underline{v}). \end{aligned}$$

Hence $T_1 + T_2 \in \mathcal{L}$.
Show that $T \in \mathcal{L}$ and $\lambda \in F \Rightarrow \lambda T \in \mathcal{L}$: If $\underline{u}, \underline{v} \in V_n(F)$ and $\alpha \in F$, then

$$\begin{aligned} (\lambda T)(\underline{u} + \underline{v}) = \lambda T(\underline{u} + \underline{v}) &= \lambda(T(\underline{u}) + T(\underline{v})), \quad \text{since } \; T \in \mathcal{L} \\ &= (\lambda T)(\underline{u}) + (\lambda T)(\underline{v}) \end{aligned}$$

Also

$$\begin{aligned} & (\lambda T)(\alpha \underline{v}) = \lambda \cdot T(\alpha \underline{v}) \\ = \;& \lambda \alpha T(\underline{v}), \text{ since } T \in \mathcal{L} \\ = \;& \alpha(\lambda T)(\underline{v}). \text{ Hence } \; \lambda T \in \mathcal{L} \end{aligned}$$

One can easily check that $\mathcal{L}$ becomes a vector space over F.

Remark: $\mathcal{L}$ is sometimes denoted by $Hom_F(V_n(F), V_m(F))$.

Proposition 2. Let $T : V_n(F) \to V_m(F)$ be a linear transformation.
Then the kernel of T is a subspace of $V_n(F)$ and the image of T is a subspace of $V_m(F)$.

Proof: Let $\underline{v}_1$, $\underline{v}_2 \in ker(T)$ and let $\alpha, \beta \in F$.
We shall show that $\alpha\underline{v}_1 + \beta\underline{v}_2 \in ker(T)$. Now $\underline{v}_1, \underline{v}_2 \in ker(T)$ imply that

$$T(\underline{v}_1 = \underline{0} \text{ and } T(\underline{v}_2) = \underline{0}$$

Hence

$$\begin{aligned} &= T(\alpha\underline{v}_1 + \beta\underline{v}_2) = \alpha T(\underline{v}_1) + \beta T(\underline{v}_2), \text{ since } T \in \mathcal{L} \\ &= \alpha \cdot 0 + \beta \cdot 0 = \underline{0} + \underline{0} = \underline{0}. \end{aligned}$$

Therefore, $\alpha\underline{v}_1 + \beta\underline{v}_2 \in ker(T)$. Hence $ker(T)$ is a subspace of $V_n(F)$.

Now let $T(\underline{v}_1)$, $T(\underline{v}_2) \in Im(T)$ and let $\alpha, \beta \in F$. We shall show that $\alpha T(\underline{v}_1) + \beta T(\underline{v}_2) \in Im(T)$. Now $T(\underline{v}_1), T(\underline{v}_2) \in Im(T)$ imply that $\underline{v}_1 \in V_n(F)$ and $\underline{v}_2 \in V_n(F)$. Hence $\alpha\underline{v}_1 + \beta\underline{v}_2 \in V_n(F)$ and $T(\alpha\underline{v}_1 + \beta\underline{v}_2) = \alpha T(\underline{v}_1 + \beta T(\underline{v}_2)$. $(V_n(F))$ being a vector space and $T \in \mathcal{L}$). Therefore $\alpha T(\underline{v}_1) + \beta T(\underline{v}_2) \in Im(T)$. Hence $Im(T)$ is a subspace of $V_m(F)$.

Proposition 3. Let $V_n(F)$ be a finite-dimensional vector space over a field F of dimension n. Then there exists an F-isomorphism $\alpha : V_n(F) \to F^n$.

Proof: Let $\{\underline{v}_1, \ldots, \underline{v}_n\}$ be a basis of $V_n(F)$. Then any vector $\underline{v} \in V_n(F)$ can be expressed uniquely as

$$\underline{v} = \sum_{i=1}^{n} a_i \underline{v}_i, \ (a_i \in F, \ 1 \le i \le n)$$

Then define $\alpha(\underline{v}) = (a_1, \ldots, a_n)$.
It is then easy to check that, with this definition of α, α is a linear transformation which is one-to-one and onto. Thus α is a vector-space isomorphism.

Corollary 1. The mapping

$$T : M_{n,m}(F) \to F^{nm}$$

defined by

$$T(a_{ij}) = (a_{11}, \ldots, a_{1m}, \ldots, a_{n1}, \ldots, a_{nm})$$

is a vector space F-isomorphism. A basis for $M_{n,m}(F)$ consists of the nm matrices $\{(a_{rs})|a_{rs} = 0 \text{ except } r = i, s = j \text{ and } a_{ij} = 1\}$.
For instance, a basis for $M_{2,3}(F)$ is the set

$$\left\{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right\}$$

Corollary 2. The mapping

$$T : F[x] \to F^{\infty}$$

defined by

$$T(a_0 + a_1x + a_2x^2 + \cdots) = (a_0, a_1, a_2, \ldots)$$

is a vector space F-isomorphism. A basis for $F[x]$ consists of the monomials

$$\{1, x, x^2, x^3, \ldots\}$$

In particular a basis for a subspace $P_n(F)$ of $F[x]$ consisting of all polynomials in $F[x]$ of degree $\leq n$ is $\{1, x, x^2, x^3, \ldots, x^n\}$.

Connection between matrices and linear transformations
Proposition 4. There is a one-to-one correspondence between $\mathcal{L} = \mathcal{L}(V_n(F), V_m(F)$ and $M_{m,n}(F)$ with respect to fixed bases for $V_n(F)$ and $V_m(F)$.

Proof. If $\{\underline{v}_1, \ldots, \underline{v}_n\}$ is a basis for $V_n(F)$, then any $\underline{v} \in V_n(F)$ can be uniquely expressed as

$$\underline{v} = \lambda_1\underline{v}_1 + \cdots + \lambda_n\underline{v}_n$$

This implies that if $T \in \mathcal{L}$, then

$$T(\underline{v}) = \lambda_1 T(\underline{v}_1) + \cdots, \lambda_n T(\underline{v}_n)$$

Hence $\{T(\underline{v}_1), \ldots, T(\underline{v}_n)\}$ spans the image $T(V_n(F))$.
Now each $T(\underline{v}_j)$ is a member of $V_m(F)$ such that if $\{\underline{u}_1, \ldots, \underline{v}_m\}$ is a basis for $V_m(F)$, then

$$T(\underline{v}_j) = a_{1j}\underline{u}_1 + \cdots + a_{mj}\underline{u}_m, \;\; 1 \leq j \leq n.$$

We now define a mapping $f : \mathcal{L} \to M_{m,n}(F)$ as $f(T) = A_T = (a_{ij})_{m,n}$ with respect to the fixed bases of $V_n(F)$ and $V_m(F)$. Note that f is well-defined and $T(\underline{v}) = A_T\underline{v}$ for any column vector $\underline{v} \in V_n(F)$, with respect to the basis $\{\underline{v}_1, \ldots, \underline{v}_n\}$ and $A_T\underline{v} \in V_m(F)$ with respect to the basis $\{\underline{u}_1, \ldots, \underline{u}_m\}$.

Show f is one-to-one
Let $S, T \in \mathcal{L}$ such that $f(S) = f(T)$, i.e. $A_S = A_T$.
Then $S(\underline{v}) = A_S\underline{v} = T(\underline{v})$.
Hence $S = T$.

Show f is onto:
Let $A \in M_{m,n}(F)$. Define a mapping
$T_A : V_n(F) \to V_m(F)$ by $T_A(\underline{v}) = A\underline{v}$, where $\underline{v}$ is a column n-vector in $V_n(F)$.
Now
$T_A(\underline{v}_1 + \underline{v}_2) = A(\underline{v}_1 + \underline{v}_2)$. (by definition of T_A)
$= A\underline{v}_1 + A\underline{v}_2$ (by the distributive law in matrices)
$= T_A(\underline{v}_1) + T_A(\underline{v}_2)$, (by definition of T_A)
Similarly,

$$T_A(\lambda\underline{v}) = A(\lambda\underline{v}) = \lambda T_A(\underline{v}) \text{ Hence } T_A \in \mathcal{L}.$$

Therefore f is a one-to-one correspondence.

Corollary 1. $f : \mathcal{L}(V_n(F), V_m(F)) \to M_{m,n}(F)$ defined by $f(T) = A_T$ with respect to fixed bases for $V_n(F)$ and $V_m(F)$ is an F-isomorphism.

Proof: Show f is a linear transformation.
$S, T \in \mathcal{L}$ and $\lambda \in F$ implies that $f(S + T) = A_{S+T}$.
However,

$$\begin{aligned} & (S+T)(\underline{v}) = S(\underline{v}) + T(\underline{v}) \\ = \;& A_S\underline{v} + A_T\underline{v} = (A_S + A_T)\underline{v} \\ \Rightarrow \;& f(S+T) = A_{S+T} = A_S + A_T \end{aligned}$$

Similarly, $f(\lambda T) = A_{\lambda T}$ and

$$(\lambda T)\underline{v} = \lambda T(\underline{v}) = \lambda A_T\underline{v} \Rightarrow f(\lambda T) = A_{\lambda T} = \lambda f(T)$$

Since, f is 1-1 and onto by Proposition 4, we conclude that f is an F-isomorphism.

Corollary 2. The matrix representing the composition $S \circ T$ of two linear transformations S and T is the product of the matrices corresponding to S and T, i.e.

$$A_{S \circ T} = A_S \cdot A_T$$

Proof.

$$\begin{aligned} A_{S \circ T}\underline{v} &= (S \circ T)(\underline{v}) = S(T(\underline{v})) \\ &= S(A_T\underline{v}) = A_S A_T \underline{v}, \text{ for all } \underline{v}. \\ &\Rightarrow A_{S \circ T} = A_S A_T \end{aligned}$$

Corollary 3. $A_{T^{-1}} = (A_T)^{-1}$.

Proof: From Corollary 2 above

$$A_{T \circ T^{-1}} = A_T \cdot A_{T^{-1}}$$

But

$$A_{T \circ T^{-1}} = A_{Id} = I$$

$$\Rightarrow \; A_T \cdot A_{T^{-1}} = I \Rightarrow A_{T^{-1}} = (A_T)^{-1}$$

Example 1. Let T be a mapping defined by

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x + 3y \\ 3x + 10y \end{bmatrix}$$

with respect to basis vectors $\underline{i}$ and $\underline{j}$.

(i) Show that T is bijective

(ii) Define the inverse mapping T^{-1}

(iii) Find the set of invariant points under T.

(iv) Write down the matrix of transformation for the mapping T with respect to the basis

$$\left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\} \text{ of } R^2.$$

The mapping T may be expressed in matrix form as

$$\left(\begin{bmatrix} x \\ y \end{bmatrix} \right) = \begin{bmatrix} 2 & 3 \\ 3 & 10 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

where we put

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 10 \end{bmatrix}$$

(i) Since $|A| = 20 - 9 = 11 \neq 0$, it follows that A is invertible and so T is invertible or bijective.

(ii) To find T^{-1}, we find A^{-1} as

$$A^{-1} = \frac{1}{11} \begin{bmatrix} 10 & -3 \\ -3 & 2 \end{bmatrix}$$

Hence

$$T^{-1} \left(\begin{bmatrix} x \\ y \end{bmatrix} \right) = \frac{1}{11} \begin{bmatrix} 10 & -3 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{10}{11}x & -\frac{3}{11}y \\ \frac{2}{11}y & -\frac{3}{11}x \end{bmatrix}$$

(iii) To find the set of invariant points under T, we solve for x and y, the following system of equations.

$$\begin{bmatrix} 2 & 3 \\ 3 & 10 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow \quad x + 3y = 0$$

Hence the set of invariant points are all the points lying on the line $x + 3y = 0$.

(iv) A point with coordinates (x, y) with respect to the basis $\left(\begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \end{bmatrix}\right)$ has position vector $x\begin{bmatrix} 1 \\ 3 \end{bmatrix} + y\begin{bmatrix} -3 \\ 1 \end{bmatrix}$ with respect to basis vectors $\underline{i}$ and $\underline{j}$.
Hence

$$T\left(x\begin{bmatrix} 1 \\ \\ 3 \end{bmatrix} + y\begin{bmatrix} -3 \\ \\ 1 \end{bmatrix}\right) = T\begin{bmatrix} x - 3y \\ \\ 3x + y \end{bmatrix} = \begin{bmatrix} 2x - 6y + 9x + 3y \\ \\ 3x - 9y + 30x + 10y \end{bmatrix}$$

$$= \begin{bmatrix} 11x - 3y \\ \\ 33x + y \end{bmatrix} \equiv a\begin{bmatrix} 1 \\ \\ 3 \end{bmatrix} + b\begin{bmatrix} -3 \\ \\ 1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 11x - 3y \\ \\ 33x + y \end{bmatrix} = \begin{bmatrix} a - 3b \\ \\ 3a + b \end{bmatrix}$$

$$\Leftrightarrow \begin{cases} a = 11x \\ \\ b = y \end{cases} \Leftrightarrow T\begin{bmatrix} x \\ \\ y \end{bmatrix} = \begin{bmatrix} 11x \\ \\ y \end{bmatrix} = \begin{bmatrix} 11 & 0 \\ \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ \\ y \end{bmatrix}$$

with respect to the basis $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 1 \end{bmatrix}$. Hence the required matrix is

$$\begin{bmatrix} 11 & 0 \\ \\ 0 & 1 \end{bmatrix}$$

Example 2. Find a linear mapping $T : R^3 \rightarrow R^3$ whose image is generated by the vectors $(1, 2, 3)$ and $(4, 5, 6)$.

To define a linear mapping, it is sufficient to define the mapping on the elements of a basis. Thus a linear mapping $T : R^3 \rightarrow R^3$ will be defined if we define

$$T(1, 0, 0), T(0, 1, 0), \text{ and } T(0, 0, 1).$$

Since $\{(1,0,0),\,(0,1,0),\,(0,0,1)\}$ is a basis for R^3. We may define

$$T(1,0,0) = (1,2,3) \text{ and } T(0,1,0) = (4,5,6).$$

Then $T(0,0,1)$ must be a linear combination of $(1,2,3)$ and $(4.5.6)$ which generate $Im(T)$. Hence we may define $T(0,0,1) = (5,7,9)$ since $(5,7,9) = (1,2,3) + (4,5,6)$. One can easily check that the mapping $T : T^3 \to R^3$ defined by

$$T(a,b,c) = a(1,2,3) + b(4.5.6) + c(5,7,9)$$

is a linear mapping and that its image is generated by $(1,2,3)$ and $(4,5,6)$.

Example 3. Let $P_2(R) = \{\text{polynomials in } R[x] \text{ whose degree is} \leq 2\}$. Exhibit a linear transformation $T : P_2(R) \to P_2(R)$ such that $ker(T) = \langle (x^2+x) \rangle$ and $Im(T) = \langle \{x^2, x\} \rangle$ where $\langle S \rangle$ is the subspace generated by the set S.

$$T \text{ will be defined if we define } T(1),\, T(x) \text{ and } T(x^2)$$

since $\{1, x, x^2\}$ is a basis for $P_2(R)$. Since $Im(T) = \langle (x^2, x) \rangle$ and $ker(T) = \langle \{x^2 + x\} \rangle$, we may define $T(1) = x$, $T(x) = x^2$, $T(x^2) = -x^2$. T defined by

$$T(a + bx + cx^2) = ax + bx^2 - cx^2$$

is then a linear transformation as can be easily checked.

To check that $ker(T) = \langle \{x+x^2\} \rangle$, one notes that $ker(T) = \{f(x) | T(f(x)) = 0\}$. Now

$$T(f(x)) = 0 \Leftrightarrow ax + bx^2 - cx^2 = 0 \Leftrightarrow a = 0, b = c.$$

Hence

$$\begin{aligned} ker(T) &= \{f(x) = bx + bx^2 | b \in \mathbb{R}\} \\ &= \{f(x) = b(x + x^2) | b \in R\} \\ &= \langle \{x + x^2\} \rangle \end{aligned}$$

To check that $Im(T) = \langle \{x, x^2\} \rangle$, one notes that

$$\begin{aligned} Im(T) &= \{T(f(x)) | f(x) \in P_2(R)\} \\ &= \{ax + (b-c)x^2 | a, b, c \in R\} \\ &= \langle \{x, x^2\} \rangle. \end{aligned}$$

Practice Exercise

1. Show that $T : V_n(F) \rightarrow V_m(F)$ is a linear transformation if and only if for every pair $\underline{u}, \underline{v} \in V_n(F)$ and $\alpha, \beta \in F$, we have

$$T(\alpha\underline{u} + \beta\underline{v}) = \alpha T(\underline{u}) + \beta T(\underline{v})$$

2. Verify whether the following mappings are linear

   (i) $T : R^3 \rightarrow R,\ (T(x, y, z) = 2x - 3y + 4z$

   (ii) $T : R^2 \rightarrow R,\ T(x, y) = xy$

3. Let $T : R^2 \rightarrow R^2$ be a mapping defined by

$$T\left(\begin{bmatrix} x \\ \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x - 5y + 3 \\ \\ -4x + 10y - 1 \end{bmatrix}$$

   (i) Verify whether the mapping T is bijective.

   (ii) Find, if any, the set of invariant points of T.

   (iii) Find the images of the point $P(0, 1)$ and the straight line $x - y + 1 = 0$.

Summary
Definitions and properties are given of a linear transformation, one-to-one, onto and bijective linear transformations as well as the kernel and image of a linear transformation. We then show that the set of all linear transformations from a vector space $V_n(F)$ to another $V_m(F)$ is a vector space which is isomorphic to the vector space $M_{m,n}(F)$ of matrices of order $m \times n$.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.

2. Ayres, F. *Matrices*, Schaum's Outline Series.

3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.

4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

5. Kuku, A.O. *Abstract Algebra,* Ibadan University Press 1980. Chapter 5.

LECTURE EIGHT

Transformations of the Plane

Introduction

We shall study the geometrical mappings of the plane R^2 to itself such as translations, reflections, rotations, enlargements, stretches and shears and their combinations. Apart from translations,all the others are examples of a linear transformation which can be defined by appropriate 2×2 matrices. We shall consider their invariant points and lines, their inverses (if they exist) and their area scale factors.

Objectives

The reader should be able to define and give properties of simple geometrical transformation of the plane.

Pre-Test

1. $\underline{P}$ is the vector which translates the point (3,5), to the point (6,4). Q is the vector which translates the point (2,8) to the point (-2,5). Express in components, the vector
   (a) $\underline{P}$ (b) $\underline{q}$ (c) $\underline{q} - \underline{p}$.

2. Consider the matrices

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

By calculating A^2, B^2, C^2 and AC, or otherwise, identify each transformation (and the scale factor for the enlargement) represented by each matrix.

3. Find the image of the unit square $O(0,0)$, $A(1,0)$, $C(1,1)$, $B(0,1)$ under the transformation given by

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

(i) Show that the area of the image is $ad - bc$.

(ii) Identify the transformation given by a matrix M if
(a) $\det M = 1$, (b) $\det M = -1$, (c) $\det M = 0$.

4. Show that (1,1) is an invariant point and $y = x$ and $x+2y = 0$ invariant lines when a transformation is given by the matrix

$$\frac{1}{3}\begin{pmatrix} -1 & 4 \\ 2 & 1 \end{pmatrix}$$

5. A, B are transformations given by the vector functions

$$A = \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} y \\ x \end{pmatrix}, \quad B = \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} x \\ -y \end{pmatrix}$$

(i) Find the images of the point $P(2,5)$ under the transformation A, B, BA and AB.

(ii) Find the inverses of the transformations A, B, BA and AB.

6. If S is the shear with invariant line $y = 0$ such that it maps point $(1, 1)$ onto $(-1, 1)$, find the images of $O(0, 0)$, $A(1, 0)$, $C(1, 1)$, $B(0, 1)$ and calculate the area of the image of the square $OACB$ under S.

7. Find the coordinates of the point $(5, 2)$ after reflection in the y-axis followed by a quarter-turn anticlockwise about the origin.

8. Write down the matrix

   (a) of an enlargement, scale factor 5, center at $(0, 0)$,

   (b) of a shear which transforms the point $(1, 1)$ onto $(4, 1)$ and for which the line $y = 0$ is invariant.

9. A point is reflected first in the line $y = x$ and then in $y = 0$. What is the final image of the point $(3, 1)$? Find the invariant points under these successive transformations.

10. Consider the transformation

$$T : \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \frac{1}{\sqrt{5}} \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Show that T leaves the distance between any two points fixed (i.e. T is an isometry).

Translations

A translation is a transformation in which all points in the plane move by a fixed vector. If $\begin{pmatrix} x \\ y \end{pmatrix}$ is a general column vector of the plane R^2 and $\begin{pmatrix} a \\ b \end{pmatrix}$ is a fixed vector, then a translation T can be expressed as

$$T : \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$$

A translation is sometimes called a rigid transformation.

Reflections
A reflection in a line, called the mirror line, is a transformation of the plane to itself such that the image of a point is at the same distance from the mirror line as the given point and the line joining a point and its image is perpendicular to the mirror line.

We shall now consider a reflection in a line passing through the origin of the plane.

Reflection T in the line $y = x\tan\theta$.
It can be shown that

$$T\begin{pmatrix} x \\ \\ b \end{pmatrix} = \begin{pmatrix} x\cos 2\theta + y\sin 2\theta \\ \\ x\sin 2\theta - y\cos 2\theta \end{pmatrix}$$

Hence

$$T : \begin{pmatrix} x \\ \\ b \end{pmatrix} \longmapsto \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} \begin{pmatrix} x \\ \\ b \end{pmatrix}$$

Since T is given by a matrix.

$$T : \begin{pmatrix} x \\ \\ y \end{pmatrix} \longmapsto A\begin{pmatrix} x \\ \\ b \end{pmatrix} \quad \text{where} \quad A = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$$

it follows that T is a linear transformation i.e. a reflection in a line through the origin is a linear transformation.

Rotations
A rotation about the origin (called the center of the rotation), is a transformation of the plane to itself such that the angle from the position vector of a point in the plane to the position vector of its image (in the anticlockwise direction) is a fixed angle for all points in the plane.

Rotation about the origin through angle α^0
It can be shown that

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\alpha - y\sin\alpha \\ x\sin\alpha + y\cos\alpha \end{pmatrix}$$

Hence

$$T : \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

Since T is given by a matrix

$$T : \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto A\begin{pmatrix} x \\ y \end{pmatrix} \quad \text{where} \quad A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

it follows that T is a linear transformation i.e. a rotation about the origin through an angle α^0 is a linear transformation.

Remark: Some special rotations are sometimes called as follows:

Rotation about the origin through 90^0 = Quarter-turn

Rotation about the origin through 180^0 = Half-turn

Rotation about the origin through 270^0 = Three-quarter turn.

Enlargements
An enlargement T from the origin, called the centre of enlargement, with scale factor k, is a transformation of the plane to itself such that the position vector of the image of a point is k times the position vector of the given point. When $-1 < k < 1$, the enlargements obtained are in fact *reductions*.
Now,

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} kx \\ ky \end{pmatrix} = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

Since T is given by a matrix, it follows that T is a linear transformation, i.e. an enlargement from the origin is a linear transformation.

Stretches

A stretch T_1 by a scale factor a in the direction of the x-axis is a transformation of the plane to itself such that the i-component of the position vector of the image of a point is a times the $\underline{i}$-component of the given point while the $\underline{j}$-component remains unchanged. Similarly, we have a stretch T_2 by a scale factor b in the direction of the y-axis. These two types of stretches are called *one-way* stretches (or *orthogonal projections*). We also have a stretch T_3 by a scale factor a in the direction of the x-axis and by a scale factor b in the direction of the y-axis, which is called a *two-way* stretch.
Now

$$T_1\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax \\ y \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$T_2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ by \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

and

$$T_3\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax \\ by \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

Since T_1, T_2 and T_3 are given by matrices it follows that they are all linear transformations, i.e. stretches (one-way and two-way) are linear transformations.

Shears

A shear T_1 with the x-axis as the axis of shear is a transformation of the pane to itself such that points on the axis of shear are left fixed by the transformation and for the other points, the $\underline{i}$-component of the position vector of the image of a point is equal to the $\underline{i}$-component of the given point plus an amount proportional to the $\underline{j}$-component of the given point and the $\underline{j}$-component remains unchanged. A shear T_2 with the y-axis as the axis of shear is defined similarly. Now,

$$T_1\begin{pmatrix} x \\ \\ y \end{pmatrix} = \begin{pmatrix} x+ky \\ \\ y \end{pmatrix} = \begin{pmatrix} 1 & k \\ \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ \\ y \end{pmatrix}$$

$$T_2\begin{pmatrix} x \\ \\ y \end{pmatrix} = \begin{pmatrix} x \\ \\ kx+y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \\ k & 1 \end{pmatrix}\begin{pmatrix} x \\ \\ y \end{pmatrix}$$

Since T_1, T_2 are given by matrices, it follows that they are linear transformations i.e. shears with the x-axis or y-axis as axis of shear, are linear transformations.

Combinations of Transformations

Transformations of the plane to itself are usually a combination of some or all the special transformations discussed above - translations, reflections, rotations, enlargements, stretches and shears. A general transformation T of the plane to itself is a combination of a linear transformation and a translation. For example,

$$T\begin{pmatrix} x \\ \\ y \end{pmatrix} = \begin{pmatrix} ax+by+h \\ \\ cx+dy+k \end{pmatrix} = \begin{pmatrix} a & b \\ \\ c & d \end{pmatrix}\begin{pmatrix} x \\ \\ y \end{pmatrix} + \begin{pmatrix} h \\ \\ k \end{pmatrix}$$

is a transformation of the plane to itself consisting of a linear transformation with matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and a translation which moves every point by a constant vector $\begin{pmatrix} h \\ \\ k \end{pmatrix}$.

Invariant Points and Lines

A point in the plane which sets mapped to itself under a transformation is called an invariant point or a fixed point. In other words if $\underline{v}$ is an invariant point of a transformation T, then $T(\underline{v}) = \underline{v}$.

We first note that the origin $\underline{0}$ is always an invariant point under a linear transformation T of the plane to itself, since if A is the matrix of T, then

$$T(\underline{0}) = A\underline{0} = \underline{0}$$

Note: Since the origin is not an invariant point the following transformations are not linear transformations:

(i) Translation by a non-zero vector,

(ii) Reflections in a line not passing through the origin,

(iii) Rotations about a point which is not the origin.

(iv) Enlargements from a point different from the origin.

To determine the invariant points of a general transformation, one will have to solve the pair of equations represented by the matrix equation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} h \\ k \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

If the equations do not have any solutions, then there are no invariant points. If the solution of the simultaneous equations is unique, then there is only one invariant point. If, however, the two equations are identical, then the invariant points are all the points on the line represented by the common equation.

If T is a transformation of the plane to itself, we say that a line L is an invariant line or a fixed line if whenever $P \in L$ implies its image $T(P) \in L$. However it does not follow that P is mapped to itself. But if every point of a line L is mapped to itself, then L is an invariant line. If P is an invariant point under a transformation T, then the line containing the position vector $\overrightarrow{OP}$ is an invariant line under T.

The inverse of a Transformation

Some transformations have inverses. Let

$$T : \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} h \\ k \end{pmatrix} \quad \text{where } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

be a general transformation, then if matrix A has an inverse A^{-1} (i.e. if A is non-singular) then T has an inverse which we denote by T^{-1} and it is given by

$$T^{-1} : \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto A^{-1} \left[\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} h \\ k \end{pmatrix} \right]$$

We can easily show that $TT^{-1} = I$ and $T^{-1}T = I$. The case when A^{-1} does not exist (i.e. if $\det A = 0$, A being singular) represents the case such as when the image of the transformation is not the whole plane but just, a line or even a point and the corresponding transformation T does not have an inverse.

Area Scale Factor of a Transformation

Let T be a general transformation such that

$$T : \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} h \\ k \end{pmatrix} \quad \text{where we put } \ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

If S is a subset of finite area of the plane, and if S' is its image under T, then we have the following relationship between area of S and S'

$$\text{area of } S' = (\text{area of } S) \times |A|$$

$|A|$ is called the area scale factor of the transformation T.

If $|A| > 0$, the orientation of S is preserved, i.e. if S is a triangle ABC such that A', B', C' are the images of A, B, C respectively, and if ABC is taken in the clockwise or anticlockwise direction, then the image $A'B'C'$ must also be in the clockwise or anticlockwise direction, respectively.

If $|A| < 0$, the orientation of S is reverse, i.e. if ABC is taken in the clockwise or anticlockwise direction, then $A'B'C'$ is in the anticlockwise or clockwise direction, respectively.

However, if $|A| = 0$, the image of T is only a straight line or just a point.

Example 1.

A is the point whose coordinates are $(2, 1)$. The transformations T, M, R and

N are defined as follows
T is a translation of $+3$ parallel to the x-axis.
M is a reflection in the x-axis.
R is a quarter-turn anticlockwise about the origin.
N is a reflection in $y = x$.

Write down the image of A under each of the transformations:
(a) T, (b) N, (c) $M \circ T$, (d) $T \circ R$, (e) $N \circ T$, (f) T^2

(a) $T : \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \end{pmatrix}$

$T(A) = T \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$

(b) $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \to \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \to \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$M : \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$

$M(A) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$

(c) $M \circ T(A) = M(T(A)) = M \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} -5 \\ 1 \end{pmatrix}$

(d) $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \to \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \to \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$R : \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$

$$T \circ R(A) = T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = T \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

(e) $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \to \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \to \begin{pmatrix} -1 \\ 0 \end{pmatrix}$

$$N : \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$N \circ T(A) = N \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -5 \end{pmatrix}$$

(f) $T^2 A = T \circ T(A) = T \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 8 \\ 1 \end{pmatrix}$

Example 2.
A certain transformation maps the points $P(3,2)$ and $Q(5,2)$ to the points $P'(0,2)$ and $Q'(0,4)$, respectively. Show that the transformation can be effected by a single rotation and determine the angle of rotation.

The transformation can be seen to be equivalent to a translation T by a vector $\overrightarrow{PP''}$ followed by an anticlockwise rotation R through an angle 90^0 about the origin, where

$$T : \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

$$R : \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Therefore, the transformation is $R \circ T$ given by

$$R \circ T : \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \left[\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -1 \\ -2 \end{pmatrix} \right]$$

i.e.

$$R \circ T : \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

In order to prove that the transformation $R \circ T$ is a rotation, we shall show that it has only one invariant point by solving the matrix equation

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\begin{pmatrix} 2-y \\ x-1 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \begin{matrix} 2-y = x \\ x-1 = y \end{matrix}$$

Therefore $2x = 3$, $x = \dfrac{3}{2}$ and $y = \frac{1}{2}$.

Thus $R \circ T$ is a rotation with center $C\left(\dfrac{3}{2}, ; \dfrac{1}{2}\right)$.

The angle of rotation α is equal to the angle between $\vec{CP}$ and $\vec{CP}'$,
$\vec{CP} = \vec{OP} - \vec{OC} = (3,2) - \left(\frac{3}{2}, \frac{1}{2}\right) = \left(\frac{3}{2}, \frac{3}{2}\right)$
$\vec{CP}' = \vec{OP}' - \vec{OC} = (0,2) - \left(\frac{3}{2}, \frac{1}{2}\right) = \left(-\frac{3}{2}, \frac{3}{2}\right)$
Using the dot product

$$\vec{CP} \cdot \vec{CP}' = |\vec{CP}| \cdot |\vec{CP}'| \cos\alpha$$

$$\frac{-9}{4} + \frac{9}{4} = \frac{18}{4}\cos\alpha, \cos\alpha = 0$$

Therefore the angle of rotation is 90^0, anticlockwise.

Example 3

Let $A(3,-1)$, $M(2,1)$, $A'(-2,4)$, $M'(\lambda,3)$ be four given points in the coordinate plane. If $A'M'$ is the image of AM under a rotation in the given plane, find the possible values of λ.
Under a rotation, distances between any two points are unchanged.

Hence

$$(A'M')^2 = AM^2 \text{ implies}$$

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

$$\lambda^2 + 4\lambda = 0 \Rightarrow \lambda = 0 \text{ or } -4$$

Practice Exercise

1. P is the point $(4,1)$. The transformation T_1 and T_2 are defined as follows:
   T_1 is a reflection in the line $x = 3$
   T_2 is a translation by the vector $(3,0)$ followed by a reflection in $y = 3$.
   Write down the image of P under each of the transformations.
   (a) T_1, (b) T_2, (c) $T_1 \circ T_2$, (d) $T_2 \circ T_1$.

2. Find the coordinates of the image of the point $(2,3)$ after a reflection in the line $y = -x$.

3. Explain the geometrical significance of each of the following matrices.

$$P = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

4. Represent each of the transformations of the plane in the form $\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow A \begin{pmatrix} x \\ y \end{pmatrix}$

   (i) S is an anticlockwise rotation about the origin through 180^0.

   (ii) T is a reflection in the x-axis.

   (iii) R is a reflection in the y-axis.

5. Solve the matrix equation $A\underline{r} = \underline{r}$ where

$$A = \begin{pmatrix} 3 & -2 \\ 6 & -5 \end{pmatrix} \quad \text{and} \quad \underline{r} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

   Interpret your answer with respect to the linear transformation $\underline{r} \rightarrow A\underline{r}$ of the plane.

6. Consider the matrix

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

   Describe the transformation of the plane which is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow A \begin{pmatrix} x \\ y \end{pmatrix}$$

   when (i) $\alpha = 0$, (ii) $\alpha = \dfrac{\pi}{2}$, (iii) $\alpha = \pi$.

7. Consider the points $A(4,0)$, $B(6,2)$, $A'(4,5)$, $B'(0,9)$ in the plane. Let A' and B' be images of A and B respectively, under the transformation of a rotation followed by an enlargement.
Find

   (a) the angle of rotation,

   (b) the scale factor of the enlargement, and

   (c) the centre of this transformation.

8. Find the matrix for the shear which transforms the points $(0,0),(2,0),(2,1),(0,1)$ to the points $(0,0),(2,3),(2,4)$ and $(0,1)$ respectively.

9. T is the shear with invariant line $y = 1$ which maps the point $(0,0)$ onto $(-2,0)$. Find the images of $A(1,0), C(1,1), B(0,1)$. What is the image of square $OACB$ under T?

10. Show that

   (a) a translation is an isometry

   (b) a rotation is an isometry.
   [**Note:** An isometry is a mapping which preserves distance.]

Summary
The following transformations and their combinations are considered
(a) translations by a fixed vector
(b) reflections in a line passing through the origin
(c) rotations about the origin through an angle
(d) enlargements from the origin
(e) stretches by a scale factor in the direction of the x-axis and y-axis.
(f) shears with either the x-axis or the y-axis as the axis of shear
It is shown that apart from the non-zero translations all the other transformations mentioned above are linear transformations which always keep the origin fixed, and sometimes a line is mapped to itself. The area scale factor of a transformation is shown to be related to the determinant of the matrix defined by the transformation.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.

2. Ayres, F. *Matrices*, Schaum's Outline Series.

3. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

LECTURE NINE

Rank and Nullity of a Linear Transformation

Introduction

We have seen in Unit 7 that the kernel and image of a linear transformation are subspaces. In this Unit, we shall study the relationship between their dimensions and compute the dimensions and bases elements for given linear transformations.

Objectives

The reader should be able to

(i) prove the relationship between the dimensions of the kernel and image of a linear transformation; and

(ii) determine the dimensions and bases elements of the kernel and image of a given linear transformation.

Pre-Test

Find the dimension and a basis for
(i) image of T (ii) kernel of T.
for each of the following linear transformations

1. $T : R^3 \to R^2$, with $T(x, y, z) = (x + y, y + z)$

2. $T : R^3 \to R^3$, with $T(a, b, c) = (a - b,\ b + c, a)$

3. $T : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 7 & 1 & 1 \\ 3 & 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

4. $\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 2 & -1 \\ 2 & 3 & -1 & 1 \\ -2 & 0 & -5 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$

5. $T : Q^3 \to Q^3$, with $T(x, y, z) = (x + 2y + 3z,\ 2x - y + z,\ 3x - 2y)$

6. $T : R^3 \to R^3$ with $T(x, y, z) = (x + 2y - z,\ y + z,\ x + y - 2z)$

7. If $S, T : V \to V$ are linear transformations over a field F, show that

$$\text{rank } (S + T) \leq \text{rank}(S) + \text{rank}(T)$$

8. Consider the linear transformation defined by

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 3 & 2 \\ 4 & 5 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

   (a) Show that the three planes whose equations are $3x + 2y + z = 14$, $x + 3y + 2z = 20$, $4x + 5y + 3z = 34$ intersect in a single line, and find the equation of this line.

   (b) Find the image and kernel of T.

   (c) What subspace of R^3 has the point $(4, 2, 6)$ as its image under T?

   (d) What is the image under T of the subspace given by the equations

$$\frac{x+1}{1} = \frac{y-5}{-5} = \frac{z+7}{7}?$$

Definitions: Given a linear transformation

$$T : V_n(F) \to V_m(F),$$

1. the kernel of T, being a subspace of $V_n(F)$, is called the *null-space* of T and its dimension is called the *nullity* of T.

2. the dimension of the image of T (i.e. the range $T(V_n(F))$ of T, a subspace of $V_m(F)$, is called the *rank* of T.

The next proposition gives a formula connecting the nullity and the rank of a linear transformation.

Proposition
Let $T : V_n(F) \to V_m(F)$ be a linear transformation. Then
(Nullity of T) + (Rank of T) = $\dim V_n(F)$
i.e. $\dim(ker(T)) + \dim(Im(T)) = \dim V_n(F)$

Proof: Let $\dim(\ker(T)) = r$ and suppose that $S = \{\underline{v}_1, \dots, \underline{v}_r\}$ is a basis for $\ker(T)$. Since $ker(T)$ is a subspace of $V_n(F)$, S can be extended to a basis (see Proposition 5 of Unit 5)

$$S' = \{\underline{v}_1, \dots, \underline{v}_r,\ \underline{v}_{r+1}, \dots, \underline{v}_n\}$$

for $V_n(F)$. We shall prove that

$$S'' = \{T(\underline{v}_{r+1}), \dots, T(\underline{v}_n)\}$$

is a basis for $Im(T)$ and the proposition will be proved.
Show S'' spans $Im(I)$: Any element of $Im(T)$ has the form $T(\underline{v})$ for some $\underline{v} \in V_n(F)$ where

$$\underline{v} = \sum_{i=1}^{n} \lambda_i \underline{v}_i, \quad \text{(since } S' \text{ is a basis for } \ V_n(F))$$

Hence, $T(\underline{v}) = \sum_{i=1}^{n} \lambda_i T(\underline{v}_i)$; (since T is a linear transformation).
Therefore

$$T(\underline{v}) = \sum_{i=r+1}^{n} \lambda_i T(\underline{v}_i) \ \text{ since } T(\underline{v}_i) = \underline{0} \text{ for } i \le i \le r$$

Hence S'' spans $Im(T)$.
Show S'' is a linearly independent set: Let

$$\lambda_{r+1}T(\underline{v}_{r+1}) + \cdots + \lambda_n T(\underline{v}_n) = \underline{0}$$

Then

$$T(\lambda_{r+1}\underline{v}_{r+1} + \cdots + \lambda_n\underline{v}_n) = \underline{0}, \text{ since } T \text{ is a linear transformation}$$

$$\begin{aligned}
\Rightarrow\ & \lambda_{r+1}\underline{v}_{r+1} + \cdots + \lambda_n\underline{v}_n) = \in ker(T) \\
\Rightarrow\ & \sum_{i+r+1}^{n} \lambda_i\underline{v}_i = \sum_{j=1}^{r} \alpha_j\underline{v}_j, \quad \text{(since } S \text{ is a basis for } ker(T)) \\
\Rightarrow\ & \lambda_i = 0,\ r+1 \le i \le n \text{ and } \alpha_j = 0,\ 1 \le j \le r \\
& \text{(since } S' \text{ is a basis for } V_n(F)) \\
\Rightarrow\ & S'' \text{ is a linearly independent set.}
\end{aligned}$$

Thus the proposition follows.

Remarks 1. The image of a linear transformation is equal to the column space of the defining matrix.

2. The null-space of a linear transformation is equal to the row-null-space of the defining matrix.

Example. Let $T : R^4 \to R^3$ be a linear transformation defined by
$T(x_1, y, z, w) = (2x - y + 3z + 2w,\ x + 2y + z,\ x + y - 2z + w)$.
Find a basis and dimension for
(a) Image of T, (b) Kernel of T.
We can express the linear transformation in matrix form as

$$T = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 2 & -1 & 3 & 2 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$$

(a) Since the image of T is equal to the column space of the defining matrix of T, it follows that a basis for the image of T may be obtained by reducing the defining matrix of T to a column-reduced echelon form

$$\begin{pmatrix} 2 & -1 & 3 & 2 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & -2 & 1 \end{pmatrix} \overset{C_1(\frac{1}{2})}{\sim} \begin{pmatrix} 1 & -1 & 3 & 2 \\ \frac{1}{2} & 2 & 1 & 0 \\ \frac{1}{2} & 1 & -2 & 1 \end{pmatrix} \underset{\substack{C_{31}(-3) \\ C_{41}(-2)}}{\overset{C_{21}(1)}{\sim}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{5}{2} & -\frac{1}{2} & -1 \\ \frac{1}{2} & \frac{3}{2} & -\frac{7}{2} & 0 \end{pmatrix}$$

$$\underset{C_3(2)}{\overset{C_2(\frac{2}{5})}{\sim}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & -1 & -1 \\ \frac{1}{2} & \frac{3}{5} & -7 & 0 \end{pmatrix} \underset{C_{42}(1)}{\overset{C_{32}(1)}{\sim}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 \\ \frac{1}{2} & \frac{3}{5} & -\frac{32}{5} & \frac{3}{5} \end{pmatrix}$$

$$\underset{C_4(5)}{\overset{C_3(\frac{-5}{32})}{\sim}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 \\ \frac{1}{2} & \frac{3}{5} & 1 & 0 \end{pmatrix} \overset{C_{43}(-3)}{\sim} \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 \\ \frac{1}{2} & \frac{3}{5} & 1 & 0 \end{pmatrix}$$

Since there are only 3 non-zero columns in the column-reduced echelon form of the defining matrix of T, it follows that dimension of image of T is 3, and a basis for image of T is

$$\{(1, \frac{1}{2}, \frac{1}{2}),\ (0, 1, \frac{3}{5}),\ (0, 0, 1)\}$$

(b) By the Proposition above, we have that

$$\text{Rank } (T) + \text{Nullity } (T) = \dim(R^4)$$

i.e. $\dim(\text{image of } T) + \dim(\text{kernel of } T) = 4$.
$\Rightarrow \dim(\text{kernel of } T) = 4 - 3 = 1$ (from (a) above).

Since the null-space of T is equal to the solution space of the system of linear equations $A\underline{x} = \underline{0}$ where A is the defining matrix of T, it follows that a basis for the kernel of T may be obtained by finding the linearly independent solutions of $A\underline{x} = \underline{0}$. First we obtain the row-reduced echelon form of A.

$$\begin{pmatrix} 2 & -1 & 3 & 2 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & -2 & 1 \end{pmatrix} \overset{R_1(\frac{1}{2})}{\sim} \begin{pmatrix} 1 & -\frac{1}{2} & 3 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & -2 & 1 \end{pmatrix}$$

$$\begin{matrix} R_2(-1) \\ \sim \\ R_3(-1) \end{matrix} \begin{pmatrix} 1 & -\frac{1}{2} & \frac{3}{2} & 1 \\ 0 & \frac{5}{2} & -\frac{1}{2} & -1 \\ 0 & \frac{3}{2} & -\frac{7}{2} & 0 \end{pmatrix} \begin{matrix} R_2(\frac{2}{5}) \\ \sim \\ R_3(2) \end{matrix} \begin{pmatrix} 1 & -\frac{1}{2} & \frac{3}{2} & 1 \\ 0 & 1 & -\frac{1}{5} & -\frac{2}{5} \\ 0 & 3 & -7 & 0 \end{pmatrix}$$

$$R_{32}(-3) \sim \begin{pmatrix} 1 & -\frac{1}{2} & \frac{3}{2} & 1 \\ 0 & 1 & -\frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & \frac{-32}{5} & \frac{6}{5} \end{pmatrix} \overset{R_3(-\frac{5}{32})}{\sim} \begin{pmatrix} 1 & -\frac{1}{2} & \frac{3}{2} & 1 \\ 0 & 1 & -\frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 & \frac{-3}{16} \end{pmatrix}$$

From the row reduced echelon form of A, we have

$$\begin{aligned} x - \frac{1}{2}y + \frac{3}{2}z + w &= 0 \\ y - \frac{1}{5}z - \frac{2}{5}w &= 0 \\ z - \frac{3}{16}w &= 0 \end{aligned}$$

Put $w = 16k$. Then $z = 3k$.

$$y = \frac{3}{5}k + \frac{32}{5}k = 7k,$$

and

$$x = \frac{7}{2}k - \frac{9}{2}k - 16k = -17k$$

Hence the solution space of the homogeneous system $A\underline{x} = \underline{0}$ is

$$\{(x, y, z, w) = k(-17, 7, 3, 16)|k \in R)\}$$

i.e. a basis for the kernel of T is

$$\{(-17, 7, 3, 16)\}$$

Practice Exercise
Find the dimension and a basis for
(i) image of T (ii) kernel of T
for each of the following linear transformations.

1. $T : R^3 \to R^3$, with $T(x, y, z) = (x, y, 0)$
2. $T : R^3 \to R^3$, with $T(\alpha, \beta, \gamma) = (\alpha + \beta, \beta + \gamma, \gamma)$
3. $T : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 4 & -1 & 2 \\ 2 & -2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$
4. $T : Q^4 \to Q^2$, with $T(x, y, z, w) = (x, y)$
5. $T : Q^3 \to Q^3$, with $T(x, y, z) = (3x + y - z,\ x - 5y + z,\ x + 3y - z)$

Summary
The rank and nullity of a linear transformation are defined and a relationship between them is proved and applied to the determination of the dimension and a basis for the image and kernel of a given linear transformation.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.

2. Ayres, F. *Matrices*, Schaum's Outline Series.

3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.

4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

5. Kuku, A.O. *Abstract Algebra,* Ibadan University Press 1980. Chapter 5.

LECTURE TEN

Homogeneous Systems of Linear Equations

Introduction

We are now in a position to be able to solve any system of m linear equations in n unknowns with zero constant terms over a field. In Unit 2 we considered the case when $m = n$ with any constant terms, and the matrix of coefficients was invertible. In such a case, there is only one solution of the system and the solution may be obtained by using Cramer's rule.

In this Unit, we shall show that the solution space of a homogeneous system of linear equation is related to the row null-space or column null-space of the matrix of coefficients. In the process, we shall prove that the row-rank and the column rank of a matrix are always equal.

Objectives.

The reader should be able to

(i) prove that the row rank and the column rank of a matrix are equal,

(ii) solve any homogeneous system of linear equations.

Pre-Test

1. Find bases for row and column null-spaces of the matrix over R.

$$\begin{pmatrix} 2 & -2 & 6 & -4 \\ 1 & 1 & 1 & 1 \\ 3 & 1 & 5 & 0 \end{pmatrix}$$

2. Solve over R.

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

3. Find a basis over R for the solution space

$$\begin{aligned} 4x - y - 2z - w &= 0 \\ 2x + 3y - z - 2w &= 0 \\ 7y - 4z - 5w &= 0 \\ 2x - 11y + 7z + 8w &= 0 \end{aligned}$$

3. Solve over R

$$\begin{aligned} x + 2y - 5z + 4w &= 0 \\ 2x - 3y + 2z + 3w &= 0 \\ 4x - 7y + z - 6w &= 0 \end{aligned}$$

Definitions

A system of linear equations of the form

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0$$

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0$$

is called a *homogeneous system* which is always consistent (i.e. solvable) since it always has at least one solution, namely

$$x_1 = 0 = x_2 = \cdots = x_n$$

This solution is called the *trivial solution.*
Any other solution of the homogeneous system is called a *non-trivial solution.*

If we put

$$A = (a_{ij})_{m,n} \text{ and } \underline{x} = (x_1, \dots, x_n)^T$$

we can write the homogeneous system in matrix form as

$$A\underline{x} = \underline{0}$$

where A is called the matrix of coefficients.

Associated with any $(m \times n)$-matrix $A \in M_{m,n}(F)$ over a field F is a linear transformation $T_A : V_n(F) \to V_m(F)$ defined by $T_A(\underline{x}) = A\underline{x}$, for any column vector $\underline{x} \in V_n(F)$. Thus, the set of solutions of the homogeneous system of linear equations $A\underline{x} = \underline{0}$ constitutes the kernel of T_A, which is a subspace of $V_n(F)$. This subspace is called the *solution space*, of the homogeneous system of linear equations or the *row null-space* of the matrix A. Similarly, the *column null-space* of matrix A is the solution space of the homogeneous system of linear equations $\underline{y}A = \underline{0}$ where $\underline{y} \in V_m(F)$ is a row vector.

The i-th equation of the system $A\underline{x} = \underline{0}$ is

$$a_{i1}x_1 + \cdots + a_{in}x_n = 0$$

i.e. $R_i\underline{x} = 0$, where $R_i = (a_{i1}, \dots, a_{in})$ is the i-th row of A. Therefore, $\underline{x} \in V_n(F)$ is a solution of the homogeneous system $A\underline{x} = \underline{0}$ if and only if $\underline{x}$ is orthogonal to each of the row vectors R_i of A for $i = 1, \dots, m$. In other words, $\underline{x}$ is a solution to $A\underline{x} = \underline{0}$ if and only if $\underline{x}$ is orthogonal to the row space R_A of A.

We then have the following Proposition and Corollary which contain, among other things, the dimension of the solution space of a homogeneous system $A\underline{x} = \underline{0}$.

Proposition. Let $A \in M_{m,n}(F)$ be a $(m \times n)$-matrix over a field F.
Then

(a) the row rank, and column rank, of A are equal to r, say; and

(b) the dimension of the solution space of the homogeneous system of linear equations $A\underline{x} = \underline{0}$ (or the dimension of the row null-space of A) is equal to $n - r$.

Proof: For $A \in M_{m,n}(F)$, we have a linear transformation $T_A : V_n(F) \to V_m(F)$ defined by $T_A(\underline{x}) = A\underline{x}$, where $\underline{x}$ is a column n-vector in $V_n(F)$. Note that the image of T_A is given by $Im(T_A) = \langle\{T(\underline{e}_1), \ldots, T(\underline{e}_n)\}\rangle$ where $\{\underline{e}_1, \ldots, \underline{e}_n\}$ is the standard basis for $V_n(F)$. Now $T(e_i)$ is the i-th column C_i of A, for $i = 1, \ldots, n$. Hence $Im(T_A)$ is the column space of A. Also the kernel of T_A is given by

$$\ker(T_A) = \{\underline{x} \in V_n(F) | T_A(\underline{x}) = A\underline{x} = \underline{0}\}$$

Hence $ker(T_A)$ is the solution space W of the system $A\underline{x} = \underline{0}$.
Therefore, using the formula contained in the Proposition in Unit 9,

$$\text{Rank}(T_A) + \text{Nullity}(T_A) = \dim(V_n(F))$$

we have

$$c_A + \dim W = n \tag{1}$$

where c_A = column rank of A.

Now we have seen that a vector $\underline{x} \in V_n(F)$ is in the solution space W of the homogeneous system $A\underline{x} = \underline{0}$, if and only if $\underline{x}$ is orthogonal to the row space R_A of A. Hence W is equal to the orthogonal complement of the row space R_A of A. Hence

$$r_A + \dim W = n \tag{2}$$

where r_A = row rank of A. Comparing (1) and (2), we have $r_A = c_A = r$, say.
Hence

$$\dim W = n - r.$$

Corollary. Let $A\underline{x} = \underline{0}$ be a homogeneous system of m linear equations in n unknowns, such that the rank of A is equal to r.
Then

(i) if $r = n$, the system has only the trivial solution, and

(b) if $r < n$, there exists $n - r$ linearly independent solutions and every other solution is a linear combination of them.

Proof. The proof follows easily from the Proposition.

Remark 1. Similarly, it can be shown that the dimension of the column null-space of A is equal to $m - r$, where $A \in M_{m,n}(F)$ and r is the rank of A.
2. A homogeneous system $A\underline{x} = \underline{0}$ of m linear equations in n unknowns has non-trivial solutions, if and only if $\text{rank}(A) < n$. In particular, this is the case if the number of equations is less than the number of unknowns, i.e. if $m < n$.

Example 1. Solve over R

$$\begin{aligned} x + 3y - 2z &= 0 \\ x - 8y + 8z &= 0 \\ 3x - 2y + 4z &= 0 \end{aligned}$$

First we shall find the rank of the matrix of coefficients,

$$A = \begin{pmatrix} 1 & 3 & -2 \\ 1 & -8 & 8 \\ 3 & -2 & 4 \end{pmatrix} \in M_{3,3}(R)$$

$$\begin{pmatrix} 1 & 3 & -2 \\ 1 & -8 & 8 \\ 3 & -2 & 4 \end{pmatrix} \underset{R_{31}(-3)}{\overset{R_{21}(-1)}{\sim}} \begin{pmatrix} 1 & 3 & -2 \\ 0 & -11 & 10 \\ 0 & -11 & 10 \end{pmatrix} \underset{R_{32}(-1)}{\overset{R_2(\frac{-1}{11})}{\sim}} \begin{pmatrix} 1 & 3 & -2 \\ 0 & 1 & -\frac{10}{11} \\ 0 & 0 & 0 \end{pmatrix}$$

Since there are only 2 non-zero rows in the row-reduced echelon form, it follows that the rank of the matrix of coefficients is 2. Hence by the Corollary to the Proposition, the homogeneous system has $(3 - 2) = 1$ linearly independent solution, and every solution is a linear combination of it. From the row-reduced echelon form of A, the system of linear equations becomes

$$x + 3y - 2z = 0$$

$$y - \frac{10}{11}z = 0$$

Put $z = 11k$. Then, $y = 10k$ and

$$x = 22k - 30k = -8k.$$

Hence the solution space of the homogeneous system is

$$\{(x, y, z) = k(-8, 10, 11) | k \in R\}$$

i.e. a basis for the solution space of the system is $\{(-8, 10, 11)\}$.

Example 2. If $A \in M_{m,n}(F)$ and $B \in M_{n,p}(F)$ are two matrices over a field F, show that

$$\text{rank}(AB) \leq \text{rank}(B)$$

$$B\underline{x} = \underline{0} \Rightarrow AB\underline{x} = A(B\underline{x}) = A \cdot \underline{0} = \underline{0}.$$

Hence each solution $\underline{x}$ of the homogeneous system of linear equations $B\underline{x} = \underline{0}$ is a solution of the homogeneous system of linear equations $AB\underline{x} = \underline{0}$ over F. Thus

$$\text{Row null-space of } B \subseteq \text{Row null-space of } AB$$

$$\Rightarrow \quad \dim(\text{row null-space of } B) \leq \dim(\text{row null-space of } AB)$$

Therefore by the Proposition, we have

$$\begin{aligned} & p - r \leq p - s \text{ where } r = \text{rank}(B),\ s = \text{rank}(AB) \\ \Rightarrow \quad & s \leq r \Rightarrow \ \text{rank}(AB) \leq \text{rank}(B) \end{aligned}$$

Practice Exercise X

1. Find bases for the row and column null-spaces of the matrix over R.

$$\begin{pmatrix} 6 & -2 & 9 \\ 4 & -1 & 6 \\ -6 & 3 & 9 \end{pmatrix}$$

2. Solver over R

$$\begin{aligned} x+y+z &= 0 \\ y+2z+2w &= 0 \end{aligned}$$

3. Find a basis over R for the solution space

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

4. Solve over R

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

Summary
We define a homogeneous system of m linear equations in n unknowns $A\underline{x} = 0$ and its solution space in terms of the row null-space of A. We then proved that the row rank and the column rank of any matrix are always equal. This fact is then applied to find the dimension of the solution space of the homogeneous system $A\underline{x} = \underline{0}$.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.
2. Ayres, F. *Matrices*, Schaum's Outline Series.
3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.

4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

5. Kuku, A.O. *Abstract Algebra,* Ibadan University Press 1980. Chapter 5.

LECTURE ELEVEN

Non-Homogeneous Systems of Linear Equations

Introduction

We now consider any system of m linear equations in n unknowns with non-zero constant terms. We shall study when such a non-homogeneous system has any solution and then obtain all the solutions.

Objectives:

The reader should be able to

(i) determine when any non-homogeneous system of linear equations is consistent, and

(ii) obtain a general solution of a consistent non-homogeneous system of linear equations.

Pre-Test

Verify if each system of linear equations is consistent or not. Solve when possible over the field R of real numbers.

1. $x - y - z = -4$
$2x + 3y - 12x = 7$
$3x - 4y - z = -15$

2. $3x + 4y + 6z + w = 1$
$3x + 4y - 3z + 2w = 2$

3. $2x + y + z = 3$
$4x + 5y - 7z = -15$
$4x - 2y + 14z = 34$

4. $2x + 4y - 6z - 2w = -4$
$x - 3y - z - 2w = 3$
$2x - 4y + 7z + \frac{5}{2}w = 5$

5. $a_1x + a_1^2y + a_1^3z = 1$
$a_2x + a_2^2y + a_2^3z = 2$
$a_3x + a_3^2y + a_3^3z = 3$
$(a_1 \neq a_2 \neq a_3 \neq a_1)$

6. $x - y + 2z = 1$
$3x + 2y - 4z = 8$
$y - 2z = 1$

7. Show that the system

$$\begin{aligned} 2x - y + 3w &= 3 \\ 2x + 8y - 6z &= 5 \\ 2x + 5y - 4z + w &= 4 \end{aligned}$$

has no solution over R.

8. Determine the values of k such that the system

$$\begin{aligned} kx + y + z &= 1 \\ x + ky + z &= 1 \\ x + y + kz &= 1 \end{aligned}$$

have

(i) a unique solution
(ii) no solution
(iii) more than one solutions, over R.

9. Determine the conditions on a, b, c such that the system in unknowns x, y, z has a solution over R.

$$x + 2y - 3z = a$$

$$\begin{aligned} 3x - y + 2z &= b \\ x - 5y + 8z &= c \end{aligned}$$

10. Show that there are two values of k for which the system is consistent

$$\begin{aligned} kx + 3y + 2z &= 1 \\ x + (k-1)y &= 4 \\ 10y + 3z &= -2 \\ 2x - ky - z &= 3 \end{aligned}$$

Find the solution of the system for each value of k over R.

Definitions
A system of linear equations of the form

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ \cdots \\ a_{mn}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = y_m \end{gathered}$$

is called a non-homogeneous system if not all the y_1 are equal to zero. If we put

$$A = (a_{ij})_{m,n},\ x = (x_1, \ldots, x_n)^T \text{ and } \underline{y} = (y_1, \ldots, y_m)^T,$$

we can write the non-homogeneous system in the form $A\underline{x} = \underline{y}$. Matrix A is called the matrix of coefficients and $\underline{x}$ the matrix of variables. The $m \times (n+1)$-matrix $(A, \underline{y})$ is called the *augmented matrix* of the non-homogeneous system of linear equations. The system is said to be consistent or solvable over a field F if it has a solution in F.

We shall prove two Propositions. The first one shows when a non-homogeneous system of linear equations $A\underline{x} = \underline{y}$ is consistent or solvable. The second one gives a general solution of such a consistent homogeneous system.

Proposition 1
Let $A \in M_{m,n}(F)$. Then a non-homogeneous system of linear equations,

$A\underline{x} = \underline{y}$ is consistent if and only if the coefficient matrix A and the augmented matrix $(A, \underline{y})$ have the same rank.

Proof.
$\Rightarrow$ $A\underline{x} = \underline{y}$ can be written in the form

$$x_1C_1 + x_2C_2 + \cdots + x_nC_n = \underline{y}$$

where the C_i are the columns of A, i.e. $\underline{y}$ is a linear combination of the columns C_i of A. If the rank of A is r, then exactly r of the C_i are linearity independent and span C_A, the column space of A. Therefore, if $A\underline{x} = \underline{y}$ is consistent, $\underline{y} \in C_A$, being a linear combination of the C_i. Hence $(A, \underline{y})$ and A must have the same rank.

$\Leftarrow$ Assume that A and $(A, \underline{y})$ have the same rank r, say. Hence A and $(A, \underline{y})$ must have the same column space and so $\underline{y} \in C_A$, i.e.

$$\underline{y} = b_1C_1 + b_2C_2 + \cdots + b_nC_n, \; b_i \in F$$

or $A\underline{b} = \underline{y}$, where $\underline{b} = (b_1, \ldots, b_n)^T \in V_n(F)$.
Hence $\underline{b}$ is a solution of $A\underline{x} = \underline{y}$ and so the non-homogeneous system is consistent or solvable.

Proposition 2
Let $A \in M_{m,n}(F)$ and let $A\underline{x} = \underline{y}$ be a non-homogeneous system of linear equations such that the coefficient matrix A has rank r. Let $\underline{v}_1, \ldots, \underline{v}_{n-r}$ be the linearly independent solutions of the corresponding homogeneous system $A\underline{x} = \underline{0}$ and let $\underline{b}$ be a solution of $A\underline{x} = \underline{y}$. Then a vector $\underline{v} \in V_n(F)$ is a solution of the non-homogeneous system $A\underline{x} = \underline{y}$, if and only if $\underline{v}$ is of the form

$$\underline{v} = \underline{b} + a_1\underline{v}_1 + \cdots + a_{n-r}\underline{v}_{n-r}, \; a_i \in F.$$

Proof.
$\Rightarrow$ $\underline{v}$ and $\underline{b}$ are solutions of $A\underline{x} = \underline{y}$ imply that $A\underline{v} = \underline{y}$ and $A\underline{b} = \underline{y}$. Hence $A(\underline{v} - \underline{b}) = \underline{0})$ which implies that $\underline{v} - \underline{b}$ is a solution of the homogeneous system $A\underline{x} = \underline{0}$ and hence

$$\underline{v} - \underline{b} = a_1\underline{v}_1 + \cdots + a_{n-r}\underline{v}_{n-r}, \; a_i \in F$$

since $\underline{v}_1, \ldots, \underline{v}_{n-r}$ generate the solution space of $A\underline{v} = \underline{0}$.
$\Leftarrow$ Suppose $\underline{v} = \underline{b} + a_1\underline{v}_1 + \cdots + a_{n-r}\underline{v}_{n-r},\ a_i \in F$.
Then

$$\begin{aligned} A\underline{v} &= A(\underline{b} + a_1\underline{v}_1 + \cdots + a_{n-r}\underline{v}_{n-r}) \\ &= A\underline{b} + a_1A\underline{v}_1 + \cdots + a_{n-r}A\underline{v}_{n-r} \\ &= \underline{y} + \underline{0} + \cdots + \underline{0} \\ &= \underline{y} \end{aligned}$$

Hence $\underline{v} = \underline{b} + a_1\underline{v}_1 + \cdots + a_{n-r}\underline{v}_{n-r}$ is a solution of the non-homogeneous system $A\underline{x} = \underline{y}$.

Remarks

1. $\underline{v} = \underline{b} + a_1\underline{v}_1 + \cdots + a_{n-r}\underline{v}_{n-r}$ in Proposition 2, is called a general solution of the non-homogeneous system of linear equations $A\underline{x} = \underline{y}$ and $\underline{b}$ is a particular solution of the system.

2. If $A \in GL_n(F)$ is an $(n \times n)$-non-singular matrix over a field F and $A\underline{x} = \underline{y}$ is a non-homogeneous system of linear equations, then we have shown in Unit 2 that the solution of the system $A\underline{x} = \underline{y}$ is unique, and is given by Cramer's rule, which is stated and proved there.

Example 1. Verify if the following system is consistent or not. Solve, if possible, over the field R of real numbers.

$$\begin{aligned} 3x + 2y - 7z - 3w &= 1 \\ 7x - 5y + 3z + 22w &= 12 \end{aligned}$$

To test for consistency, we shall find the ranks of the coefficient matrix and the augmented matrix as follows.

$$\left(\begin{array}{cccc|c} 3 & 2 & -7 & -3 & 1 \\ 7 & -5 & 3 & 22 & 12 \end{array}\right) \overset{R_1(\frac{1}{3})}{\sim} \left(\begin{array}{cccc|c} 1 & \dfrac{2}{3} & -\dfrac{7}{3} & -3 & \dfrac{1}{3} \\ 7 & -5 & 3 & 22 & 12 \end{array}\right)$$

$$\overset{R_{21}(-7)}{\sim}\left(\begin{array}{cccc|c} 1 & \frac{2}{3} & -\frac{7}{3} & -1 & \frac{1}{3} \\ 0 & -\frac{29}{3} & \frac{58}{3} & 29 & \frac{29}{3} \end{array}\right) \overset{R_2(-\frac{2}{29})}{\sim}\left(\begin{array}{cccc|c} 1 & \frac{2}{3} & -\frac{7}{3} & -1 & \frac{1}{3} \\ 0 & 1 & -2 & -3 & -1 \end{array}\right)$$

It follows that

$$\text{Rank of coefficients matrix} = \text{Rank of augmented matrix} = 2$$

Hence the system is consistent. It also follows that the corresponding homogeneous system has (4-2) = 2 linearly independent solutions. From the row-reduced echelon form of the coefficient matrix, the homogeneous part of the system becomes.

$$\begin{aligned} x + \frac{2}{3}y - \frac{7}{3}z - w &= 0 \\ y - 2z - 3w &= 0 \end{aligned}$$

Put $w = a$, $z = b$. Then $y = 3a + 2b$ and

$$x = a + \frac{7}{3}b - 2a - \frac{4}{3}b = -a + b$$

Hence the solution of the homogeneous part of the system is

$$\{(x, y, z, w) = a(-1, 3, 0, 1) + b(1, 2, 1, 0) | a, b \in R\}$$

i.e. a basis for the solution space of the homogeneous part of the system is

$$\{(-1, 3, 0, 1), (1, 2, 1, 0)\}$$

We shall obtain next a particular solution of the non-homogeneous system. From the row-reduced echelon form of the augmented matrix, the non-homogeneous system becomes,

$$x + \frac{2}{3}y - \frac{7}{3}z - w = \frac{1}{3}$$

$$y - 2z - 3w = 1$$

Put $w = 0$, $z = 0$. Then $y = -1$ and $x = \frac{1}{3} + \frac{2}{3} = 1$.
Hence a particular solution is

$$(x, y, z, w) = (1, -1, 0, 0)$$

Therefore, a general solution of the given non-homogeneous system is, by Proposition 2,

$$\{(x, y, z, w) = (1, -1, 0, 0) + a(-1, 3, 0, 1) + b(1, 2, 1, 0) | a, b \in R\}$$

Remark. Note that several particular solutions may be obtained depending on your substitutions but the resultant general solution sets are the same.

We may, however, adopt the following convention in obtaining the general solution of a non-homogeneous system of linear equations. We combine the two steps above into one by considering the row-reduced echelon form of only the augmented matrix as follows:

$$\begin{aligned} x + \frac{2}{3}y - \frac{7}{3}z - w &= \frac{1}{3} \\ y - 2z - 3w &= -1 \end{aligned}$$

Since there are two linearly independent solutions, we have two free variables or two degrees of freedom. Hence put $w = a$ and $z = b$. Then

$$y = 3a + 2b - 1, \text{ and } x = a + \frac{7}{3}b - \frac{2}{3}(+2b - 1) + \frac{1}{3}$$
$$= -a + b + 1.$$

Hence the general solution is

$$\{(x, y, z, w) = (1, -1, 0, 0) + a(-1, 3, 0, 1) + b(1, 2, 1, 0) | a, b \in R\}$$

Example 2. Find a values of λ for which the system of linear equations

$$\begin{aligned} (2 - \lambda)x + 2y + 3 &= 0 \\ 2x + (4 - \lambda)y + 7 &= 0 \\ 2x + 5y + 6 - \lambda &= 0 \end{aligned}$$

are consistent.

The system can be expressed in matrix form as

$$\begin{pmatrix} 2-\lambda & 2 & 3 \\ 2 & 4-\lambda & 7 \\ 2 & 5 & 6-\lambda \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

i.e. $A = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \underline{0}$, where $A = \begin{pmatrix} 2-\lambda & 2 & 3 \\ 2 & 4-\lambda & 7 \\ 2 & 5 & 6-\lambda \end{pmatrix}$.

If A^{-1} exists, then $\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = A^{-1}0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. This implies $1 = 0$, which is absurd. This means that the system is inconsistent, if A^{-1} exists. Hence for the system to be consistent, A^{-1} must not exist and so the determinant $|A| = 0$, i.e.

$$\begin{vmatrix} 2-\lambda & 2 & 3 \\ 2 & 4-\lambda & 7 \\ 2 & 5 & 6-\lambda \end{vmatrix} = 0$$

$\Rightarrow\ (\lambda-1)(\lambda+2)(\lambda-21) = 0 \Rightarrow \lambda = 1,\ -1, 12.$

Practice Exercise

Verify if each system of linear equations is consistent or not. Solve where possible over the field R of real numbers.

1. $x + y + z = -2$
$2x + 3y - 4z = -3$
$y - 6z = 1$

2. $x + y - 5z = -1$
$2x - y - z = -8$
$3x + 5y + 21z = 1$

3. $2x + 3y - 13z = 0$
$3x + y - 9z = -7$
$x + 2y - 8z = 1$

4. $x + 2y - 4z = 3$
$2x - y + 7z = -4$
$3x + 4y - 6z = 5$

5. $3x + y + z = 8$
$2x - 4y = -10$
$x + y - 3z = 1$

6. $2z - y + z = 3$
$x - y + 2z = -1$
$3x + y - z = 2$

7. $5x + 2y + 7z = 1$
$-3x + 4y + z = 5$
$x + 2y + 3z = -1$

8. $2x - 3y = 5$
$x + 3y + 2z = -1$
$2x + 4z = 2$

9. $x + 2y - 3z = 6$
$2x - y + 4z = 2$
$4x + 3y - 2z = 14$

10. $x + y + z = -2$
$2x + 3y - 4z = -3$
$y - 6z = 1$

11. $x + y + z = 4$
$2x + 5y + 2z = 3$
$3x + 2y - z = 5$

12. $x + 2y + 3z = 0$
$x + 2y + 4z = 2$
$x + 2y + 5z = 5$

13. $x + 2y + 3z = 1$
$-x + 2y - 3z = -1$
$4x + y - 2z = 0$

14. Determine the values of k such that the following system of linear equations.

$$\begin{aligned} x - 3z &= -3 \\ 2x + ky - z &= -2 \\ x + 2y - kz &= 1 \end{aligned}$$

has (i) a unique solution

(ii) no solution

(iii) more than one solution, over R

15. Determine the condition on a, b, c such that the system in unknowns x, y, z has a solution over R.

$$\begin{aligned} x - 2y + 4z &= a \\ 2x + 3y - z &= b \\ 3x + y + 2z &= c \end{aligned}$$

16. Prove that two systems of linear equations have the same solution space if and only if their augmented matrices are row-equivalent.

Summary
The coefficient and augmented matrices of a non-homogeneous system of linear equations are defined. Conditions are then obtained for the consistency or otherwise, of such a non-homogeneous system. A general solution of any consistent non-homogeneous system is also given.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.

2. Ayres, F. *Matrices*, Schaum's Outline Series.

3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.

4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

5. Kuku, A.O. *Abstract Algebra,* Ibadan University Press 1980. Chapter 5.

LECTURE TWELVE

Eigenvalues and Eigenvectors

Introduction

We shall study how to find the invariant lines of a linear transformation $T : V_n(F) \to V_n(F)$ by considering the characteristic polynomial, equation and roots of the matrix defining T. We shall also prove the important fact, known as Cayley-Hamilton theorem, that every square matrix satisfies its characteristic equation.

Objectives

The reader should be able to

(i) determine the eigenvalues and eigenvectors of a linear transformation of a vector space to itself, and

(ii) prove and verify the Cayley-Hamilton theorem for any square matrix.

Pre-Test

Find the characteristic roots and the corresponding characteristic vectors of each of the following matrices over R.

1. $\begin{pmatrix} 1 & 4 \\ 0 & -2 \end{pmatrix}$

2. $\begin{pmatrix} -9 & -8 & -16 \\ 4 & 3 & 2 \\ 4 & 4 & 7 \end{pmatrix}$

Find the eigenvalues and the eigenvectors of each matrix.

3. $\begin{pmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{pmatrix}$

4. $\begin{pmatrix} 1 & -1 & -1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$

5. Compute the eigenvalues and the eigenvectors of the linear transformation. $T : R^3 \to R^2$, with $T(x, y, z) = (x + 4y - z,\ y + z,\ x + y - 2z)$

Verify the Cayley-Hamilton theorem for each matrix over R.

6. $\begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$

7. $\begin{pmatrix} 2 & -2 & 3 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{pmatrix}$

8. $\begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 3 & 1 & 1 \end{pmatrix}$

Definitions

The eigenvalue problem is the following:

Given a linear transformation $T : V_n(F) \to V_n(F)$, determine those scalars $\lambda \in F$ and those non-zero vectors $\underline{v} \in V_n(F)$ which satisfy the equation.

$$T(\underline{v}) = \lambda \underline{v}.$$

λ is called an *eigenvalue* or *characteristic value* of T and the vector $\underline{v}$ is called an *eigenvector* or *characteristic vector* of T corresponding to the eigenvalue λ.

If $A \in M_n(F)$ is an$(n \times n)$-matrix, a scalar $\lambda \in F$ for which there is some non-zero column n-vector $\underline{x}$ (with entries from F) such that $A\underline{x} = \lambda\underline{x}$ is said to be an eigenvalue of A and $\underline{x}$ is said to be an *eigenvector* of A.

If matrix $A \in M_n(F)$ corresponds to the linear transformation T, then their eigenvalues and eigenvectors coincide. Note that eigenvectors are the fixed directions or invariant lines of the linear transformation T. Also if $\underline{x}$ is an eigenvector, then all the vectors in the 1-dimensional vector space over F generated by $\underline{x}$ are also eigenvectors.

Since $T(\underline{v}) = \lambda\underline{v}$ is equivalent to $(T - \lambda I)\underline{v} = \underline{0}$ where I represents the identity linear transformation on $V_n(F)$, it follows that an eigenvector $\underline{v}$ of T belongs to the null-space of the linear transformation $T - \lambda I$. The set of all eigenvectors of T corresponding to the eigenvalue λ is called the *eigenspace* of T corresponding to the eigenvalue λ. It is therefore, the null-space of $T - \lambda I$ and so a subspace of $V_n(F)$.

Similarly, $A\underline{x} = \lambda\underline{x}$ is equivalent to $(A - \lambda I)\underline{x} = \underline{0}$, where I is the identity matrix of order n, and so the set of all eigenvectors of matrix A corresponding to the eigenvalue λ is called the eigenspace or row null-space of A corresponding to the eigenvalue λ.

$(A - \lambda I)\underline{x} = \underline{0}$ corresponds to a homogeneous system of linear equations. Besides the trivial solution $\underline{x} = \underline{0}$, non-trivial solutions exist if $|A - \lambda I| = 0$, i.e. if matrix $A - \lambda I$ is singular. This last equation is called the *characteristic equation* of the matrix A and the n roots $\lambda_1, \ldots, \lambda_n$ are the *characteristic roots* or *eigenvalues* of A. We refer to the polynomial $|A - \lambda I|$ in λ, as the *characteristic* polynomial of A and we denote it by $f(\lambda)$ or $C_A(\lambda)$.

To find the characteristic polynomial or characteristic equation of a linear transformation T, one first finds the matrix representing T with respect to a basis for $V_n(F)$. However, the characteristic polynomial of T is independent of the basis used to compute its corresponding matrix. Note that

$$C_A(\lambda) = |A - \lambda I| = (-1)^n\lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0$$

Proposition **(Cayley-Hamilton Theorem)**
Every square matrix A satisfies its characteristic equation i.e. $C_A(A) = 0$.

Proof. Let

$$C_A(\lambda) = |A - \lambda I| = \sum_{i=0}^{n} b_i \lambda^i$$

Also let $D(\lambda)$ be the adjoint of $A - \lambda I$. The entries of $D(\lambda)$ are cofactors of the matrix $A - \lambda I$ and hence are polynomials in λ of degree not exceeding $n - 1$. Thus we may put

$$D(\lambda) = \sum_{i=0}^{n-1} D_i \lambda^i$$

where the D_i are $(n \times n)$-matrices with constant entries which are independent of λ. Since

$$(A - \lambda I) \cdot D(\lambda) = |A - \lambda I| \cdot I$$

we obtain that

$$(A - \lambda I)(D_{n-1}\lambda^{n-1} + \cdots + D_1\lambda + D_0) = (b_n\lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0)I$$

Equate coefficients of corresponding powers of λ

$$\begin{aligned}
&-D_{n-1} = b_n I \\
&AD_{n-1} - D_{n-2} = b_{n-1} I \\
&AD_{n-2} - D_{n-3} = b_{n-2} I \\
&\vdots \\
&AD_1 - D_0 = b_1 I \\
&AD_0 = b_0 I
\end{aligned}$$

Multiply the above matrix equations by A^n, $A^{n-1}I, \ldots, A, I$, respectively:

$$\begin{aligned}
&-A^n D_{n-1} = b_n A^n \\
&A^n D_{n-1} - A^{n-1} D_{n-2} = b_{n-1} A^{n-1} \\
&A^{n-1} D_{n-2} - A^{n-2} D_{n-3} = b_{n-2} A^{n-2} \\
&\vdots \\
&A^2 D_1 - AD_0 = b_1 A \\
&AD_0 = b_0 I.
\end{aligned}$$

Add the above matrix equations and obtain that

$$0 = b_n A^n + b_{n-1} A^{n-1} + \cdots + b_1 A + b_0 I$$

This implies that $C_A(A) = 0$.

Example.
Consider the matrix

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

(a) Find its characteristic polynomial

(b) Verify the Cayley-Hamilton theorem for A

(c) Determine the eigenvalues and eigenvectors of A.

(a) The characteristic polynomial is given by

$$\begin{aligned} C_A(\lambda) &= |A - \lambda I = \begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{vmatrix} \\ &= -\lambda^3 + 6\lambda^2 - 9\lambda + 4 = (\lambda - 1)^2(4 - \lambda) \end{aligned}$$

(b) To verify the Cayley-Hamilton Theorem for A, we have to show that $C_A(A) = 0$, i.e. show

$$-A^3 + 6A^2 - 9A + 4I = 0$$

or

$$(A - I)^2(4I - A) = 0$$

$$(A - I)^2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{pmatrix}$$

Hence,

$$\begin{aligned} L.H.S. &= (A - I)^2(4I - A) \\ &= \begin{pmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{pmatrix} \begin{pmatrix} -2 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= R.H.S. \end{aligned}$$

(c) The characteristic roots are the roots of the characteristic equation. Therefore

$$(\lambda - 1)^2(4 - \lambda) = 0 \Rightarrow \lambda = 1, 1, 4$$

i.e. the characteristic roots of A are $\lambda = 1, 1, 4$.
Next find the characteristic vectors, which are bases for the eigenspace corresponding to $\lambda = 1, 1, 4$.
$\lambda = 1$: Since the eigenspace of A corresponding to $\lambda = 1$ is equal to the row null-space of $A - I$, it is therefore, equal to the solution space of the homogeneous system of linear equations.

$$(A - I)\underline{x} = \underline{0}$$

First we obtain the row-reduced echelon form of $A - I$.

$$A - I = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{matrix} R_{21}(-1) \\ \sim \\ R_{31}(-1) \end{matrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

From the row-reduced echelong form of $A - I$, we have $x + y + z = 0$. Put $z = a, y = b$. Then $x = -a - b$. Hence the solution space of the homogeneous system $(A - I)\underline{x} = \underline{0}$ is

$$\{(x, y, z) = a(-1, 0, 1) = b(-1, 1, 0)|a, b \in R)\}$$

i.e., a basis for the eigenspace of A corresponding to $\lambda = 1$ is

$$\{(-1, 0, 1), (-1, 1, 0)\}$$

$\lambda = 4$: The eigenspace of A corresponding to $\lambda = 4$ is equal to the solution space of the homogeneous system of linear equation $(A - 4)\underline{x} = \underline{0}$. First we obtain the row-reduced echelon form of $A - 4I$.

$$A - 4I \begin{pmatrix} 2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -1 \end{pmatrix} \begin{matrix} R_1(-\frac{1}{2}) \\ \sim \\ R_{21}(\frac{1}{2}) \\ R_{31}(\frac{1}{2})) \end{matrix} \begin{pmatrix} 1 & -\dfrac{1}{2} & \dfrac{1}{2} \\ 0 & -\dfrac{3}{2} & \dfrac{3}{2} \\ 0 & \dfrac{3}{3} & -\dfrac{3}{2} \end{pmatrix}$$

$$\begin{matrix} R_{21}(-\frac{2}{3}) \\ \sim \\ R_{32}(1) \end{matrix} \begin{pmatrix} 1 & -\dfrac{1}{2} & -\dfrac{1}{2} \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

From the row-reduced echelon form of $A - 4I$, we have

$$\begin{aligned} x - \frac{1}{2}y - \frac{1}{2}z &= 0 \\ y - z &= 0 \end{aligned}$$

Put $z = a$. Then $y = a$ and $x = \dfrac{1}{2}a + \dfrac{1}{2}a = a$.
Hence the solution space of the homogeneous system $(A - 4I)\underline{x} = \underline{0}$ is

$$\{(x, y, z) = a(1, 1, 1)|a \in R\}$$

i.e. a basis for the eigenspace of A corresponding to $\lambda = 4$ is $\{(1, 1, 1)\}$.
Thus, a basis for the characteristic vectors of A is

$$\{(-1, 0, 1), (-1, 1, 0), (1, 1, 1)\}$$

Practice Exercise

1. Find the characteristic roots and the corresponding characteristic vectors of the matrix.
$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$$

2. Find the eigenvalues and the eigenvectors of the matrix
$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

3. Compute the eigenvalues and the eigenvectors of the linear transformation
$$T : R^2 \to R^2, \text{ with } T(a, b) = (b - 3a, 18a + 4b)$$
Verify the Cayley-Hamilton theorem for each matrix over R.

4. $\begin{pmatrix} -3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

5. $\begin{pmatrix} 2 & 0 & 0 \\ 1 & -3 & 0 \\ 2 & -10 & 2 \end{pmatrix}$

Summary
The eigenvalues (or characteristic values) and eigenvectors (or characteristic vectors) of linear transformations and matrices are defined. By calculating the characteristic polynomial and equation, its characteristic roots and the solution of the resultant homogeneous system of linear equations, we are able to determine the eigenvalues and eigenvectors of a linear transformation or a matrix. It is also shown that every square matrix satisfies its characteristic equation, a fact known as the Cayley-Hamilton Theorem.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.

2. Ayres, F. *Matrices*, Schaum's Outline Series.

3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.

4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

LECTURE THIRTEEN

The Minimal Polynomial

Introduction

We wish to consider the set M of all polynomials $f(x)$ over a field F such that $f(T) = 0$ or $f(A) = 0$, where T is a linear transformation $T : V_n(F) \to V_n(F)$ and A is a square matrix. The minimal polynomial of T or A is then a generator of M since M is an ideal of the polynomial ring $F[x]$ which is a principal ideal domain.

In Unit 12, we have found that the characteristic polynomial $C_A(x) \in M$ since by the Cayley-Hamilton Theorem, $C_A(A) = 0$. We shall study how to obtain the minimal polynomial and its application to the computation of matrix polynomials and identities as well as matrix inverses in the case of non-singular matrices.

Objectives

The reader should be able to

(i) determine the minimal polynomial of any square matrix or linear transformation of a vector space to itself, and

(ii) apply the minimal polynomial to the computation of matrix polynomials, identities and matrix inverses in the case of non-singular matrices.

Pre-Test
Find the minimal polynomial of each matrix

1. $\begin{pmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{pmatrix}$ (2) $\begin{pmatrix} 2 & 0 & 0 \\ 1 & -3 & 0 \\ 2 & -10 & 2 \end{pmatrix}$

3. Consider the matrix

$$B = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{pmatrix}$$

By using the minimal polynomial:

(i) compute $3B^3 + 2A^2 - 5I$

(ii) show that $16B^{-3} - 26B^{-2} + 11B^{-1} - I = 0$

(iii) find B^{-1}

4. Consider the matrix

$$C = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$$

By using the minimal polynomial

(i) compute $2C^6 - C^5 - C^4 + 2C^2 - I$

(ii) show that $6C^{-3} - 11C^{-2} + 6C^{-1} - I = 0$

(iii) find C^{-1}.

Matrix Polynomials
Every $(n \times n)$-matrix $A \in M_n(F)$ yields polynomials.
$f(A) = a_0 I + a_1 A + \cdots + a_k A^k, (a_i \in F)$ in A which are again members of $M_n(F)$. Since there are exactly n^2 linearly independent $(n \times n)$-matrices over F (i.e. $\dim M_n(F) = n^2$), the $n^2 + 1$ matrices I, A, $A^2, \ldots, A^{n^2}$ are linearly

dependent and the dependence relation provides a non-zero polynomial $f(x)$ of degree at most n^2 with $f(A) = 0$.

Since there is an F isomorphism between $M_n(F)$ and the vector space $\mathcal{L}(V_n(F), V_n(F))$ of all linear transformations from $V_n(F)$ to $V_n(F)$, we have that for every linear transformation T of $V_n(F)$, there exists a non-zero polynomial $f(x)$ with $f(T) = 0$.

Now consider the set M of all polynomials $f(x)$ over F such that $f(T) = 0$ or $f(A) = 0$. Then M is an ideal of $F[x]$ i.e. M is an Abelian group under addition and M is closed under multiplication by any polynomial in $F[x]$. Thus $M = \langle m(x) \rangle$ is an ideal generated by a unique monic polynomial $m(x)$ (since $F[x]$ is a principal ideal domain where every ideal is principal).

Definition
The minimal polynomial of a linear transformation T, denoted by $m(x)$, is the generator of the ideal of the polynomials $f(x)$ over F satisfying $f(T) = 0$. Thus $m(x)$ satisfies the properties

$$m(T) = 0 \text{ and } f(T) = 0 \Rightarrow m(x) | f(x).$$

The minimal polynomial of a matrix is defined in a similar way. It is identical with the minimal polynomial of the corresponding linear transformation T_A of $V_n(F)$.

By the Cayley-Hamilton Theorem considered in Unit 12, since $C_A(A) = 0$, if follows that the minimal polynomial of A divides the characteristic polynomial of A, i.e. $m(x)|C_A(x)$. This fact then provides a method of determining the minimal polynomial.

The minimal polynomial may also be found as follows. One tries to solve the equation (beginning with $i = 1$).

$$A^i = \alpha_{i-1} A^{i-1} + \cdots + a_1 A + a_0 I, \quad i = 1, 2, 3, \ldots$$

for the coefficients $\alpha_{i-1}, \ldots, \alpha_1, \alpha_0$. The first i for which solutions exist provides the minimal polynomial which is then

$$m(x) = x^i - \alpha_{i-1} x^{i-1} - \cdots - \alpha_1 x - \alpha_0.$$

The Companion Matrix
For each monic polynomial

$$f(x) = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1} + x^n$$

of degree n, we can construct an $(n \times n)$-matrix with minimal polynomial $f(x)$. This matrix is called the companion matrix of $f(x)$ and is given by

$$C_f = \begin{pmatrix} 0 & 1 & & & & & & \\ & 0 & 1 & & & & & \\ & & 0 & 1 & & & & \bigcirc \\ & & & 0 & 1 & & & \\ & & & & \ddots & \ddots & & \\ & & & & & 0 & 1 & \\ \bigcirc & & & & & & 0 & 1 \\ -c_0 c_1 & & \cdots & & & & & -c_n - 1 \end{pmatrix}$$

with zero entries except for entries 1 in the diagonal just above the main diagonal and entries $-c_0, -c_1, \ldots, -c_{n-1}$ in the last row. We then have the following fact:

For each monic polynomial $f(x)$, the companion matrix C_f has minimal polynomial $f(x)$ and characteristic polynomial $(-1)^n f(\lambda)$.
Note that $|C_f - \lambda I| = (-1)^n f(\lambda)$, with $(-1)^n \lambda^n$ as its leading term.

Example 1. Find the minimal polynomial of the matrix over R

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

The characteristic polynomial of A is given by

$$\begin{aligned} C_A(\lambda) &= |A - \lambda I| = \begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{vmatrix} \\ &= -\lambda^3 + 6\lambda^2 - 9\lambda + 4 = (\lambda - 1)^2(4 - \lambda) \end{aligned}$$

Since the minimal polynomial $m(x)$ of A must divide the characteristic polynomial $C_A(x)$ of A, it follows that $m(x)$ is one of the divisors of $(x-1)^2(4-x)$, namely.

$$x-1, x-4, (x-1)^2, (x-1)(x-4), -C_A(x).$$

Note that $m(x)$ must be a monic polynomial i.e. its leading coefficient must be equal to $+1$. By testing each of the above divisors to know which one gives 0 when $x = A$, we find that the minimal polynomial is $m(x) = (x-1)(x-4) = x^2 - 5x + 4$ since

$$(A-I)(A-4I) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Aliter
The minimal polynomial may also be obtained as follows.
First try to solve $A = \alpha_0 I$. There is no solution. Next try to solve $A^2 = \alpha_1 A + \alpha_0 I$. Then $\alpha_1 = 5$, $\alpha_0 = -4$ are solutions.
Hence the minimal polynomial is

$$m(x) = x^2 - 5x + 4.$$

Example 2.
Consider the matrix

$$B = \begin{pmatrix} -6 & -2 & 2 \\ -2 & 3 & 3 \\ 2 & -1 & 3 \end{pmatrix}$$

By using the minimal polynomial

(i) compute $B^5 - 2B^4 + B - 4I$,

(ii) show that $16B^{-3} + 6B^{-2} - 9B^{-1} + I = 0$, and

(iii) find B^{-1}

The minimal polynomial of B can be found to be

$$m(x) = (x-2)(x-8) = x^2 - 10x + 16$$

(i) First divide $x^5 - 2x^4 + x - 4$ by $m(x)$ and find the remainder.
We obtain

$$x^5 - 2x^4 + x - 4 = (x^2 - 10x + 16)(x^3 + 8x^2 + 64x + 512) + 4097x - 8196$$

Since $m(B) = B^2 - 10B + 161 = 0$, we have that

$$B^5 - 2B^4 + B - 4I = 4097B - 8196 = \begin{pmatrix} 16386 & -8194 & 8194 \\ -8194 & 4095 & -4097 \\ 8194 & -4097 & 4095 \end{pmatrix}$$

(ii) Consider $16 + 6B - 9B^2 + B^3$ or $x^3 - 9x^2 + 6x + 16$.
Divide $x^3 - 9x^2 + 6x + 16$ by $m(x)$ and obtain

$$x^3 - 9x^2 + 6x + 16 = (x^2 - 10x + 16)(x + 1)$$

Since

$$m(B) = B^2 - 10B + 161 = 0$$
$$B^3 - 9B^2 + 6B + 161 = 0$$

$\Leftrightarrow \quad I - 9B^{-1} + 6B^{-2} + 16B^{-3} = 0$, multiplying both sides by B^{-3}.

(iii) $m(B) = B^2 - 10B + 161 = 0$
Multiply by B^{-1}: $B - 10I + 16B^{-1} = 0$

$$\Rightarrow B^{-1} = \frac{1}{16}(10I - B) = \begin{pmatrix} \frac{1}{4} & \frac{1}{8} & -\frac{1}{8} \\ \frac{1}{8} & \frac{7}{16} & \frac{1}{16} \\ -\frac{1}{8} & \frac{1}{16} & \frac{7}{16} \end{pmatrix}$$

Remark
The above example (part (iii) gives a third method for finding the inverse of

a non-singular matrix. The other two methods considered in Unit 3 are the adjoint method and the method of elementary operations.

Practice Exercise
Find the minimal polynomial of each matrix

1. $\begin{pmatrix} -3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ (2) $\begin{pmatrix} 1 & -1 & -1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$

3. $\begin{pmatrix} -9 & -8 & -16 \\ 4 & 3 & 2 \\ 4 & 4 & 7 \end{pmatrix}$

4. Consider the matrix
$\begin{pmatrix} 2 & 0 & 0 \\ 1 & -3 & 0 \\ 2 & -10 & 2 \end{pmatrix}$

By using the minimal polynomial

(i) compute $A^5 - 2A^4 + A - 4I$

(ii) show that $6A^{-3} + 5A^{-2} - 14A^{-1} + I + 2A = 0$

(iii) find A^{-1}.

Summary
It is shown that the set of all polynomials $f(x)$ over a field F such that $f(T) = 0$ or $f(A) = 0$ is an ideal of the principal ideal domain $F[x]$. This ideal is then generated by a monic polynomial which is called the minimal polynomial of T or A. The minimal polynomial is then calculated using a direct method or the Cayley-Hamilton Theorem and it is then applied to the computation of matrix polynomials, identities and matrix inverses in the case of non-singular matrices.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.

2. Ayres, F. *Matrices*, Schaum's Outline Series.

3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.

4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

LECTURE FOURTEEN

Similar and Diagonal Matrices

Introduction

The fact that the matrix representing a linear transformation depends on the bases used indicates that a linear transformation may be represented by several different matrices which are referred to as similar matrices. The question that naturally arises is, what are the common properties of these matrices. This question will be answered in this Unit. It is also of particular interest to determine what diagonal matrix, if any, is similar to a given matrix, since the algebra of diagonal matrices is very simple.

Objectives:

The reader should be able to:

(i) State and prove the properties common to similar matrices.

(ii) obtain a diagonal matrix similar to a given matrix.

Pre-Test

1. Show that the relation of similarity among matrices in $M_n(F)$ is an equivalence relation.

2. Diagonalise the real matrix.

$$\begin{pmatrix} -9 & -8 & -16 \\ 4 & 3 & 8 \\ 4 & 4 & 7 \end{pmatrix}$$

3. Show that the real matrix

$$\begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{pmatrix}$$

   is reducible to a diagonal form under similarity transformations. Exhibit the diagonal form and the relevant non-singular matrix of transition.

4. Find a matrix P over R such that $P^{-1}AP$ is a diagonal matrix, where

$$A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$$

   and exhibit the diagonal form.

5. Show that similar matrices form a multiplicative group.

Definition:

Two square matrices A and B in $M_n(F)$ are *similar* if there is a non-singular matrix P in $GL_n(F)$ such that $B = P^{-1}AP$.

One can show that the relation of similarity among matrices in $M_n(F)$ is an equivalence relation satisfying the reflexive, symmetric and transitive laws (the RST laws) and partitioning the set of similar matrices into disjoint equivalence classes.

The relationship between the matrix representation, A_T of a linear transformation, $T : V_n(F) \to V_n(F)$ with respect to a basis $\{\underline{v}_1, \ldots, \underline{v}_n\}$ and its matrix representation B_T with respect to another basis $\{\underline{v}'_1, \ldots, \underline{v}'_n\}$ is that

$$B_T = P^{-1} A_T P$$

where P is the non-singular matrix associated with the change of coordinates from the basis $\{\underline{v}_1, \ldots, \underline{v}_n\}$ to the basis $\{\underline{v}'_1, \ldots, \underline{v}'_n\}$, i.e.

$$P = (P_{ij})_{n,n'} \text{ where } \underline{v}'_j = \sum_{i=1} P_{ij}\underline{v}_i.$$

It follows that two similar matrices represent the same linear transformation with respect to two bases which may be different. We then have the following facts about similar matrices.

Proposition 1
Similar matrices in $M_n(F)$ have the same

(a) determinant;

(b) characteristic roots;

(c) characteristic polynomial;

(d) minimal polynomial in $F[x]$; and

(e) rank.

Proof: Let A and B be similar matrices in $M_n(F)$. Hence there exists a non-singular matrix P in $GL_n(F)$ such that $B = P^{-1}AP$.

(a)
$$\begin{aligned} B = P^{-1}AP \;&\Rightarrow\; PB = AP \\ &\Rightarrow\; \det(P)\cdot\det(B) = \det(A)\cdot\det(P) \text{ in } F \\ &\Rightarrow\; \det(B) = det(A) \end{aligned}$$

(b) Let λ be a characteristic root of A such that $\underline{v} \neq \underline{0}$ is its characteristic vector. Then $A\underline{v} = \lambda\underline{v}$.
If $B = P^{-1}AP$ is similar to A, then

$$A = PBP^{-1} \Rightarrow (PBP^{-1})\underline{v} = \lambda\underline{v} = B(P^{-1}\underline{v}) = \lambda(P^{-1}\underline{v}).$$

Since $\underline{v} \neq \underline{0}$ and P is non-singular, it follows that $P^{-1}\underline{v} \neq \underline{0}$. This implies that $P^{-1}\underline{v}$ is a characteristic vector of B belonging to the same characteristic root λ. Similarly any characteristic root of B is also a characteristic root of A.

(c) Let $B = P^{-1}AP$ be similar to A. Since $P^{-1}P = I$, then $|P^{-1}| \cdot |P| = 1$, which implies that $|P^{-1}| = |P|^{-1}$.
Now since $|P^{-1}| = |P|^{-1}$ and $|P|$ are scalars in F, they commute and so the characteristic polynomial of B is given by

$$\begin{aligned} |B - \lambda I| &= |P^{-1}AP - \lambda I| = |P^{-1}AP - \lambda P^{-1}IP| \\ &= |P^{-1}(A - \lambda I)P| = |P^{-1}| \cdot |A - \lambda I| \cdot |P| \\ &= ||A - \lambda I| \\ &= \text{characteristic polynomial of } A. \end{aligned}$$

(d) Since $B = P^{-1}AP$ implies that $B^i = P^{-1}A^iP$, $i \in N$, it follows that a polynomial satisfied by A is also satisfied by B, and vice versa, i.e.

$$\begin{aligned} \sum a_i A^i &= 0 \Leftrightarrow \sum a_i B^i = \sum a_i P^{-1}A^iP \\ &= P^{-1}\left(\sum a_i A^i\right)P = P^{-1} \cdot 0 \cdot P = 0 \end{aligned}$$

Hence it follows that they must have the same minimal polynomials.

(e) Since the rank of a matrix is not affected by pre- or post-multiplication by a non-singular matrix, it follows that similar matrices have the same ranks.

Diagonal Matrices
The algebra of diagonal matrices is very easy. To add or multiply two diagonal matrices, we just add or multiply corresponding diagonal matrices.

Thus it is of interest to know which matrices are similar to diagonal matrices and which pairs of diagonal matrices are similar to each other. We have the following results.

Proposition 2
A matrix A in $M_n(F)$ is similar to a diagonal matrix D if and only if the characteristic vectors of A span $V_n(F)$; and if this is the case, the characteristic roots of A are the diagonal entries in D. In particular, it follows that the characteristic roots of a diagonal matrix are the entries on the diagonal.

Proof
$\underline{\Rightarrow}$: Assume that a matrix A in $M_n(F)$ is similar to a diagonal matrix D given by

$$D = \text{diagonal } (d_1, \ldots, d_n)$$

Then consider $\underline{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, 1 in the i-th coordinate $i = 1, \ldots, n$. Since $D\underline{e}_i = d_i\underline{e}_i,\ i = 1, \ldots, n$, it follows that the unit vector $\underline{e}_i$ are characteristic vectors for D and the diagonal entries $d_1, \ldots, d_n$ are the corresponding characteristic roots of D and hence of A. Since $A = P^{-1}DP$ and $D\underline{e}_i = d_i\underline{e}_i$ imply that $A(P^{-1}\underline{e}_i) = d_iP^{-1}\underline{e}_i$, we have that the characteristic vectors $P^{-1}\underline{e}_i, \ldots, P^{-1}\underline{e}_n$ of A are linearly independent (P and P^{-1} being non-singular) and so span $V_n(F)$. (Note that since A and D are similar, A and D have the same characteristic roots).

$$\begin{aligned} T_A(\underline{v}_i) &= A\underline{v}_i, \text{ (by definition of } T_A) \\ &= \lambda_i\underline{v}_i,\ i = 1, \ldots, n \end{aligned}$$

for characteristic roots $\lambda, \ldots, \lambda_n$. Thus the linear transformation T_A is represented, relative to this basis, by the diagonal matrix.

$$D = \text{doaginal } (\lambda_1, \ldots, \lambda_n),\ \text{since } T_A(\underline{v}_i) = D\underline{v}_i$$

In other words T_A is represented by either of the matrices A or D and hence A is similar to D.

Proposition 3
If matrices A and B in $M_n(F)$ have the same set of n distinct characteristic roots, then they are similar.

Proof.
Let the distinct characteristic roots be $\lambda_1, \ldots, \lambda_n$. We shall show that the corresponding vectors $\underline{x}_1, \ldots, \underline{x}_n$ are linearly independent and so span $V_n(F)$. We shall show this by induction. When $n = 2$, we have that

$$A\underline{x}_1 = \lambda_1 \underline{x}_1 \text{ and } A\underline{x}_2 = \lambda_2 \underline{x}_2$$

and suppose that $\underline{x}_1$ and $\underline{x}_2$ are linearly dependent so that $\underline{x}_1 = \alpha \underline{x}_2$ for some non-zero scalar $\alpha \in F$. Then

$$A(\alpha \underline{x}_2) = \lambda_1(\alpha \underline{x}_2) \Leftrightarrow A\underline{x}_2 = \lambda_1 \underline{x}_2$$

Now,

$$A\underline{x}_2 = \lambda_1 \underline{x}_2 \text{ and } A\underline{x}_2 = \lambda_2 \underline{x}_2 \Rightarrow \lambda_1 = \lambda_2$$

This is a contradiction to the fact that the roots are distinct. Hence it must be that $\underline{x}_1, \underline{x}_2$ are linearly independent.

Next assume as an inductive hypothesis that the corresponding vectors $\underline{x}_1, \ldots, \underline{x}_{n-1}$ to the distinct roots $\lambda_1, \ldots, \lambda_{n-1}$ are linearly independent. Now consider the characteristic vectors $\underline{x}_1, \ldots, \underline{x}_n$ corresponding to the distinct roots $\lambda_1, \ldots, \lambda_n$. Assume that the vectors $\underline{x}_1, \ldots, \underline{x}_n$ are linearly dependent. Then we can express $\underline{x}_n$ as a linear combination of $\underline{x}_1, \ldots, \underline{x}_{n-1}$ as

$$\underline{x}_n = \sum_{i=1}^{n-1} \alpha_i \underline{x}_i, \alpha_i \in F, \ i = 1, \ldots, n-1.$$

Now we know that $A\underline{x}_i = \lambda_i \underline{x}_i, \ \ i = 1, \ldots, n.$
Hence $A\underline{x}_n = \lambda_n \underline{x}_n$ implies that

$$\begin{aligned} & A\left(\sum_{i=1}^{n-1} \alpha_i \underline{x}_i\right) = \lambda_n \left(\sum_{i=1}^{n-1} \alpha_i \underline{x}_i\right) \\ \Leftrightarrow \ & \sum_{i=1}^{n-1} (\alpha_i \lambda_i) \underline{x}_i = \sum_{i=1}^{n-1} (\lambda_n \alpha_i) \underline{x}_i \\ \Leftrightarrow \ & \sum_{i=1}^{n-1} (\alpha_i \lambda_i - \lambda_n \alpha_i) \underline{x}_i = \underline{0} \end{aligned}$$

Since $\underline{x}_1, \ldots, \underline{x}_{n-1}$ are linearly independent, then the only solutions of the last vector equation are

$$\alpha_i \lambda_i - \lambda_n \alpha_i = 0, \;\; i = 1, \ldots, n-1$$

$$\Rightarrow \;\; \lambda_n = \lambda_i \;\text{ for at least one } i \in \{1, \ldots, n-1\}$$

This is a contradiction to the fact that the roots are all distinct. Hence it must be that $\underline{x}_1, \ldots, \underline{x}_n$ are linearly independent. Therefore, it follows that A and B are similar to the same diagonal matrix D, by Proposition 2 above. Since the relation of similarity is an equivalence relation, it follows that A and B are similar matrices.

Proposition 4
If a matrix P in $M_n(F)$ is a matrix whose columns are linearly independent characteristic vectors of a matrix A in $M_n(F)$, then P is non-singular and $P^{-1}AP$ is a diagonal matrix.

Proof. Let the column vectors $\underline{v}_1, \ldots, \underline{v}_n$ be the characteristic vectors of A such that

$$A\underline{v}_i = \lambda_i \underline{v}_i, \;\; i = 1, \ldots, n$$

where $\lambda_1, \ldots, \lambda_n$ are characteristic roots of A. The matrix P with columns $\underline{v}_1, \ldots, \underline{x}_n$ is non-singular since the columns are linearly independent. Hence

$$\begin{aligned} AP &= A(\underline{v}_1, \ldots, \underline{v}_n) = (\lambda_1 \underline{v}_1, \ldots, \lambda_n \underline{v}_n) \\ &= (\underline{v}_1, \ldots, \underline{v}_n) \text{ diagonal } (\lambda_1, \ldots, \lambda_n) \\ &= PD, \text{ where } D = \text{diagonal } (\lambda_1, \ldots, \lambda_n) \end{aligned}$$

Hence $D = P^{-1}AP$, which implies that $P^{-1}AP$ is a diagonal matrix.

Remarks.

1. To construct, (if it exists), a diagonal matrix similar to a given matrix, one computes the characteristic roots and vectors.

2. It can also be shown that a matrix is similar to a diagonal matrix if and only if the linear factors of its minimal polynomial are distinct.

3. There are matrices which are NOT similar to any diagonal matrix.

   e.g. $\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$, $c \neq 0$ over R.

Example

Find a matrix P over the field R of real numbers such that $P^{-1}AP$ is a diagonal matrix where

$$A = \begin{pmatrix} 8 & -8 & -2 \\ 4 & -3 & -2 \\ 3 & -4 & 1 \end{pmatrix} \in M_3(R)$$

and exhibit the diagonal form.

First obtain the characteristic polynomial of A.

$$|A - \lambda I| = \begin{vmatrix} 8-\lambda & -8 & -2 \\ 4 & -3-\lambda & -2 \\ 3 & -4 & 1-\lambda \end{vmatrix} \begin{array}{l} = -\lambda^3 + 6\lambda^2 - 11\lambda + 6 \\ = (\lambda - 1)(\lambda - 2)(3 - \lambda) \end{array}$$

This implies that the characteristic roots of A are the roots of the characteristic equation,

$$(\lambda - 1)(\lambda - 2)(3 - \lambda) = 0$$

Hence the characteristic roots of A are 1,2,3.

Next, find the characteristic vectors which form bases for the eigenspaces corresponding to $\lambda = 1, 2, 3$.

$\lambda = I$: Since the eigenspace of A corresponding to $\lambda = 1$ is equal to the row null-space of $A - I$, it is therefore, equal to the solution space of the homogeneous system of linear equations $(A - I)\underline{x} = \underline{0}$. Obtain the row-reduced echelon form of $A - I$.

$$A-I = \begin{pmatrix} 7 & -8 & -2 \\ 4 & -4 & -2 \\ 3 & -4 & 0 \end{pmatrix} \overset{R_1(\frac{1}{7})}{\sim} \begin{pmatrix} 1 & -\frac{8}{7} & -\frac{2}{7} \\ 4 & -4 & -2 \\ 3 & -4 & 0 \end{pmatrix}$$

$$\underset{R_{31}(-3)}{\overset{R_{21}(-4)}{\sim}} \begin{pmatrix} 1 & -\frac{8}{7} & -\frac{2}{7} \\ 4 & -\frac{4}{7} & -\frac{6}{7} \\ 0 & -\frac{4}{7} & \frac{6}{7} \end{pmatrix} \underset{R_{32}(1)}{\overset{R_2(\frac{7}{4})}{\sim}} \begin{pmatrix} 1 & -\frac{8}{7} & -\frac{2}{7} \\ 0 & 1 & -\frac{3}{2} \\ 0 & 0 & 0 \end{pmatrix}$$

From the row-reduced form of $A - I$, we have

$$\begin{aligned} x - \frac{8}{7}y - \frac{2}{7}x &= 0 \\ y - \frac{3}{2}z &= 0 \end{aligned}$$

Put $z = 2a$. Then $y = 3a$ and $z = \dfrac{24a}{7} + \dfrac{4a}{7} = 4a$. Hence the solution space of the homogeneous system $(A - I)\underline{x} = \underline{0}$ is

$$\{(x, y, z) = a(4, 3, 2) | a \in R\}$$

i.e. a basis for the eigenspace of A correspondence to $\lambda = 1$ is equal to $\{(4, 3, 2)\}$.

$\lambda = 2$: The eigenspace of A corresponding to $\lambda = 2$ is equal to the solution space of the homogeneous system of linear equations $(A - 2I)\underline{x} = \underline{0}$. Obtain the row-reduced echelon form of $A - 2I$.

$$A-2I = \begin{pmatrix} 6 & -8 & -2 \\ 4 & -5 & -2 \\ 3 & -4 & 1 \end{pmatrix} \overset{R_2(\frac{1}{6})}{\sim} \begin{pmatrix} 1 & -\frac{4}{3} & -\frac{1}{3} \\ 4 & -5 & -2 \\ 3 & -4 & -1 \end{pmatrix}$$

$$\overset{R_{21}(-4)}{\underset{R_{31}(-3)}{\sim}} \begin{pmatrix} 1 & -\frac{4}{3} & -\frac{1}{3} \\ 0 & \frac{1}{3} & -\frac{2}{3} \\ 0 & 0 & 0 \end{pmatrix} \overset{R_2(3)}{\sim} \begin{pmatrix} 1 & -\frac{4}{3} & -\frac{1}{3} \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

From the row-reduced form of $A-2I$, we have

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

Put $z = a$. Then $y = 2a$ and $a = \frac{8}{3}a + \frac{1}{3}a = 3a$.
Hence the solution space of the homogeneous system $(A-2I) underline x = \underline{0}$ is

$$\{(x, y, z) = a(3, 2, 1) | a \in R\}$$

i.e. a basis for the eigenspace of A corresponding to $\lambda = 2$ is $\{(4, 3, 2)\}$.
$\lambda = 3$: The eigenspace of A corresponding to $\lambda = 3$ is equal to the solution space of the homogeneous system of linear equations $(A-3I)\underline{x} = \underline{0}$. Obtain the row-reduced echelon form of $A-3I$.

$$A-3I = \begin{pmatrix} 5 & -8 & -2 \\ 4 & -6 & -2 \\ 3 & -4 & -2 \end{pmatrix} \overset{R_1(\frac{1}{5})}{\sim} \begin{pmatrix} 1 & -\frac{8}{5} & -\frac{2}{5} \\ 4 & -6 & -2 \\ 3 & -4 & -2 \end{pmatrix}$$

$$\overset{R_{21}(-4)}{\underset{R_{31}(-3)}{\sim}} \begin{pmatrix} 1 & -\frac{8}{5} & -\frac{2}{5} \\ 0 & -\frac{2}{5} & -\frac{2}{5} \\ 0 & -\frac{4}{7} & \frac{4}{5} \end{pmatrix} \overset{R_2(\frac{5}{2})}{\underset{R_{32}(-2)}{\sim}} \begin{pmatrix} 1 & -\frac{8}{5} & -\frac{2}{5} \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

From the row-reduced echelon form of $A - 3I$, we have

$$\begin{aligned} x - \frac{8}{5}y - \frac{2}{5}x &= 0 \\ y - z &= 0 \end{aligned}$$

Put $z = a$. Then $y = a$ and $x = \frac{8}{5}a + \frac{2}{5}a = 2a$.
Hence the solution space of the homogeneous system $(A - 3I)\underline{x} = \underline{0}$ is

$$\{(x, y, z) = a(2, 1, 1)|a \in R\}$$

i.e. a basis for the eigenspace of A corresponding to $\lambda = 3$ is $\{(2, 1, 1)\}$.
The required matrix is

$$P = \begin{pmatrix} 4 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix}$$

Note that $|P| = -1$ which implies that P^{-1} exists. Hence A is reducible to a diagonal form given by

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Remark: Note that by Proposition 4, we may obtain $P^{-1}AP$, not by direct calculation and substitution for P^{-1}, A and P but by just writing down the characteristic roots 1,2,3 along the principal diagonal in the same order as the corresponding characteristic vectors are written as column vectors of P.

Practice Exercise XIV
Find a matrix P over R such that $P^{-1}AP$ is a diagonal matrix, where

$$A = \begin{pmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ 2 & -4 & 4 \end{pmatrix}$$

and exhibit the diagonal form.

Summary
We defined similar matrices and proved properties possessed by them. We then indicated those matrices which are similar to diagonal matrices.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.
2. Ayres, F. *Matrices*, Schaum's Outline Series.
3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.
4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.

LECTURE FIFTEEN

Triangular Matrices

Introduction

Triangular matrices are important since every square matrix can be expressed as a product of two triangular matrices. Their inverses are also relatively easy to compute thus giving rise to another method of finding the inverse of a non-singular matrix. Triangular matrices are also useful in the solution of systems of linear equation $A\underline{x} = \underline{y}$ when A is non-singular.

Objectives:

The reader should be able to

(i) decompose a square matrix into a product of two triangular matrices; and

(ii) apply the decomposition to the computation of inverses of non-singular matrices and the solution of systems of linear equations with non singular matrix of coefficients.

Pre-Test

1. Decompose the matrix over R

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

into triangular matrices as
(i) $A = L_1U_1$ and (ii) $A = U_2L_2$
where L_1 and L_2 are lower triangular matrices and U_1 and U_2 are upper triangular matrices.

2. If A and B are triangular matrices such that

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 1 & -2 & 0 \\ -2 & 4 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}.$$

Find A^{-1} and B^{-1}. Hence, find the matrix $C = AB$ and its inverse C^{-1}.

3. Decompose over R

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

into triangular matrices in two ways. Hence, find A^{-1} and solve the system of linear equations

$$\begin{aligned} x - y + z &= 1 \\ x + y - z &= 2 \\ x + y + z &= 5 \end{aligned}$$

Definitions
There are two types of triangular matrices.
These are

(a) lower triangular matrices which are square matrices having zeroes in all entries above the principal diagonal, and

(b) upper triangular matrices which are square matrices, having zeroes in all entries below the principal diagonal.

Note that a diagonal matrix is both a lower and an upper triangular matrix. For example,

$$\begin{pmatrix} a & 0 & 0 \\ d & b & 0 \\ e & f & c \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a & d & e \\ 0 & b & f \\ 0 & 0 & c \end{pmatrix}$$

are lower and upper triangular matrices, respectively.

We have the following important facts about the decomposition of any square matrix.

Every square matrix A can be decomposed as $A = L_1U_1$, where L_1 is a lower triangular matrix and U_1 is an upper triangular matrix or as $A = U_2L_2$ where U_2 is an upper triangular matrix and L_2 is a lower triangular matrix.

The above decompositions of a square matrix A are useful in the calculation of the inverse matrix A^{-1} if A is non-singular. They are also useful in the solution of systems of linear equations

$$A\underline{x} = \underline{y}$$

when A is non-singular, since if $A = LU$, then

$$LU\underline{x} = \underline{y} \Leftrightarrow U\underline{x} = L^{-1}\underline{y}$$

and the last equation is a reduced system of linear equations which can be easily solved.

If a triangular matrix is of the form

$$L = \begin{pmatrix} a & 0 & 0 \\ d & b & 0 \\ e & f & c \end{pmatrix} \text{ or } U = \begin{pmatrix} a & d & e \\ 0 & b & f \\ 0 & 0 & c \end{pmatrix}$$

where $a \neq 0$, $b \neq 0$ and $c \neq 0$, then L and U are non-singular matrices and their inverses are

$$L^{-1} = \begin{pmatrix} \frac{1}{a} & 0 & 0 \\ \frac{d}{ab} & \frac{1}{b} & 0 \\ \frac{df - be}{abc} & \frac{-f}{bc} & \frac{1}{c} \end{pmatrix} \text{ and } U^{-1} = \begin{pmatrix} \frac{1}{a} & \frac{d}{ab} & \frac{df - be}{abc} \\ 0 & \frac{1}{b} & \frac{-f}{bc} \\ 0 & 0 & \frac{1}{c} \end{pmatrix}$$

Example
Decompose the matrix over R,

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

into triangular matrices as
(i) $A = L_1U_1$ and (ii) $A = U_2L_2$
where L_1 and L_2 are lower triangular matrices and U_1 and U_2 are upper triangular matrices. Hence, obtain the inverse of A, if it exists. Furthermore, use either of your decompositions to solve the system of linear equations

$$\begin{aligned} 2x - y + 3z &= 7 \\ x + 2y - z &= 1 \\ x + y + z &= 3 \end{aligned}$$

(i) Assume that the decomposition is of the form

$$L_1U_1 = \begin{pmatrix} a & 0 & 0 \\ d & b & 0 \\ e & f & c \end{pmatrix} \begin{pmatrix} 2 & -1 & 3 \\ 1 & b^1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 2 & -1 \\ 1 & 1 & 1 \end{pmatrix} = A$$

Then, by expanding L_1U_1 using the first column of U_1, we obtain that

$$a = 1, \quad d = \frac{1}{2}, \quad e = \frac{1}{2}$$

By using the third column of U_1, we obtain that $b = \frac{5}{2}$. Next use the second column of U_1 to obtain that $b = 1$, $f = \frac{3}{2}$. Finally, use the third column of U_1 to obtain that $c = 1$. Hence the required decomposition is

$$\begin{pmatrix} 2 & -1 & 3 \\ 1 & 2 & -1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{5}{2} & 0 \\ \frac{1}{2} & \frac{3}{2} & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

(ii) Assume that the decomposition is of the form

$$U_2L_2 = \begin{pmatrix} a & d & e \\ 0 & b & f \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 1 & b' & 0 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 2 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

Then $e = 3$, $f = -1$, $c = 1$, $b = 2$, $b = \frac{3}{2}$, $d = -\frac{8}{3}$, $a = \frac{5}{6}$.

Hence the required decomposition is

$$\begin{pmatrix} 2 & -1 & 3 \\ 1 & 2 & -1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{5}{6} & -\frac{8}{3} & 3 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 1 & \frac{3}{2} & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

By (i) above L_1 and U_1 are non-singular and so $A = L_1 U_1$ is also non-singular and

$$A^{-1} = U_1^{-1} L_1^{-1} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{5} & \frac{2}{5} & 0 \\ -\frac{1}{5} & -\frac{3}{5} & 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} & -1 \\ -\frac{2}{5} & -\frac{1}{5} & 1 \\ -\frac{1}{5} & -\frac{3}{5} & 1 \end{pmatrix}$$

The given system of linear equations can be written as $A\underline{x} = \underline{y}$
$\Leftrightarrow$ (by (ii) above) $U_2 L_2 \underline{x} = \underline{y} \Leftrightarrow L_2 \underline{x} = U_2^{-1} \underline{y}$

$$\Leftrightarrow \begin{pmatrix} 2 & 0 & 0 \\ 1 & \frac{3}{2} & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{6}{5} & \frac{8}{5} & -2 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 7 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix}$$

$$\begin{aligned} \Leftrightarrow \quad & 2x = 4 \text{ i.e. } x = 2 \\ & x + \frac{3}{2}y = 2, \text{ i.e. } y = \frac{2}{3}(2-2) = 0 \\ & x + y + z = 3 \text{ i.e. } z = 3 - 2 = 1 \end{aligned}$$

The solutions are

$$x = 2, \ y = 0, \ z = 1.$$

Practice Exercise

1. If A and B are triangular matrices such that

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 & 0 \\ 1 & -3 & 0 \\ 2 & -10 & 2 \end{pmatrix}$$

   Find A^{-1} and B^{-1}. Hence, find the matrix $C = AB$ and its inverse C^{-1}.

2. Decompose over R

$$B = \begin{pmatrix} 1 & 1 & -2 \\ 1 & -1 & 2 \\ 2 & 1 & 1 \end{pmatrix}$$

   into triangular matrices in two ways. Hence, find B^{-1} and solve the system of linear equations

$$\begin{aligned} x + y - 2z &= 7 \\ x - y + 2z &= 3 \\ x + y + z &= 5 \end{aligned}$$

Summary
Two methods are given for decomposing a square matrix into a product of 2 triangular matrices. The decomposition is then used to find the inverse of a non-singular matrix and to solve a system of linear equations when the matrix of coefficients is non-singular.

Post-Test
See Pre-Test at the beginning of the Unit.

References

1. Ayres, F. *Modern Algebra*, Schaum's Outline Series.

2. Ayres, F. *Matrices*, Schaum's Outline Series.

3. Birkoff, G. and S. MacLane, *A survey of Modern Algebra*, Macmillan Co. 1965.

4. Ilori, S.A. and O. Akinyele, *Elementary Abstract* and *Linear Algebra*, Ibadan University Press, 1986, pp. 231-263. Reprinted 2006.