SURESH
GYAN VIHAR
UNIVERSITY
Accredited by NAAC with 'A+' Grade

Master of Science Mathematics

(M.Sc. Mathematics)

MMT-203

Linear Algebra

Semester-II

Author- Dr. Harsh Vardhan Harsh

SURESH GYAN VIHAR UNIVERSITY
Centre for Distance and Online Education
Mahal, Jagatpura, Jaipur-302025

Published by:
S. B. Prakashan Pvt. Ltd.
WZ-6, Lajwanti Garden, New Delhi: 110046
Tel.: (011) 28520627 | Ph.: 9205476295
Email: info@sbprakashan.com | Web.: www.sbprakashan.com

Every effort has been made to avoid errors or omissions in the publication. In spite of this, some errors might have crept in. Any mistake, error or discrepancy noted may be brought to our notice and it shall be taken care of in the next edition. It is notified that neither the publishers nor the author or seller will be responsible for any damage or loss of any kind, in any manner, therefrom.

For binding mistakes, misprints or for missing pages, etc., the publishers' liability is limited to replacement within one month of purchase by similar edition. All expenses in this connection are to be borne by the purchaser.

Designed & Graphic by : S. B. Prakashan Pvt. Ltd.

Printed at :

Suresh Gyanvihar University

Department of Mathematics

School of Science

M.Sc., Mathematics - Syllabus – I year – II Semester (ODL Mode)

COURSE TITLE	**:**	LINEARALGEBRA
COURSE CODE	**:**	**MMT- 203**
COURSE CREDIT	**:**	**4**

COURSE OBJECTIVES

While studying the **LINEARALGEBRA**, the Learner shall be able to:

CO 1: Discuss the concept of null space and range of a linear transformation.

CO 2: Review the concept of a algebra over a field.

CO 3: Represent linear transformation on a vector space by matrices.

CO 4: Describe the concept of direct sum and interior direct sum.

CO 5:Review the concept of companion matrix.

COURSE LEARNING OUTCOMES

After completion of the **LINEARALGEBRA**, the Learner will be able to:

CLO 1: Interpret the idea of linear transformation, identify them to represent the linear transformation by matrices.

CLO 2: Describe the prime factorization of a polynomial and write each polynomial as the product of prime polynomials.

CLO 3: Enable to find the characteristic value and characteristic vectors of a linear transformation.

CLO 4: Interpret the idea of linear transformation; identify them to represent the ordered basis by triangular matrix.

CLO 5: Interpret the ideas of Jordan forms and rational forms of real matrices.

BLOCK I: LINEAR TRANSFORMATIONS

Linear transformations – Isomorphism of vector spaces – Representations of linear transformations by matrices – Linear functionals.

BLOCK II: ALGEBRA OF POLYNOMIALS

The algebra of polynomials –Polynomial ideals - The prime factorization of a polynomial - Determinant functions.

BLOCK III: DETERMINANTS

Permutations and the uniqueness of determinants – Classical adjoint of a (square) matrix – Inverse of an invertible matrix using determinants – Characteristic values – Annihilating polynomials.

BLOCK IV: DIAGONALIZATION

Invariant subspaces – Simultaneous triangulations – Simultaneous diagonalization – Direct-sum decompositions – Invariant direct sums – Primary decomposition theorem.

BLOCK V: THE RATIONAL AND JORDAN FORMS

Cyclic subspaces – Cyclic decompositions theorem (Statement only) – Generalized Cayley – Hamilton theorem - Rational forms – Jordan forms.

REFERENCE BOOKS :

1. Kenneth M Hoffman and Ray Kunze, Linear Algebra, 2nd Edition, Prentice-Hall of India Pvt. Ltd, New Delhi, 2013.

UNIT	Chapter(s)	Sections
I	3	3.1 – 3.5
II	4 & 5	4.1, 4.2, 4.4, 4.5 and 5.1, 5.2
III	5 & 6	5.3, 5.4 and 6.1 – 6.3
IV	6	6.4 – 6.8
V	7	7.1 – 7.3

2. M. Artin, "*Algebra*", Prentice Hall of India Pvt. Ltd., 2005.
3. S.H. Friedberg, A.J. Insel and L.E Spence, "*Linear Algebra*", 4th Edition, Pritice-Hall of India Pvt. Ltd., 2009.
4. I.N. Herstein, "*Topics in Algebra*", 2nd Edition, Wiley Eastern Ltd, New Delhi, 2013.
5. J.J. Rotman, "*Advanced Modern Algebra*", 2nd Edition, Graduate Studies in Mathematics, Vol. 114, AMS, Providence, Rhode Island, 2010.
6. G. Strang, "*Introduction to Linear Algebra*", 2ndEdition, Prentice Hall of India Pvt. Ltd,2013.

Web resources

https://www.youtube.com/watch?v=Ts3o2I8_Mxc

https://www.youtube.com/watch?v=Yaijk7zegFg

https://www.youtube.com/watch?v=9pqhfDyzbhw

https://www.youtube.com/watch?v=ZvL9aDGNHqA

https://www.youtube.com/watch?v=NHTI0SOpEpU https://www.youtube.com/watch?v=3ROzG6n4yMc

https://www.youtube.com/watch?v=7gWP96bL9jw

https://www.youtube.com/watch?v=Fg7_mv3izR0

https://www.youtube.com/watch?v=XF55ilf9ZpQ

https://www.youtube.com/watch?v=aRewVVUzJ2c

https://www.youtube.com/watch?v=r9smdgQcpC8

https://www.youtube.com/watch?v=sJV0QyHoRio

https://www.youtube.com/watch?v=GR4TTzq12Uk

https://www.youtube.com/watch?v=MWYifkq9hWs

https://www.youtube.com/watch?v=btLbluS_Qp4

Contents

BLOCK - I

Unit - 1: Linear Transformations-I.

Unit - 2: Linear Transformations-II.

Block-I

UNIT-1

LINEAR TRANSFORMATIONS-I

Overview

In this unit, we will illustrate the basic concepts of linear transformations.

Objectives

After successful completion of this lesson, students will be able to

understand the concept of linear transformation.

explain the concept of null space and range of linear transformation.

explain the concept of linear operator.

1.1. Linear Transformations

In this section, we shall study the concept of linear transformations.

Definition 1.1 (Linear Transformation). Let V and W be vector spaces over the field F. A linear transformation from V into W is a function T from V into W such that

$$T(c \quad + \quad) = c(T \quad) + T(\quad)$$

for all and in V and all scalars c in F.

Example 1.1. If V is any vector space, the identity transformation I; defined by $I \quad = \quad$; is a linear transformation from V into V.

Example 1.2. If V is any vector space, the zero transformation O; defined by $O \quad = O$; is a linear transformation from V into V.

Example 1.3. Let F be a field and let V be the space of polynomial functions f from F into F; given by

$$f(x) = c_0 + c_1x + \quad + c_kx^k \tag{1.1}$$

$$\text{Let} \quad (Df)x = c_1 + 2c_2x + \quad + kc_kx^{k \quad 1} \tag{1.2}$$

Then D is a linear transformation from V into V and it is also called the differentiation transformation.

If $f(x)$ is a polynomial over the field F, then $Df(x)$ is also a polynomial over the field F.

Thus, if $f(x)$ 2 V, then $Df(x)$ 2 V. Therefore D is a function from V into V.

Also, if $f(x); g(x) \in V$ and $a; b \in F$ then

$$D\left(af(x) + g(x)\right) = aDf(x) + Dg(x) \tag{1.3}$$

$\Rightarrow$ D is a linear transformation from V into V.

Example 1.4. Let A be a fixed $m \times n$ with entires in the field F. Define a function $T : F^{n\times 1} \to F^{m\times 1}$ by $T(X) = AX$:

Then T is a linear transformation from $F^{n\times 1}$ into $F^{m\times 1}$.

The function U defined by $U(\alpha) = \alpha A$ is a linear transformation from F^m into F^n.

Example 1.5. Let P be a fixed $m \times m$ matrix with entries in the field F and let Q be a fixed $n \times n$ matrix over F. Define a function $T : F^{m\times n} \to F^{m\times n}$ by

$$T(A) = PAQ \tag{1.4}$$

Then T is a linear transformation.

Proof.

$$\begin{aligned} T(cA + B) &= P(cA + B)Q \\ &= (cPA + PB)Q \\ &= cPAQ + PBQ \\ &= cT(A) + T(B) \end{aligned}$$

Thus, T is a linear transformation from $F^{m\times n}$ into $F^{m\times n}$.

Example 1.6. Let R be the field of real numbers and let V be the space of all functions from R into R which are continuous. Define $T : V \to V$ by

$$(Tf)x = \int_0^x f(t)dt$$

Then,

$$\begin{aligned} T(af + g)(x) &= \int_0^x \left(af + g\right)(t)dx \\ &= \int_0^x af(t)dx + \int_0^x g(t)dt \\ &= a\int_0^x f(t)dt + \int_0^x g(t)dt \\ &= aT(f)x + T(g)x \end{aligned}$$

$\Rightarrow$ T is a linear transformation from V into V. It is also called integral transformation.

Note 1.1. If T is a linear transform V into W, then $T(0) = 0$, because

$$\begin{aligned} T(0) &= T(0+0) \\ &= T(0) + T(0) \\ \Rightarrow \quad T(0) &= 0 \end{aligned}$$

Note 1.2. If T is a linear transformation from V into W. If $\,_1; \ \,_2; \ + \ \,_n$ are vectors in V and $c_1; c_2; \ ; c_n$ are scalars, then

$$T(c_1\,_1 + c_2\,_2 + \quad + c_n\,_n) = c_1T(\,_1) + c_2T(\,_2) + \quad + c_nT(\,_n)$$

Theorem 1.1. Let V be a finite-dimensional vector space over the field F and let $\{\,_1; \ \,_2; \quad ; \ \,_n\}$ be an ordered basis for V. Let W be a vector space over the same field F and let $\{\,_1; \ \,_2; \quad ; \ \,_n\}$ be any vectors in W. Then there is a precisely one linear transformation T from V into W such that

$$T\,_j = \,_j; \qquad j = 1; 2; \ldots; n \tag{1.5}$$

Proof. First we shall prove that there is a linear transformation T with

$$T(\,_j) = \,_j$$

Let $\ \in V$, then there is a unique n-tuples $(x_1; x_2; \ldots; x_n)$ such that

$$= x_1\,_1 + x_2\,_2 + \ldots + x_n\,_n$$

For this vector $\ $, we define

$$T(\ ) = x_1\,_1 + x_2\,_2 + \ldots + x_n\,_n$$

Obviously $T(\ )$ as defined above is a unique element of V. Therefore T is well defined rule for associating with each vector $\ $ in V a unique vector $T(\ )$ in W. Thus, T is a function V into W.

The unique representation of $\,_i \in V$ as a linear combination of the vectors is

$$_i = 0\,_1 + 0\,_2 + \ldots + 1\,_i + \ldots + 0\,_n$$

Therefore, according to definition of T, we have

$$\begin{aligned} T(\,_i) &= 0\,_1 + 0\,_2 + \ldots + 1\,_i + \ldots + 0\,_n \\ i.e.; \quad T(\,_i) &= \,_i; \quad i = 1; 2; \ldots; n \end{aligned}$$

Now, our aim is to prove that T is a linear transformation.

Let $= y_1 \ _1 + y_2 \ _2 + ::: + y_n \ _n$ 2 V and c be any scalar.

Then

$$\begin{aligned} T(c \ + \ ) &= (cx_1 + y_1) \ _1 + (cx_2 + y_2) \ _2 + ::: + (cx_n + y_n) \ _n \\ &= c(x_1 \ _1 + x_2 \ _2 + ::: + x_n \ _n) + y_1 \ _1 + y_2 \ _2 + ::: + y_n \ _n \\ &= cT(\ ) + T(\ ) \end{aligned}$$

) T is a linear transformation from V into W such that. Thus, there exists a linear transformation from V into W such that

$$T(\ _i) = \ _i; \qquad i = 1; 2; ::: ; n$$

It remains to prove that the uniqueness of T .

Let U be a linear transformation from V into W such that

$$U(\ _i) = \ _i; \qquad i = 1; 2; ::: ; n$$

For the vector $= x_1 \ _1 + x_2 \ _2 + ::: + x_n \ _n$ 2 V , we have

$$\begin{aligned} U(\ ) &= U(x_1 \ _1 + x_2 \ _2 + ::: + x_n \ _n) \\ &= x_1U(\ _1) + x_2U(\ _2) + ::: + x_nU(\ _n) \\ &= x_1 \ _1 + x_2 \ _2 + ::: + x_n \ _n \\ &= T(\ ) \end{aligned}$$

Thus, T is a unique linear transformation from V into W such that

$$T(\ _i) = \ _i \qquad i = 1; 2; ::: ; n$$

Example 1.7. Find the linear transformation $T : R^2$! R^2 such that $T(2; 3) = (4; 5)$ and $T(1; 0) = (0; 0)$:

Solution. First we shall show that the set f(2; 3); (1; 0)g is a basis of R^2 .

First we shall prove that the linear independence of this set.

Let

$$\begin{aligned} a(2; 3) + b(1; 0) &= (0; 0) \qquad \text{where } a; b \ 2 \ R \\ ) \quad (2a + b; 3a) &= (0; 0) \\ ) \quad 2a + b &= 0; \quad 3a = 0 \\ ) \quad a = 0; \quad b = 0 & \end{aligned}$$

Hence, the set f(2; 3); (1; 0)g is a linearly independent.

Next, we shall prove that the set $\{(2;3);(1;0)\}$ spans R^2.

Let $(x_1;x_2) \in R^2$ and let

$$(x_1;x_2) = a(2;3)+b(1;0) = (2a+b;3a)$$

Then $2a+b = x_1; \quad 3a = x_2$

$$\Rightarrow \quad a = \frac{x_2}{3}; \quad b = \frac{3x_1-2x_2}{3};$$

Thus, we have $(x_1;x_2) = \frac{x_2}{3}(2;3) + \frac{3x_1-2x_2}{3}(1;0)$:

From the above relation, we see that the set $\{(2;3);(1;0)\}$ spans R^2. Hence this is a basis for R^2:

Now, let $(x_1;x_2)$ be any member of R^2, then we can find a formula for $T(x_1;x_2)$ with the conditions that $T(2;3) = (4;5)$ and $T(1;0) = (0;0)$:

We have

$$\begin{aligned} T(x_1;x_2) &= T\left[\frac{x_2}{3}(2;3) + \frac{3x_1-2x_2}{3}(1;0)\right] \\ &= \frac{x_2}{3}T(2;3) + \frac{3x_1-2x_2}{3}T(1;0) \\ &= \frac{x_2}{3}(4;5) + \frac{3x_1-2x_2}{3}(0;0) \\ &= \left(\frac{4x_2}{3};\frac{5x_2}{3}\right) \end{aligned}$$

If T is a linear transformation from V into W, then the range of T is not only a subset of W and also it is a subspace of W. Let R_T be the range of T then

$$R_T = \{\beta \in W : T(\alpha) = \beta \text{ for some vector } \alpha \text{ in } W\}$$

Our wish is to prove that R_T is a subspace of W.

For this, let $\beta_1;\beta_2 \in R_T$ and let c be a scalar. Then there exists a vectors α_1 and α_2 in V such that

$$\begin{aligned} T(\alpha_1) &= \beta_1 \\ T(\alpha_2) &= \beta_2 \end{aligned}$$

Consider

$$\begin{aligned} T(c\alpha_1+\alpha_2) &= cT(\alpha_1)+T(\alpha_2) \\ &= cT(\alpha_1)+T(\alpha_2) \\ &= c\beta_1+\beta_2 \in R_T \end{aligned}$$

Thus, R_T is a subspace of W.

If T is a linear transformation from V into W.

Let $N = \{ \ \in V/T(\ ) = 0\}$:

Clearly, N is non -empty, since $T(0) = 0$ and $0 \in N$.

Now, our claim is to prove that N is a subspace of V.

Let $_1;\ _2 \in V$ and c be a scalar, then

$$T(\ _1) = 0$$

$$T(\ _2) = 0$$

Consider

$$\begin{aligned} T(c\ _1 + \ _2) &= cT(\ _1) + T(\ _2) \\ &= c(0) + 0 \\ &= 0 \end{aligned}$$

$$\Rightarrow \quad c\ _1 + \ _2 \in N$$

Thus, N is a subspace of V.

Definition 1.2. Let V and W be vector spaces over the field F and let T be a linear transformation from V into W. The null space of T is the set of all vectors in V such that $T(\ ) = 0$.

If V is finite-dimensional, the rank of T is the dimension of the range of T and the nullity of T is the dimension of the null space of T.

Theorem 1.2. Let V and W be vector spaces over the field F and let T be a linear transformation from V into W. Suppose that V is finite-dimensional. Then

$$\text{rank}(T) + \text{nullity}(T) = \dim V$$

Proof. Let V and W be a vector space over the field F and given that V is finite-dimensional.

Let us assume that $\dim V = n$:

We know that N is the null space of T, is a subspace of V.

$\Rightarrow \dim N \ \ n$: Hence , we assume that $\dim N = k\ (\ \ n)$.

It remains to prove that $\dim R(T) = n\ \ k$:

Let $\{\ _1;\ _2; : : : ;\ _k$ be a basis for N. Then the set $\{\ _i\}$ can be extended to become a basis of V . So, we assume that f $_1$; $_2$; : : : ; $_k$; $_{k+1}$; : : : ; $_n$g is a basis of V .

Claim: The set fT($_{k+1}$); T($_{k+2}$); : : : ; T($_n$)g for the range of T:

i:e:; We can prove that the set fT($_{k+1}$); T($_{k+2}$); ; T($_n$)g are linearly independent and that they span the range of T .

Let 2 R(T) (Range of T).

Thus, by definition of range of T , there exists a vector 2 V such that

$$T( ) = \qquad (1.6)$$

We know that the set f $_1$; $_2$; : : : ; $_k$; $_{k+1}$; : : : ; $_n$g forms a basis of V . Also 2 V .

Hence, every element of V can be expressed as a linear combination of the set f $_1$; $_2$; : : : ; $_k$; $_{k+1}$; : : : ; $_n$g.

Therefore, $= c_1 \ _1 + c_2 \ _2 + : : : + c_n \ _n$:

From equation (1.6), we have

$$= T (c_1 \ _1 + c_2 \ _2 + : : : + c_n \ _n)$$
$$= c_1T(\ _1) + c_2T(\ _2) + : : : + c_nT(\ _n)$$
$$= \text{ a linear combination of } T(\ _1); T(\ _2); : : : ; T(\ _n)$$

i:e:; Every element of range of T is a linear combination of T($_1$); T($_2$); : : : ; T($_n$) .

Similarly, we can show that every linear combination of

T($_1$); T($_2$); : : : ; T($_n$) is an element of range of T .

Thus, the set T($_1$); T($_2$); : : : ; T($_n$) spans R(T) .

But, the set T($_1$); T($_2$); : : : ; T($_k$) is a basis for the null space N .

$$) \quad \ _1; \ _2; : : : ; \ _k{}_j \ 2 \ N$$
$$) \quad T(\ _j) = 0; \quad j = 1; 2; : : : ; k$$

Therefore, the set T($_{k+1}$); T($_{k+2}$); : : : ; T($_n$) spans R(T) .

Claim: The set T($_{k+1}$); T($_{k+2}$); : : : ; T($_n$) are linearly independent.

Assume that, there exist a scalars $c_{k+1}; c_{k+2}; : : : ; c_n$ such that

$$c_{k+1}T(\ _{k+1}) + c_{k+2}T(\ _{k+2}) + : : : + c_nT(\ _n) = 0 \qquad (1.7)$$

Now, our wish is to prove that $c_{k+1} = c_{k+2} = ::: = c_n = 0$:

From equation (1.7), we have

$$\sum_{i=k+1}^{n} c_i T(\ _i) = 0$$

$$\Rightarrow T\left\{\sum_{i=k+1}^{n} c_i\ _i\right\} = 0$$

$$\Rightarrow \sum_{i=k+1}^{n} c_i\ _i \in N$$

$$\Rightarrow \ \in N \qquad \text{where} \quad = \sum_{i=k+1}^{n} c_i\ _i \tag{1.8}$$

Hence there exists a scalars $b_1; b_2; ::: ; b_k$ such that

$$= \sum_{i=1}^{k} b_i\ _i$$

$$\sum_{i=k+1}^{n} c_i\ _i = \sum_{i=1}^{k} b_i\ _i$$

$$\Rightarrow \sum_{i=k+1}^{n} c_i\ _i \quad \sum_{i=1}^{k} b_i\ _i = 0$$

$$\Rightarrow (b_1\ _1 + b_2\ _2 + ::: + b_k\ _k) \quad (c_{k+1}\ _{k+1} + c_{k+2}\ _{k+2} + ::: + c_n\ _n) = 0$$

Since, the set $\{\ _1; \ _2; ::: ; \ _n\}$ are linearly independent, thus we have

$$b_1 = b_2 ::: = b_k = \quad c_{k+1} = \quad c_{k+2} = ::: = \quad c_n = 0$$

$$\Rightarrow b_1 = b_2 ::: = b_k = c_{k+1} = c_{k+2} = ::: = c_n = 0$$

Thus, the set $\{T(\ _1; \ _2; ::: ; \ _n)\}$ are linearly independent.

Hence, the set $\{T(\ _1; \ _2; ::: ; \ _n)\}$ is a basis of range of T.

i:e:; $\dim R(T) = n\ k$

Let $r = \dim R(T)$

$r = n\ k$

$\Rightarrow n = r + k$:

Hence $\dim V =$ rank of T + nullity of T:

This completes the proof of the theorem.

Theorem 1.3. Uf A is an $m \ n$ matrix with entries in the field F, then

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

Proof. Let T be the linear transformation from $F^{n\ 1}$ into $F^{m\ 1}$ defined by $T(X) = AX$:

Suppose $AX = 0$

i:e:; X is the solution space of the system $AX = 0$.

i:e:; The set of all column matrices X such that $AX = 0$:

(The null space of $T = \{\ \ \in V/T(\ \ ) = 0\}$) .

$\Rightarrow$ $T(X) = AX$ and $AX = 0$ which implies that $T(X) = 0$.

Thus, the null space of T is the solution space for the system $AX = 0$.

Suppose $AX = Y$:

i:e:; The set of all column matrices X such that $AX = Y$:

Thus, the range of T is the set of all $m\ 1$ column matrices Y such that $AX = Y$ has a solution $X = A^{\ 1}Y$:

Let $A_1; A_2; : : : ; A_n$ be the columns of the matrix A

Let

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

Then

$$\begin{aligned} AX &= A_1X_1 + A_2X_2 + : : : + A_nX_n \\ Y &= A_1X_1 + A_2X_2 + : : : + A_nX_n \end{aligned}$$

Y is an arbitrary elements of range of T and Y is spanned by $A_1; A_2; : : : ; A_n$.

i:e:; Range of T is the subspace spanned by the columns of A .

In otherwords, the range of T is the column space of A .

Therefore rank (T) = column rank(A) .

But, if S is the solution space for the system $AX = 0$, then

$$\dim S + \text{column rank}(A) = n$$

If r is the dimension of the row space of A , then the solution space S has a basis consisting of $n\ r$ vectors.

$$\dim S = n\ \ \text{row rank}(A)$$

$)$; we have row rank(a) = column rank(A):

This completes the proof of the theorem.

1.1.1. Examples

Example 1.8. Find a linear map $F : R^3 \; ! \; R^4$ whose image is generated by $(1;\ 1; 2; 3)$ and $(2; 3;\ 1; 0)$

Solution. Consider the usual basis of R as given below:

$f e_1 = (1; 0; 0); e_2 = (0; 1; 0); e_3 = (0; 0; 1)g$

Write $F(e_1) = (1;\quad 1; 2; 3))$, $F(e_2) = (2; 3;\quad 1; 0))$ and

$F(e_3) = (0; 0; 0; 0)$

Clearly,

$$\begin{aligned}
(x; y; z) &= xe_1 + ye_2 + ze_3 \\
F(x; y; z) &= F(xe_1 + ye_2 + ze_3) \\
&= xF(e_1) + yF(e_2) + zF(e_3) \\
&= (x;\quad x; 2x; 3x) + (2y; 3y;\quad y; 0) + (0; 0; 0; 0) \\
&= (x + 2y;\quad x + 3y; 2x\quad y; 3x)
\end{aligned}$$

Example 1.9. Show that the mappint $T : V_2(R) \; ! \; V_3(R)$ defined as

$$T(a; b) = (a + b; a\quad b; b)$$

is a linear transformation from $V_2(R)$ into $V_3(R)$. Find the range, rank, null-space and nullity of T.

Solution. Let $\quad = (a_1; b_1); \quad = (a_2; b_2) \; 2 \; V_3(R)$:

Then $T(\ ) = T(a_1; b_1) = (a_1 + b_1;\ a_1\ b_1;\ b_1)$

and $T(\ ) = T(a_2; b_2) = (a_2 + b_2;\ a_2\ b_2;\ b_2)$.

Also let $a; b \; 2 \; R$. Then $a\quad + b \; 2 \; V_3(R)$ and

$$\begin{aligned}
T(a\quad + b\ ) &= T\ [a(a_1; b_1) + b(a_2; b_2)] \\
&= T\ (aa_1 + ba_2; ab_1 + bb_2) \\
&= (aa_1 + ba_2 + ab_1 + bb_2; aa_1 + ba_2\ \ ab_1\quad bb_2; ab_1 + b_2) \\
aT(\ ) + bT(\ ) &= a(a_1 + b_1; a_1\ \ b_1; b_1) + b(a_2 + b_2; a_2\ \ b_2; b_2)
\end{aligned}$$

Therefore, T is a linear transformation from $V_3(R)$ into $V_3(R)$.

Now $\{(1; 0); (0; 1)\}$ is a basis for $V_3(R)$.

Thus, we have

$$\begin{aligned} T(1; 0) &= (1 + 0; 1 - 0; 0) = (1; 1; 0) \\ T(0; 1) &= (0 + 1; 0 - 1; 0) = (1; -1; 0) \end{aligned}$$

The vectors $T(0; 1)$ and $T(1; 0)$ span the range of T.

Thus, the range of T is the subspace of $V_3(R)$ spanned by the vectors $(1; 1; 0); (1; -1; 1)$.

Now the vectors $(1; 1; 0); (1; -1; 1) \in V_3(R)$ are linearly independent because if $x; y \in R$, then

$$\begin{aligned} x(1; 1; 0) + y(1; -1; 0) &= (0; 0; 0) \\ \Rightarrow (x + y; x - y; 0) &= (0; 0; 0) \\ \Rightarrow x + y = 0; \ x - y = 0 \ \Rightarrow & \ x = 0; \ y = 0 \end{aligned}$$

$\Rightarrow$ The vectors $(1; 1; 0); (1; -1; 0)$ form a basis for range of T.

Hence rank T = dim of range of $T = 2$:

Nullity of T = dim of $V_2(R)$ -rank $T = 2 - 2 = 0$.

$\Rightarrow$ null space of T must be the zero subspace of $V_2(R)$.

1.2. The Algebra of Linear Transformations

In the study of linear transformations from V into W, it is of fundamental importance that the set of these transformations inherits a natural vector space structure. The set of linear transformations froma space V into itself has even more algebraic structure, because ordinary composition of functions provides a multiplication of such transformations. Now, we shall see these ideas in this section.

Theorem 1.4. Let V and W be vector spaces over the field F. Let T and U be linear transformations from V into W. The function $(T + U)$ defined by

$$(T + U)(\alpha) = T(\alpha) + U(\alpha)$$

is a linear transformation from V into W. If c is any element of F, the function

CT is defined by

$$(cT)(\ ) = cT(\ )$$

is a linear transformation from V into W. The set of all linear transformations from V into W, together with the addition and scalar multiplication defined above, is a vector space over the field F.

Proof. Given that $T : V \ ! \ W$ and $U : V \ ! \ W$ are linear transformations such that

$$(T + U)(\ + \ ) = T(\ ) + U(\ )$$

Consider

$$\begin{aligned} (T + U)(c \ + \ ) &= T(c \ + \ ) + U(c \ + \ ) \\ &= c(T(\ ) + T(\ ) + c(U(\ ) + U(\ ) \\ &= c\,[T(\ ) + U(\ )] + (T(\ ) + U(\ ) \\ &= c(T + U)(\ ) + (T + U)(\ ) \end{aligned}$$

) $T + U$ is a linear transformation.

Similarly

$$\begin{aligned} (cT)(d \ + \ ) &= c\ T(d \ + \ ) \\ &= c\ dT(\ ) + T(\ ) \\ &= d\,[cT(\ )] + c(T(\ )) \\ &= d\,[(cT)(\ )] + (CT) \end{aligned}$$

) cT is a linear transformation.

Next our wish to prove that the set of all linear transformation from V into W is a vector space over F with respect to the vector addition and scalar multiplication.

$$(T + U)(\ ) = T(\ ) + U(\ ); \quad 8 \ \ 2\, V \tag{1.9}$$

$$(cT)(\ ) = c(T(\ )); \quad 8\ c\ 2\ F; \ \ 2\, V \tag{1.10}$$

Addition is Cummutative:

$$\begin{aligned} \text{Consider } (T + U)(\ ) &= T(\ ) + U(\ ) \\ &= U(\ ) + T(\ ) \\ &= (U + T)(\ ) \end{aligned}$$

Addition is Associative: Let $S : V \to W$ be any linear transformation.

$$\begin{aligned} \text{Consider } (T + (U + S))(\ ) &= T(\ ) + (U + S)(\ ) \\ &= T(\ ) + U(\ ) + S(\ ) \\ &= (T + U)(\ ) + S(\ ) \\ &= ((T + U) + S)(\ ) \end{aligned}$$

Identity transformation under addition:

Define Zero transformation $0 : V \to W$ by $0(\ ) = 0$.

For this unique linear transformation, $0 : V \to W$, we have $T + 0 = T$, for all T.

Inverse transformation under addition:

For each linear transformation $T : V \to W$; there exists a unique linear transformation T such that $T + (T) = 0$ where T is the inverse linear transformation.

Identity transformation under multiplication:

$1(T) = T \ \forall T : V \to W$ is a linear transformation.

Cummutative under addition: Let $c_1; c_2 \in F$ and $T : V \to W$ be a linear transformation.

$$\begin{aligned} \text{Consider } [(c_1c_2)T](\ ) &= (c_1c_2)(T(\ )) \\ &= c_1 [c_2T(\ )] \\ &= [c_1(c_2T)](\ ) \\ \Rightarrow \quad (c_1c_2)T &= c_1(c_2T) \end{aligned}$$

Distribution Law:

(i) Let $c \in F$ and let $T : V \to W$ and $U : V \to W$ be linear transformations.

$$\begin{aligned} \text{Consider } [c(T + U)(\ )] &= c[(T + U)(\ )] \\ &= c[(T + U)(\ )] \\ &= c[T(\ ) + U(\ )] \\ &= c[T(\ )] + c[U(\ )] \\ &= (cT + cU)(\ ) \\ \text{Thus, } c(T + U) &= cT + cU \end{aligned}$$

(ii) Let $c_1; c_2$ 2 F and $T : V$! W be a linear transformation.

$$
\begin{aligned}
\text{Consider}\quad [(c_1 + c_2)T](\ ) &= (c_1 + c_2)T(\ )\\
&= c_1T(\ ) + c_2T(\ )\\
&= (c_1T)(\ ) + (c_2T)(\ )\\
&= (c_1T + c_2T)(\ )\\
)\quad (c_1 + c_2)T &= c_1T + c_2T
\end{aligned}
$$

Thus, the set of all linear transformations from V into W is a vector space.

This completes the proof of the theorem.

Definition 1.3. Let $V; W$ are vector spaces over the same field F. Then the set of all linear transformations from V into W is denoted by $L(V; W)$:

Note 1.3. $L(V; W)$ is a vector space over F.

Theorem 1.5. Let V be an n-dimensional vector space over the field F and let W be an m-dimensional vector space over F, then prove that the space $L(V; W)$ is finite dimensional vector space and has dimension mn.

Proof. Given that V is an n-dimensional vector space over F.

$i:e:;\ \dim_F V = n$) every basis of V has n elements.

Let $\mathcal{B} =$ f $_1; \ _2; : : : : \ _n$g be an ordered basis for V.

Also, Given that W is an m-dimensional vector space over F.

$i:e:;\ \dim_F W = m$) every basis of W has m elements.

Let $\mathcal{B}^0 =$ f $_1; \ _2; : : : : \ _n$g be an ordered basis for W.

Now, our wish is to prove that $L(V; W)$ is finite-dimensional and has dimension mn:

$i:e:;$ to prove that every basis of $L(V; W)$ has mn elements.

For each pair of integers $(p; q)$ with $1 \ p \ m$ and $1 \ q \ n$; we define a linear transformation $E^{p;q}$ from V into W by

$$
\begin{aligned}
E^{p;q} &= \begin{cases} 0; & \text{if } x \ 6\!= 1\\ \ _p\,; & \text{if } x = 1 \end{cases}\\
&= \ _{iq} \ _p
\end{aligned}
$$

According to Theorem 1:1, there is a unique linear transformation from V into W.

i:e:; $E^{p;q} : V \rightarrow W$ such that $E^{p;q}(\ _i) = \ _{iq}\ _p$.

Since p varies from 1 to m and q varies from 1 to n and hence linear transformations $E^{p;q}$ are totally mn in number.

Claim: These mn linear transformations $E^{p;q}$ form a basis for $L(V; W)$:

i:e:; to prove that
(i) These mn linear transformations are linearly independent over F .

(ii) These mn linear transformations spans $L(V; W)$ over F .

First, we shall prove that mn linear transformations span $L(V; W)$ over F .Sub-Claim:1 These mn linear transformations E^{pq} span $L(V; W)$ over F .

Let $T \in L(V; W)$

i:e:; let T be a linear transformation from V into W .

For each j , 1 j n: let $A_{1\,j}; A_{2\,j}; : : : ; A_{m\,j}$ be the coordinates of the vector $T(\ _j)$ in the order basis $B^0 = \{\ _1;\ _2; : : : :\ _m\}$:

$$\text{i:e:;}\quad T(\ _j) = \sum_{p=1}^{m} A_{p\,j}B_p \qquad (1.11)$$

Now, we shall prove that every element of $L(V; W)$ is some linear combination of the mn linear transformations E^{pq} .

i:e:; to prove that

$$T = \sum_{p=1}^{m}\sum_{q=1}^{n} A_{pq}E^{pq} \qquad (1.12)$$

Let U be the linear transformation in the right hand member of (1.12).

Then for each j

$$\begin{aligned} U(\ _j) &= \sum_{p=1}^{m}\sum_{q=1}^{n} A_{pq}E^{pq}(\ _j) \\ &= \sum_{p=1}^{m}\sum_{q=1}^{n} A_{pq}\ _{jq}\ _p \\ &= \sum_{p=1}^{m}\left(\sum_{q=1}^{n} A_{pq}\ _{jq}\right)\ _p \\ &= \sum_{p=1}^{m} A_{p\,j}\ _p \\ &= T(\ _j) \\ \Rightarrow\quad U &= T \end{aligned}$$

Sub-Claim:2 The mn linear transformation E^{pq} are linearly independent over F.

i:e:; to prove that any linear combination of $E^{pq} = 0$:

i:e:; to prove that $\sum_{p=1}^{m} A_{pj} \; _p = 0$ then $A_{pj} = 0; \; \forall p$ & j

Let $U = \sum_{p=1}^{m}\sum_{q=1}^{n} A_{pq}E^{pq}$ be the zero transformation, then by definition

$U(\; _j) = 0; \; \forall j$

i:e:; $\sum_{p=1}^{m} A_{pj} \; _p = 0$

The independence of $_p$ implies that $A_{pj} = 0 \; \forall \; p$ & j.

This completes the proof of the theorem.

Theorem 1.6. Let $V; W;$ and Z be vector spaces over the field F. Let T be a linear transformation from V into W and U a linear transformation from W into Z. Then the composed function UT defined by $(UT)(\;) = U(T(\;))$ is a linear transformation from V into Z.

Proof. Our wish is to prove that UT is a linear transformation from V into Z.

i:e:; to prove that $(UT)(c \; + \;) = c(UT)(\;) + (UT)(\;) \quad \forall c \in F; \; ; \in V$

$$\begin{aligned} (UT)(c \; + \;) &= U \; T(c \; + \;) \\ &= U\,(cT(\;) + T(\;) \\ &= c\,[U(T(\;))] + T(\;) \\ &= c(UT)(\;) + (UT)(\;) \end{aligned}$$

This completes the proof of the theorem.

Definition 1.4. If V is a vector space over the field $F;$ a linear operator on V is a linear transformation from V into V.

Note 1.4. If $V = W = Z$, then by theorem (1.6) we see that both U and T are linear operators on the space V.

Also, we see that the composition UT is a linear transformation on V.

In other words, the space $L(V; V)$ has a multiplication defined on it by the composition.

Note that, in general, $TU \neq UT$ (or) $TU \quad UT \neq 0$.

Also, we take a special note that if T is a linear operator on V, then we can compose T with T:

We shall use the notation that $T^2 = TT$ and in general $T^n = T \quad T$ (n times) for $n = 1; 2; 3; : : :$.

We define $T^0 = I$ if T 6= 0:

Lemma 1.1. Let V be a vector space over the field F; let $U; T_1$ and T_2 be linear operators on V; let c be an element of F.

(a) $IU = UI = U$;

(b) $U(T_1 + T_2) = UT_1 + UT_2 = (T_1 + T_2)U = T_1U + T_2U$;

(c) $c(UT_1) = (cU)T_1 = U(cT_1)$:

Proof. (a) Since $I : V$! V is defined by $I( ) = ;$ 8 vectors V.

$$\begin{aligned} IU(\ ) &= I(U(\ )) \\ &= U(\ ) \\ )\quad IU &= U \end{aligned}$$

Similarly, we can prove that $UI = U$. Thus the proof of (a) is complete.

(b) Let 2 V

Consider

$$\begin{aligned} [U(T_1 + T_2)](\ ) &= U[(T_1 + T_2)](\ ) \\ &= U(T_1(\ ) + T_2(\ )) \\ &= (UT_1)(\ ) + (UT_2)(\ ) \\ &= (UT_1 + UT_2)(\ ) \\ U(T_1 + T_2) &= UT_1 + UT_2 \end{aligned}$$

Similarly,

$$\begin{aligned} [(T_1 + T_2)U](\ ) &= (T_1 + T_2)U(\ ) \\ &= (T_1 + T_2)U(\ ) \\ &= T_1(U(\ )) + T_2(U(\ )) \\ &= (T_1U + T_2U)(\ ) \\ )\quad (T_1 + T_2)U &= T_1U + T_2U \end{aligned}$$

This proves (b).

(c) Consider

$$\begin{aligned} [c(UT_1)](\ ) &= [(cU)T_1](\ ) \\ &= (cU)T_1(\ ) \\ &= [(cU)T_1](\ ) \\ )\quad c(UT_1) &= (cU)T_1 \end{aligned}$$

In a similar way, we can prove that $(cU)T_1 = U(cT_2)$:

Thus, the proof of (c) is complete.

Hence the lemma is proved.

Example 1.10. If A is an $m\ n$ matrix with entries in F . Then, we have the linear transformation T from $F^{n\ 1}$ into $F^{m\ 1}$ and is defined by $T(X) = AX$:

If B is a $p\ m$ matrix, then we have have the linear transformation U from $F^{m\ 1}$ into $F^{p\ 1}$ and defined by $U(Y) = BY$:

The composition of UT can be easily described as follows:

$$\begin{aligned} (UT)(X) &= U(T(X)) \\ &= U(AX) \\ &= B(AX) \\ &= (BA)X \end{aligned}$$

Thus, UT is left multiplication by the product matrix BA^{00}:

Example 1.11. Let F be a field and V the vector space of all polynomial function from F into F . Let D be the differentiation operator defined in example (1.3), and let T be the linear operator multiplication by x :

$$(Tf)x = xf(x)$$

We can easily seen that $DT\ 6{=}\ TD$.

In fact, we can easily verify that $DT\ \ TD = I$; the identity operator.

Example 1.12. Let $\mathsf{B} = \text{f}\ \ _1; : : : ;\ \ _n\text{g}$ be an ordered basis for a vector space V . Consider the linear operators $E^{p;q}$ which arose in the proof of the Theorem 1.5.

$$i:e:;\quad E^{p;q}(\ _i) = \ _{iq}\ \ _p$$

These n^2 linear operators form a basis for the space of linear operators on V .

What is $E^{p;q}E^{r;s}$?

$$\begin{aligned}(E^{p;q}E^{r;s})(\;_i) &= E^{p;q}(\;_{is}\;_r)\\ &= \;_{is}E^{p;q}(\;_r)\\ &= \;_{is}\;_{iq}\;_p\end{aligned}$$

Therefore, we have

$$Ep;\,qE^{r;s} = \begin{cases}0; & \text{if } r \;\; q\\ E^{p;s}; & \text{if } q = r\end{cases}$$

Let T be a linear operator on V.

$$\begin{aligned}\text{if}\quad A_j &= \left[T(\;_j)\right]_B\\ A &= [A_1; A_2; \dots; A_n]\\ \text{then}\quad T &= \sum_p\sum_q A_{pq}E^{p;q}\end{aligned}$$

If $U = \sum_r\sum_s B_{rs}E^{r;s}$ is another linear operator on V, then by the above lemma we have

$$\begin{aligned}TU &= \left(\sum_p\sum_q A_{pq}E^{p;q}\right)\left(\sum_r\sum_s B_{rs}E^{r;s}\right)\\ &= \sum_p\sum_q\sum_r\sum_s A_{pq}B_{rs}E^{p;q}E^{r;s}\end{aligned}$$

When $q = r$; and since $E^{p;r}E^{r;s} = E^{p;s}$, then we have

$$\begin{aligned}TU &= \sum_p\sum_s\left(\sum_r A_{pr}B_{rs}\right)E^{p;s}\\ &= \sum_p\sum_s (AB)_{ps}E^{p;s}\end{aligned}$$

Thus, the effect of composing T and U is to multiply the matrices A and B.

Definition 1.5. The function T from V into W is called invertible, if there exists a function U from W into V such that UT is the identity function on V and TU is the identity function on W. If T is invertible, the function U is unique and is denoted by T^{1}:

Further, T is invertible if and only if

1. T is $1:1$, that is $T(\;) = T(\;)$ implies $\;=\;$;
2. T is onto, that is the range of T is (all of) W.

Theorem 1.7. Let V and W be vector spaces over the field F and let T be a linear transformation from V into W . If T is invertible, then the inverse function T^{1} is a linear transformation from W onto V .

Proof. Given that $T : V \to W$ is a linear transformation.

For all $_1; _2 \in V$ and $c \in F$, $T(c _1 + _2) = cT(_1) + T(_2)$

Let $_1; _2$ be the unique vectors in V such that $T(_i) = _i$.

Also given that T is invertible which implies that T^{1} exists.

$$_1 = T^{1}(_1)$$
$$_2 = T^{1}(_2)$$

Now, our wish is to prove that T^{1} is linear transformation from W onto V .

Since T is linear and $T(_i) = _i$, thus we have

$$T(c _1 + _2) = cT(_1) + T(_2) = c _1 + _2$$

Since $_1$ and $_2$ are the unique vectors in V which implies that $c _1 + _2$ is the unique vector in V which is sent by T into $c _1 + _2$ and so

$$\begin{aligned} T^{1}(c _1 + _2) &= c _1 + _2 \\ &= c T^{1}(_1) + T^{1}(_2) \end{aligned}$$

Therefore T^{1} is linear transformation.

Since, T is invertible which implies that T is onto.

Thus, T^{1} is onto linear transformation.

Note 1.5. Suppose that T is an invertible transformation from V onto W and an invertible transformation U from W onto Z . Then UT is also an invertible transformation.

Moreover, $(UT)^{1} = T^{1}U^{1}$:

This conclusion does not require the linearity nor does it involve checking separately that UT is $1 : 1$ and onto. But it requires that $T^{1}U^{1}$ is both a left and right inverse for UT:

Note 1.6. If T is linear then $T(\quad) = T( )\, T( )$:

Hence $T( ) = T( )$ if and only if $T(\quad) = 0$:

Thus, T is one-to-one then $=$ implies that $T( ) = T( )$:

i.e., T is one-to-one if and only if $T( ) = 0$.

Definition 1.6. A linear transformation T is non-singular if $T(\ ) = 0$ implies $\ = 0.$

i.e., if the null space of linear transformation T is $\{0\}.$

Clearly T is $1:1$ if and only if T is non-singular.

Theorem 1.8. Let T be a linear transformation from V into W . Then T is non-singular if and only if T carries each linearly independent subset of V onto a linearly independent subset of W .

Proof. Assume that T is non-singular.

Let S be a linearly independent subset of V.

If ${}_1; {}_2; \ldots; {}_k$ are vectors in S.

$$\begin{aligned} \text{i.e., If } \ c_1\ {}_1 + c_2\ {}_2 + \ldots + c_k\ {}_k &= 0 \\ \Rightarrow \quad c_1 = c_2 = \ldots = c_k &= 0 \end{aligned} \tag{1.13}$$

Now, we shall prove that $T(\ {}_1); T(\ {}_2); \ldots; T(\ {}_k)$ are linearly independent vectors.

$$\begin{aligned} \text{Let } \ c_1T(\ {}_1) + c_2T(\ {}_2) + \ldots + c_kT(\ {}_k) &= 0 \\ \Rightarrow \quad T(c_1\ {}_1) + T(c_2\ {}_2) + \ldots + T(c_k\ {}_k) &= 0 \\ \Rightarrow \quad T(c_1\ {}_1 + c_2\ {}_2 + \ldots + c_k\ {}_k) &= 0 \\ \Rightarrow \quad c_1\ {}_1 + c_2\ {}_2 + \ldots + c_k\ {}_k &= 0 \quad (\because T \text{ is non-singular}) \\ \Rightarrow \quad c_1 = c_2 = \ldots = c_k &= 0 \end{aligned}$$

Thus, the vectors $\{T(\ {}_1); T(\ {}_2); \ldots; T(\ {}_k)\}$ are linearly independent vectors.

Thus, the image of S under T is independent.

Conversely, Assume that T carries a linearly independent subset of V onto linearly independent subsets of W.

Now, we shall prove that T is non-singular.

Let $\ $ be a non-zero vector in $V.$

If $S = \{\ \}$, then the set S is linearly independent. (Since the set consisting of single vector is linearly independent)

By assumption, the set $\{T(\ )\}$ is linearly independent.

Therefore $T(\ ) \neq 0.$

Thus, T is non-singular.

Example 1.13. Let F be a field and let T be the linear operator on F^2 defined by

$$T(x_1; x_2) = (x_1 + x_2; x_1)$$

If $T(x_1; x_2) = 0$, then we have

$$x_1 + x_2 = 0$$
$$x_1 = 0$$

Thus, we have $x_1 = 0; x_2 = 0$.

Therefore T is non-singular.

Let $T : F^2 \to F^2$ and let $(z_1; z_2)$ be any vector in F^2.

Now, our wish is to prove that T is onto.

i:e:; to prove that $(z_1; z_2)$ is in the range of T.

i:e:; we must find scalars x_1 and x_2 such that $T(x_1; x_2) = (z_1; z_2)$:

$$(x_1 + x_2; x_1) = (z_1; z_2)$$
$$\Rightarrow \quad x_1 + x_2 = z_1 \text{ and } x_1 = z_2$$

Upon solving these equations, we get $x_1 = z_1; \quad x_2 = z_1 - z_2$:

Thus T is onto.

Therefore, the explicit formula for computing T^{-1} is

$$T^{-1}(z_1; z_2) = (z_2; z_1 - z_2)$$

Theorem 1.9. Let V and W be finite-dimensional vector spaces over the field F such that $\dim V = \dim W$: If T is a linear transformation from V into W, the following conditions are equivalent:

(i) T is invertible.

(ii) T is non-singular.

(iii) T is onto, that is, the range of T is W.

Proof. Let $\dim V = \dim W = n$:

By Theorem 1.2, we have

$$\operatorname{rank}(T) + \operatorname{nullity}(T) = n \tag{1.14}$$

Assume that T is non-singular. Now, we shall prove that T is onto.

Given that T is non-singular which implies that nullity$(T) = 0$, then from equation (1.14), we have

$$\begin{aligned} \text{rank}(T) &= n \\ \Rightarrow \quad \text{Range of } T &= W \\ \Rightarrow \quad T \text{ is onto.} \end{aligned}$$

Thus the condition (iii) is proved.

Now, we shall assume that T is onto, i.e.; the range of T is $W.$

Given that T is onto, which implies that range of $T = \dim W = n.$

Thus, from equation (1.14), we have

$$\begin{aligned} \text{nullity}(T) &= 0 \\ \Rightarrow \quad T \text{ is non-singular} \end{aligned}$$

Thus, the condition (ii) is proved.

Next, we shall prove that the conditions (ii) and (iii) $\Rightarrow$ (i).

Assume that T is non-singular and T is onto.

We know that T is non-singular if and only if T is $1:1$.

By condition (iii) we have T is onto.

Thus, T is $1:1$ and onto.

Also, we know that T is invertible if and only if T is $1:1$ and onto.

Therefore T is invertible.

This completes the proof of the theorem.

Note 1.7. The above theorem cannot be applied except in the case of finite-dimensionality and $\dim V = \dim W.$

Remark 1.1. Under the hypothesis of Theorem 1.9, the conditions (i); (ii) and (iii) are also equivalent to the following conditions.

(iv) If $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is basis for V, then $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ is a basis for $W.$

(v) There is some basis $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ for V such that $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ is a basis for $W.$

Now, we shall give the proof of the equivalance of five conditions which contains a different proof that (i); (ii) and (iii) are equivalent.

Proof. (i)) (ii) : If T is invertible, then obviously T is non-singular.

(ii)) (iii) : Assume that T is non-singular.

Let f $_1$; $_2$; ; $_n$g be a basis for V.

Then by Theorem 1.8, f$T(\ _1); T(\ _2); :::; T(\ _n)$g is a linearly independent vectors set of vectors in W and since the dimension of W is also n , this set of vectors is a basis for W.

Let f g be any vectors in W. Then there is a scalars $c_1; c_2; :::; c_n$ such that

$$\begin{aligned} &= c_1T(\ _1) + c_2T(\ _2) + ::: + c_nT(\ _n) \\ &= T(c_1\ _1 + c_2\ _2 + ::: + c_n\ _n) \\ ) \quad & 2 \quad \text{range of } T \end{aligned}$$

Thus, T is onto.

(iii)) (iv) : Assume that T is onto.

If f $_1$; $_2$; ldots; $_n$g is any basis for V, the vectors f$T(\ _1); T(\ _2); :::; T(\ _n)$g span the range of T, which is all of W, since T is onto.

Since the dimension of W is n and hence these set of n vectors must be linearly independent.

Thus the set f$T(\ _1); T(\ _2); :::; T(\ _n)$g is a basis for W.

(iv)) (v) : This is quite obvious.

(v)) (i) : Assume that there is some basis f $_1$; $_2$; ldots; $_n$g for V such that f$T(\ _1); T(\ _2); :::; T(\ _n)$g is a basis for W.

Since $T(\ _i)$ spans W and moreover, range of T is all of W.

If $= c_1\ _1 + c_2\ _2 + ::: + c_n\ _n$ is in the null space of T, then

$$\begin{aligned} T(c_1\ _1 + c_2\ _2 + ::: + c_n\ _n) &= 0 \\ ) \quad c_1T(\ _1) + c_2T(\ _2) + \quad + c_nT(\ _n) &= 0 \\ ) \quad c_1 = c_2 = \quad = c_n &= 0 \quad (* \ T(\ _i) \text{ are independent}) \\ ) &= 0 \end{aligned}$$

Therefore nullity of T is f0g:

Thus, the range of T is W and also T is non-singular.

Hence T is invertible.

This completes the proof of the theorem.

Definition 1.7. A group consists of the following:

1. A set G;
2. A rule (or operation) which associates with each pair of elements $x; y$ in G is an element xy in G such a way that
   (a) $x(yz) = (xy)z$ for all $x; y;$ and z in G (associativity)
   (b) there is an element e in G such that $ex = xe = x;$ for every x in G;
   (c) to each element x in G there corresponds an element x^{1} in G such that $xx^{1} = x^{1}x = e:$

Note 1.8. A set of all invertible operators on V together the operation $(U; T)$! UT where $U; T$ are invertible linear operators and the composition UT is an invertible linear operator on V.

1. Composition is an associative operation;

2. The identity operator I satisfies $IT = TI = I$ for each T;

3. For an invertible operator T, then by theorem there is an invertible linear operator T^{1} such that $TT^{1} = T^{1}T = I:$

Thus, the set of invertible linear operators on V together with this operation is a group.

Another example for a group is the set of $n\ n$ matrices with matrix multiplication.

Definition 1.8. A group is called commutative if it satisfies the condition $xy = yx$ for each x and y.

Remark 1.2. The above two examples are not commutative groups.

Let us Sum Up:

In this unit, the students acquired knowledge to the concepts of linear transformation.

the concepts of existence of inverse linear transformation.

Check Your Progress:

1. Find the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T(1;0) = (1;1)$ and $T(0;-1) = (-1;2)$: Prove that T maps the square with vertices $(0;0);(1;0);(1;1)$ and $(0;1)$ into a parallelogram.

2. Let T be a linear transformation on $\mathbb{R}^3$ defined by $T(a;b;c) = (3a; a-b; 2a+b+c)$. Is T invertible? If so, find a rule for T^{-1} like the one which defines T.

Suggested Readings:

1. M. Artin, *Algebra* , Prentice Hall of India Pvt. Ltd., 2005.

2. S.H. Friedberg, A.J. Insel and L.E Spence, *Linear Algebra* , 4th Edition, Prentice-Hall of India Pvt. Ltd., 2009.

3. I.N. Herstein, *Topics in Algebra* , 2nd Edition, Wiley Eastern Ltd, New Delhi, 2013.

4. J.J. Rotman, *Advanced Modern Algebra* , 2nd Edition, Graduate Studies in Mathematics, Vol. 114, AMS, Providence, Rhode Island, 2010.

5. G. Strang, *Introduction to Linear Algebra* , 2nd Edition, Prentice Hall of India Pvt. Ltd, 2013.

Block-I

UNIT-2

LINEAR TRANSFORMATIONS-II

Overview

In this unit, we will illustrate the basic concepts of isomorphsm and linear functionals.

Objectives

After successful completion of this lesson, students will be able to

- understand the concept of linear functionals.
- understand the concept of representation of transformation by matrices.
- explain the concept of isomorphism.

2.1. Isomorphism

Definition 2.1. If V and W are vector spcaes over the field F, any one-one linear transformation T of V onto W is called an isomorphism of V onto W. If there exists an isomorphism of V onto W, we say that V is isomorphic to W.

Note 2.1.

1. The identity operator being an isomorphism of V onto V.
2. If V is isomorphic to W via an isomorphism T; then W is isomorphic to V, because T^{-1} is an isomorphism of W onto V.
3. If V is isomorphic to W and W is isomorphic to Z, then V is isomorphic to Z.

Theorem 2.1. Every n-dimensional vector space over the field F is isomorphic to the space F^n.

Proof. Let V be a vector space over F and let $\dim V = n$.

To prove that $V \cong F^n$.

Let $\{u_1, u_2, \ldots, u_n\}$ be an ordered basis for V.

Every element of $u \in V$ is uniquely expressible as a linear combination of vectors $\{u_1, u_2, \ldots, u_n\}$.

Let $u = \sum_{i=1}^{n} a_i u_i$; $v = \sum_{i=1}^{n} b_i u_i$

Let $a, b \in R$ be arbitrary.

Define $T : V \to F^n$ as follows:

$$f(u) = f\left(\sum_{i=1}^{n} a_i u_i\right) = (a_1; a_2; \ldots; a_n)$$

f is linear:

$$\begin{aligned} f(au+bv) &= f\left(\sum_{i=1}^{n} aa_i u_i + bb_i u_i\right) \\ &= \sum_{i=1}^{n} [aa_i + bb_i]\, u_i \\ &= (aa_1 + bb_1; \ldots; aa_n + bb_n) \\ &= (aa_1; aa_2; \ldots; aa_n) + (bb_1; bb_2; \ldots; bb_n) \\ &= a\,(a_1; a_2; \ldots; a_n) + b\,(b_1; b_2; \ldots; b_n) \\ &= a\,f(u) + b\,f(v) \end{aligned}$$

f is one-one:

$$\begin{aligned} f(u) &= f(v) \\ f\left(\sum_{i=1}^{n} a_i u_i\right) &= f\left(\sum_{i=1}^{n} b_i u_i\right) \\ (a_1; a_2; \ldots; a_n) &= (b_1; b_2; \ldots; b_n) \\ \Rightarrow \quad a_i &= b_i \quad \forall i \\ \Rightarrow \quad \sum_{i=1}^{n} a_i u_i &= \sum_{i=1}^{n} b_i u_i \\ \Rightarrow \quad u &= v \end{aligned}$$

f is onto:

For any given $(a_1; a_2; \ldots; a_n) \in F^n$, there exist $\sum_{i=1}^{n} a_i u_i \in V$ such that

$$f\left(\sum_{i=1}^{n} a_i u_i\right) = (a_1; a_2; \ldots; a_n)$$

Thus, f is one-one and linear map of V onto F^n, which implies that f is an isomorphism of V onto F^n.

$\Rightarrow$ V is isomorphic to F^n.

This completes the proof of the theorem.

2.2. Representation of Transformations by Matrices

Let V be an n-dimensional vector space over the field F and let W be an m-dimensional vector space over F. Let $B = \{ {}_1; {}_2; \; ; {}_n\}$ be an ordered basis for V and $B^0 = \{ {}_1; {}_2; \; ; {}_m\}$ an ordered basis for W. If T is any linear transformation from V into W, then T is determined by its action on the vectors ${}_j$: Each of the n vectors $T({}_j)$ is uniquely expressible as a linear combination of the ${}_i$

$$\text{i:e:;} \quad T({}_j) = \sum_{i=1}^{m} A_{ij} \; {}_i \qquad (2.1)$$

The scalars $A_{1j}; \; ; A_{mj}$ being the coordinates of $T({}_j)$ in the ordered basis B^0: Accordingly, the transformation T is determined by the mn scalars A_{ij} by using the formula ((2.1)). The $m \; n$ matrix A defined by $A(i; j) = A_{ij}$ is called the the matrix of T relative to the pair of ordered basis B and B^0

Theorem 2.2. Let V be an n-dimensional vector space over the field F and W an m-dimensional vector space over F. Let B be an ordered basis for V and B^0 an ordered basis for W. For each linear transformation T from V into W, there is an $m \; n$ matrix A with entries in F such that

$$[T(\;)]_{B^0} = A[\;]_B$$

for every vector in V. Furthermore, $T \rightarrow A$ is a one-one correspondence between the set of all linear transformation from V into W and the set of all $m \; n$ matrices over the field F.

Proof. Let T be a linear transfromation from V into W such that $\dim V = n$ and $\dim W = m$.

Let $B = \{ {}_1; {}_2; \; ; {}_n\}$ and $B^0 = \{ {}_1; {}_2; \; ; {}_m\}$ be an ordered basis for V and W respectively.

If $\in V$ then $= x_1 \, {}_1 + x_2 \, {}_2 + \; + x_n \, {}_n$

$$
\begin{aligned}
T(\ ) &= T\left(\sum_{j=1}^{n} x_j \ {}_j\right)\\
&= \sum_{j=1}^{n} x_j \ T(\ {}_j)\\
&= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} A_{ij} \ {}_i\\
&= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} A_{ij}x_j\right) {}_i
\end{aligned}
$$

If X is the coordinate matrix of in the ordered basis B, then X is an $n \quad 1$ matrix. The product AX is the coordinate matrix of the vector $T(\ )$ in the ordered basis B^0, AX will be an $m\ 1$ matrix.

The j^{th} entry of this column matrix AX will be

$$\sum_{j=1}^{n} A_{ij}x_j$$

If A is any $m \quad n$ matrix over the field F, then

$$T\left(\sum_{j=1}^{n} x_j \ {}_j\right) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} A_{ij}x_j\right) {}_i$$

defines a linear transformation T from V into W , the matrix A is relative to the ordered basis $B;\ B^0$:

Theorem 2.3. Let V be an n-dimensional vector space over the field F and let W be an m-dimensional vector space over F. For each pair of ordered bases $B;\ B^0$ for V and W respectively, the function which assigns to a linear transformation T its matrix relative to $B;\ B^0$ is an isomorphism between the space $L(V;W)$ and the space of all $m\ n$ matrices over the field F.

Proof. Let $B = \{\ {}_1;\ {}_2;\ \ :\ {}_n\}$ and $B^0 = \{\ {}_1;\ {}_2;\ \ ;\ {}_m\}$ be an ordered basis for V and W respectively.

Let M be the vector space of all $m\ n$ matrices over the field F:

Define : $L(V;W) \to M$ by

$$(T) = A_{ij}$$

Let $T_1; T_2 \in L(V;W)$, and let

$$T_1(\ _j) = \sum_{i=1}^{m} a_{ij}\ _i; \quad i = 1; 2; \quad ; n$$

$$\text{and} \quad T_2(\ _j) = \sum_{i=1}^{m} b_{ij}\ _i; \quad i = 1; 2; \quad ; n$$

Now, our claim is to prove that is an isomorphism from $L(V; W)$ onto M .

 is One-One:

$$\begin{aligned} (T_1) &= (T_2) \\ [a_{ij}]_{m\ n} &= [b_{ij}]_{m\ n} \\ \Rightarrow \quad a_{ij} &= b_{ij} \quad \text{for } i = 1; 2; \quad ; m \text{ and } j = 1; 2; \quad ; n \\ \Rightarrow \quad \sum_{i=1}^{m} a_{ij}\ _i &= \sum_{i=1}^{m} b_{ij}\ _i \quad \text{for } j = 1; 2; \quad ; n \\ \Rightarrow \quad T_1(\ _j) &= T_2(\ _j) \quad \text{for } j = 1; 2; \quad ; n \\ \Rightarrow \quad T_1 &= T_2 \end{aligned}$$

$\Rightarrow$ is one-one.

 is onto:

Let $[c_{ij}]_{m\,n} \in M$, then there exists a linear transformation T from V into W such that

$$\begin{aligned} T(\ _j) &= \sum_{i=1}^{m} c_{ij}\ _i; \quad j = 1; 2; \quad ; n \\ (T) &= [c_{ij}]_{m\ n} \end{aligned}$$

$\Rightarrow$ is onto.

Obviously is a linear transformation.

Hence is an isomorphism of $L(V; W)$ onto $M:$

Example 2.1. Let F be a field and let T be the operator on F^2 defined by

$$T(x_1; x_2) = (x_1; 0)$$

show that T is a linear operator on F^2:

Solution. If $\ _1 = (x_1; x_2)$; $\ _2 = (y_1; y_2)$ and c is any scalar.

Given that $T(x_1; x_2) = (x_1; 0) \Rightarrow T(\ _1) = (x_1; 0)$.

Similarly, $T(y_1; y_2) = (y_1; 0) \Rightarrow T(\ _2) = (y_1; 0)$.

$$\begin{aligned} cT(\ _1) + T(\ _2) &= c(x_1; 0) + (y_1; 0) \\ &= (cx_1; 0) + (y_1; 0) \\ &= (cx_1 + y_1; 0) \qquad (2.2) \end{aligned}$$

Now,

$$\begin{aligned} c\ _1 + \ _2 &= c(x_1; x_2) + (y_1; y_2) \\ &= (cx_1; cx_2) + (y_1; y_2) \\ &= (cx_1 + y_1; cx_2 + y_2) \\ ) \quad T(c\ _1 + \ _2) &= (cx_1 + y_1; 0) \qquad (2.3) \end{aligned}$$

From (2.2) and (2.3), we have

$$T(c\ _1 + \ _2) = cT(\ _1) + T(\ _2)$$

) T is a linear transformation from F^2 into F^2:

Let B be the standard ordered basis for F^2.

i:e:; $B = \{\ _1;\ _2\}$, where $\ _1 = (1; 0)$ and $\ _2 = (0; 1)$.

$$\begin{aligned} T(\ _1) &= T(1; 0) \\ &= (1; 0) = 1(1; 0) + 0(0; 1) \\ &= 1(\ _1) + (0)(\ _2) \\ \text{Similarly, } T(\ _2) &= T(0; 1) \\ &= (0; 0) = 0(1; 0) + 0(0; 1) \\ &= (0)\ _1 + (0)\ _2 \\ \text{i:e:; } T(\ _1) &= 1\ _1 + 0\ _2 \\ T(\ _2) &= 0\ _1 + 0\ _2 \end{aligned}$$

) The matrix of T in the ordered basis B is

$$[T]_B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Example 2.2. Let V be the space of all polynomial functions from R into R of the form

$$f(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 \qquad (2.4)$$

that is, the space of polynomial functions of degree three or less.

Let D be the differentiation operator, then D maps V into V.

Let $\mathcal{B}$ be the ordered basis for V containing four functions $f_1; f_2; f_3$ and f_4 defined by $f_j(x) = x^{j-1}$; for $j = 1; 2; 3; 4$:

$$\begin{aligned} \text{i.e.,}\quad f_1(x) &= x^{1-1} = x^0 = 1 \Rightarrow Df_1(x) = 0 \\ f_2(x) &= x^{2-1} = x^1 = x \Rightarrow Df_2(x) = 1 \\ f_3(x) &= x^{3-1} = x^2 \Rightarrow Df_3(x) = 2x \\ f_4(x) &= x^{4-1} = x^3 \Rightarrow Df_4(x) = 3x^2 \end{aligned}$$

Then using (2.4), we have

$$\begin{aligned} Df_1 &= 0f_1 + 0f_2 + 0f_3 + 0f_4 \\ Df_2 &= 1f_1 + 0f_2 + 0f_3 + 0f_4 \\ Df_3 &= 0f_1 + 2f_2 + 0f_3 + 0f_4 \\ Df_4 &= 0f_1 + 0f_2 + 3f_3 + 0f_4 \end{aligned}$$

$\Rightarrow$ The matrix of the operator D in the ordered basis $\mathcal{B}$ is

$$[D]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Theorem 2.4. *Let $V; W$ and Z be finite-dimensional vector spaces over the field F. Let T be a linear transformation from V into W and U a linear transformation from W into Z. If $\mathcal{B}$, $\mathcal{B}'$ and $\mathcal{B}''$ are ordered bases for the vector spaces $V; W$ and Z respectively, if A is the matrix of T relative to the pair $\mathcal{B}$, $\mathcal{B}'$ and $\mathcal{B}''$ is the matrix of U relative to the pair $\mathcal{B}'$; $\mathcal{B}''$; then the matrix of the composition UT relative to the pair $\mathcal{B}; \mathcal{B}''$ is the product matrix $C = BA$:*

Proof. Given that $V; W$ and Z are finite-dimensional vector spaces.

Let $\dim_F V = n$; $\dim_F W = p$; $\dim_F Z = p$:

Also, given that T is a linear transformation from V into W and U is a linear transformation from W into Z.

Let $\mathcal{B}; \mathcal{B}'; \mathcal{B}''$ are ordered bases for the vector spaces $V; W$ and Z respectively.

$$\begin{aligned}
\text{Let } \mathsf{B} &= \{\ {}_1;\ {}_2;\ {}_3;\ \ ;\ {}_n\} \\
\mathsf{B}^0 &= \{\ {}_1;\ {}_2;\ {}_3;\ \ ;\ {}_n\} \\
\mathsf{B}^{00} &= \{\ {}_1;\ {}_2;\ {}_3;\ \ ;\ {}_n\}
\end{aligned}$$

Since, A is the matrix of T relative to the pair $\mathsf{B};\mathsf{B}^0$:

and B is the matrix of U relative to the pair $\mathsf{B}^0;\mathsf{B}^{00}$:

Using our usual convention, that if $\ \in V$, we get

$$[T(\ )]_{\mathsf{B}^0} = A[\ ]_{\mathsf{B}} \tag{2.5}$$

$$[U(T(\ ))]_{\mathsf{B}^{00}} = B[T(\ )]_{\mathsf{B}^0} \tag{2.6}$$

$$\begin{aligned}
\text{Consider}\quad [(UT)(\ )]_{\mathsf{B}^{00}} &= [U(T(\ ))]_{\mathsf{B}^{00}} \\
&= B[T(\ )_{\mathsf{B}^0}] \\
&= BA[\ ]_{\mathsf{B}} \\
\text{i:e:;}\quad [(UT)(\ )]_{\mathsf{B}^{00}} &= BA[\ ]_{\mathsf{B}}
\end{aligned}$$

If C is the matrix of the composition UT relative to the pair $\mathsf{B};\ \mathsf{B}^{00}$, then

$$C = BA$$

For if,

$$\begin{aligned}
(UT)(\ {}_j)\quad (j = 1;2;\ \ ;n) &= U(T(\ {}_j)) \\
&= U\left(\sum_{k=1}^{m} A_{kj}\ {}_k\right) \\
&= \sum_{k=1}^{m} A_{kj}\,(U(\ {}_k)) \\
&= \sum_{k=1}^{m} A_{kj}\left(\sum_{i=1}^{p} B_{ik}\ {}_i\right) \\
&= \sum_{i=1}^{p}\left(\sum_{k=1}^{m} B_{ik}A_{kj}\right)\ {}_i \\
(UT)(\ {}_i) &= \sum_{i=1}^{p} C_{ij}\ {}_i;\quad (j = 1;2;\ \ ;n)\quad \text{where } C_{ij} = \sum_{k=1}^{m} B_{ik}A_{kj}
\end{aligned}$$

If C is the matrix of UT , then $C = BA$.

i:e:; The matrix of the composition UT is the product matrix $C = BA$:

This completes the proof of the theorem.

Note 2.2.

1. If T and U are linear operators on V and we are representing by a single ordered basis $\mathcal{B}$, then above theorem assumes the simple form

$$[UT]_{\mathcal{B}} = [U]_{\mathcal{B}}[T]_{\mathcal{B}}$$

2. The linear operator T is invertible if and only if $[T]_{\mathcal{B}}$ is an invertible matrix.

3. The identity operator I is represented by the identity matrix in any order basis, and thus

$$UT = TU = I$$

is equivalent to

$$[U]_{\mathcal{B}}[T]_{\mathcal{B}} = [T]_{\mathcal{B}}[U]_{\mathcal{B}} = I$$

4. When T is invertible

$$\left[T^{1}\right]_{\mathcal{B}} = [T]_{\mathcal{B}}^{1}$$

Theorem 2.5. Suppose P is an $n \quad n$ invertible matrix over V. Let V be an n-dimensional vector space over F and let $\mathcal{B}$ be an ordered basis of V. Then there is a unique ordered basis $\mathcal{B}^0$ of V such that

(i) $[\ ]_{\mathcal{B}} = P[\ ]_{\mathcal{B}^0}$

(ii) $[\ ]_{\mathcal{B}^0} = P^{1}[\ ]_{\mathcal{B}}$ for every vector in V.

The proof of theorem is not included in the syllabus.

Theorem 2.6. Let V be a finite-dimensional vector space over the field F, and let

$$\begin{aligned} \mathcal{B} &= \{\ _1;\ _2;\ \ ;\ _i\} \\ \text{and} \quad \mathcal{B}^0 &= \{\ ^0_1;\ ^0_2;\ \ ;\ ^0_n\} \end{aligned}$$

be ordered basis for V.

Suppose T is a linear operator on V. If $P = [P_1;\ \ ; P_n]$ is the $n \quad n$ matrix which columns $P_j = \left[\ ^0_j\right]_{\mathcal{B}}$; then

$$[T]_{B'} = P^{-1}[T]_B P.$$

Alternatively, if U is the invertible operator on V defined by $U(\alpha_j) = \alpha'_j$; $j = 1, 2, \cdots, n$, then

$$[T]_{B'} = [U]_B^{-1}[T]_B[U]_B$$

Proof. Let T be a linear operator on the finite dimensional space V, and let

$$\mathcal{B} = \{\alpha_1, \alpha_2, \cdots, \alpha_n\} \text{ and } \mathcal{B}' = \{\alpha'_1, \alpha'_2, \cdots, \alpha'_n\}$$

be two ordered bases for V.

Now, the question is, how are the matrices $[T]_B$ and $[T]_B^0$ are related?

By above theorem, there is a unique (invertible) $n \times n$ matrix P such that

$$[\alpha]_B = P[\alpha]_{B'}; \quad \forall \alpha \in V \tag{2.7}$$

Here P is the matrix $P = [P_1, P_2, \cdots, P_n]$ where $P_j = [\alpha'_j]_B$

By definition

$$[T(\alpha)]_B = [T]_B[\alpha]_B \tag{2.8}$$

Applying (2.7) to the Vector $T(\alpha)$, we have

$$[T(\alpha)]_B = P[T(\alpha)]_{B'} \tag{2.9}$$

Combining (2.7), (2.8) and (2.9), we obtain

$$[T]_B P[\alpha]_{B'} = P[T(\alpha)]_{B'}$$

Premultiplying P^{-1}, we get

$$\begin{aligned} P^{-1}[T]_B P[\alpha]_{B'} &= P^{-1} P[T(\alpha)]_{B'} \\ \Rightarrow P^{-1}[T]_B P[\alpha]_{B'} &= [T(\alpha)]_{B'} \\ \Rightarrow P^{-1}[T]_B P[\alpha]_{B'} &= [T]_{B'}[\alpha]_{B'} \\ \Rightarrow [T]_{B'} &= P^{-1}[T]_B P \end{aligned}$$

This proves the first part of the theorem.

If U is a linear operator, which carries $\mathcal{B}$ onto $\mathcal{B}^0$ is defined by

$$U(\alpha_j) = \alpha'_j \quad (j = 1, 2, \cdots, n) \tag{2.10}$$

i:e:; U carries a basis B onto another basis B^0 of V .

i:e:; U is invertible.

The matrix P (above) is precisely the matrix of the operator U in the ordered basis B .

For if, P is defined by

$$
\begin{aligned}
{}^{0}_{j} &= \sum_{i=1}^{n} P_{ij\ i} \\
\Rightarrow \quad U(\ _j) &= \sum_{i=1}^{n} P_{ij\ i} \qquad [* \ U(\ _j) = {}^{0}_{j}] \\
\Rightarrow \quad P &= [U]_B
\end{aligned}
$$

By first part, we have

$$[U]_B^{\ 1}[T]_B[U]_B = [T]_{B^0} \tag{2.11}$$

Hence the theorem.

Example 2.3. Let T be the linear operator on R^2 defined by $T(x_1; x_2) = (x_1; 0)$ with respect to the ordered basis $B = (\ _1;\ _2)$. What is the matrix T with respect to the ordered basis $B^0 = \{\ {}^0_1 = (1; 1);\ {}^0_2 = (2; 1)\}$

Solution. From Example 2.1, we showed tht the matrix of T in the standard basis $B = \{\ _1;\ _2\}$ is

$$[T]_B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

Suppose B^0 is the ordered basis for R^2 consisting of the vectors

${}^0_1 = (1; 1);\ {}^0_2 = (2; 1)$: Then

$$
\begin{aligned}
{}^0_1 &= \ _1 + \ _2 \\
{}^0_2 &= 2\ _1 + \ _2 \\
P &= \begin{bmatrix} & \\ & \end{bmatrix}
\end{aligned}
$$

so that the matrix P is

We can easily compute that

$$
\begin{array}{ccccc}
 & 1 & 2 & & \\
6 & & 7 & & \\
2 & 1 & 1 & & 3 \\
 & & 1 & 2 & \\
P^{1} = & 6 & & & 7 \\
 & 1 & & 1 &
\end{array}
$$

Thus, we have

$$\begin{aligned}
[T]_{B'} &= P^{-1}[T]_B P \\
&= \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix}
\end{aligned}$$

Example 2.4. Let V be the space of polynomial functions from R into R which have degree less than or equal to three. Let D be the differentiation operator on V, and let

$$\mathcal{B} = \{f_1; f_2; f_3; f_4\}$$

be the ordered basis for V defined by $f_i(x) = x^{i-1}$.

Let t be a real number and define $g_i(x) = (x+t)^{i-1}$, that is

$$\begin{aligned}
g_1 &= f_1 \\
g_2 &= t f_1 + f_2 \\
g_3 &= t^2 f_1 + 2t f_2 + f_3 \\
g_4 &= t^3 f_1 + 3t^2 f_2 + 3t f_3 + f_4
\end{aligned}$$

The matrix P is

$$P = \begin{bmatrix} 1 & t & t^2 & t^3 \\ 0 & 1 & 2t & 3t^2 \\ 0 & 0 & 1 & 3t \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

We can easily compute the invertible matrix P^{-1} is

$$P^{-1} = \begin{bmatrix} 1 & -t & t^2 & -t^3 \\ 0 & 1 & -2t & 3t^2 \\ 0 & 0 & 1 & -3t \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Thus, $\mathcal{B}' = \{g_1; g_2; g_3; g_4\}$ is an ordered basis for V.

We can easily found that the matrix D in the ordered basis $\mathcal{B}$ is

$$[D]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The matrix of D in the ordered basis $\mathcal{B}'$ is

$$P^{-1}[D]_{\mathcal{B}}P = \begin{bmatrix} 1 & -t & t^2 & -t^3 \\ 0 & 1 & -2t & 3t^2 \\ 0 & 0 & 1 & -3t \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & t & t^2 & t^3 \\ 0 & 1 & 2t & 3t^2 \\ 0 & 0 & 1 & 3t \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

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

Thus D represented by the same matrix in the ordered basis $\mathcal{B}$ and $\mathcal{B}'$.

Definition 2.2. Let A and B be $n \times n$ (square) matrices over the field F. We say that B is similar to A over F if there is an invertible $n \times n$ matrix P over F such that $B = P^{-1}AP$.

Note 2.3. According to Theorem 2.6, we have the following observations:

If V is an n-dimensional vector space over F and if $\mathcal{B}$ and $\mathcal{B}'$ are two ordered bases for V, then for each linear operator T on V the matrix $B = [T]_{\mathcal{B}'}$ is similar to the matrix $A = [T]_{\mathcal{B}}$.

Thus the matrix B is similar to A means that on each n-dimensional vector space over F, the matrices A and B represent the same linear transformation in two (possibly) different ordered bases.

Note 2.4.

(i) Note that each $n \times n$ matrix A is similar to itself, by using $P = I$.

(ii) If B is similar to A, then A is similar to B.

(iii) If A is similar to B and B is similar to C, then A is similar to C. Thus, similarity is an equivalance relation on the set of $n \times n$ matrices over the field F.

(iv) The only matrix similar to the identity matrix I is I itself.

(v) The only matrix similar to the zero matrix is the zero matrix itself.

2.3. Linear Functionals

The concept of linear functional is important in the study of finite-dimensional spaces because it helps to organize and clarifty the discussion of subspaces, linear equations, and coordinates.

Definition 2.3. If V is a vector space over the field F, a linear transformation from V into the scalar field F is also called a linear functional on V such that

$$f(c \quad + \quad) = cf(\;) + f(\;)$$

for all vectors and in V and all scalars c in F:

Example 2.5. Let F be a field and let $a_1; a_2; \quad ; a_n$ be scalars in F. Define a function f on F^n by

$$f(x_1; x_2; \quad :x_n) = a_1x_1 + a_2x_2 + \quad + a_nx_n$$

Then f is a linear functional on F^n:

Example 2.6. If A is an $n \; n$ matrix with entries in F, the trace of A is the scalar

$$tr(A) = A_{11} + A_{22} + \quad + A_{nn}$$

The trace function is a linear functional on the matrix space $F^{n \; n}$.

Definition 2.4. If V is a vector space, the collection of all linear functionals on V forms a vector space in a natural way. It is the space $L(V; F)$. We denote this space by V and call it the dual space of V.

$$V \quad = L(V; F)$$

Note 2.5.

1. If V is finite-dimensional, then $\dim V = \dim V$.

2. Let $\mathsf{B} = \{ \;_1; \;_2; \quad ; \;_n \}$ be a basis for V, then there is (for each i) a unique linear functional f_i on V such that

$$f_i(\;_i) = \;_{ij}$$

In this way, we obtain from B a set of n distinct linear functionals $f_1; f_2; \quad ; f_n$ on V . These functionals are also linearly independent.

3. If V has a finite-dimensional and $f_1; f_2; \quad ; f_n$ are linearly independent functionals, and we know that V has dimension n; it must be that B $= \{f_1; f_2; \quad ; f_n\}$ is a basis for V : This basis is called the dual basis of B:

Theorem 2.7. Let V be a finite-dimensional vector space over the field F , and let B $= \{ \ _1; \ _2; \quad ; \ _n\}$ be a basis for V: Then there is a unique dual basis B $= \{f_1; f_2; \quad ; f_n\}$ for V such that $f_i(\ _j) = \ _{ij}$. For each linear functional f on V we have

$$f = \sum_{i=1}^{n} f(\ _i) f_i$$

and for each vector in V we have

$$= \sum_{i=1}^{n} f_i(\ ) \ _i$$

Proof. We have seen above from the note, that there is a unique basis which is dual to B .

If f is a linear functional on V , then f is equal to some linear combination of $f_1; f_2; \quad ; f_n$ where the scalars c_j are given by $c_j = f(\ _j)$.

If $= \sum_{i=1}^{n} x_i \ _i$ is a vector in V , then

$$\begin{aligned} f_j(\ ) &= \sum_{i=1}^{n} x_j f_j(\ _i) \\ &= \sum_{i=1}^{n} x_j \ _{ij} \\ &= x_{ij} \end{aligned}$$

Therefore, the unique expression for is a linear combination of the $\ _1; \ _2; \quad ; \ _n$ is

$$\begin{aligned} &= \sum_{i=1}^{n} x_i \ _i \\ &= \sum_{i=1}^{n} f_i(\ ) \ _i \end{aligned}$$

Remark 2.1.

1. The equation $= \sum_{i=1}^{n} f_i(\ )\ _i$ provides us with a nice way of describing what the dual basis.

   If $B = \{\ _1;\ _2;\ \ ;\ _n\}$ is an ordered basis for V and if $B = \{f_1; f_2;\ \ ; f_n\}$ is the dual basis, then f_i is precisely the function which to each vector in V the i^{th} coordinate of relative to the ordered basis B: Thus, we may also call the f_i; the coordinate functions for B.

2.
$$
\begin{aligned}
f &= \sum f(\ _i)\, f_i \\
&= \sum f_i(\ )\ _i
\end{aligned}
$$

   The formula tells us the following:

$$
\begin{aligned}
\text{Let } f(\ _i) &= a_i; \text{ then} \\
&= x_1\ _1 + x_2\ _2 + \quad + x_n\ _n \\
f(\ ) &= f(x_1\ _1 + x_2\ _2 + \quad + x_n\ _n) \\
&= x_1 f(\ _1) + x_2 f(\ _2) + \quad + x_n f(\ _n) \\
f(\ ) &= a_1 x_1 + a_2 x_2 + \quad + a_n x_n \qquad (2.12)
\end{aligned}
$$

   Therefore, we conclude that if we have choosen an ordered basis B for V and describe each vector in V by its n -tuple of coordinates $(x_1; x_2; \ ; x_n)$ relative to B, then every linear functional on V has the form (2.12).

Now, we shall discuss the relationship between linear functionals and subspaces.

Let f be a non-zero linear functionals.

Note that the co-domain of f is a scalar field F:

Now, f is non-zero, the range of f is non-zero.

) The range of f is non-zero subspace of f; which is a scalar field.

i:e:; The range of $f = 1$:

i:e:; dimension of range of $f = 1$:

i:e:; rank of $f = 1$:

Let V be a finite-dimensional. Then we know that

$$\begin{aligned} \text{rank of } f + \text{nullity of } f &= \dim V \\ \Rightarrow \quad \text{rank of } f + \dim N_f &= \dim V \\ \Rightarrow \quad \dim N_f &= \dim V \quad 1 \end{aligned}$$

Note 2.6.

1. Every hyperspace is an null space of some linear functionals.
2. Each d-dimensional subspace of an n-dimensional space is the intersection of the null spaces of $(n \;\; d)$ linear functionals.

Definition 2.5. If V is a vector space over the field F and S is a subset of V; the annihilator of S is the set S^0 of linear functionals f on V such that $f(\;) = 0$ for every in S:

i:e:; $S^0 = \{ f \in V = f(\;) = 0; \;\forall \;\; \in S \}$

Note 2.7.

1. S^0 is a subspace of V; whether S is a subspace of V or not.
2. If $S = \{0\}$; then $S^0 = V$
3. If $S = V$; then S^0 the zero subspace of V:

Theorem 2.8. Let V be a finite-dimensional vector space over the field F, and let W be a subspace of V. Then

$$\dim W + \dim W^0 = \dim V$$

Proof. Let $\dim W = k$ and $\dim V = n$.

$\Rightarrow$ Let $\{\;_1; \;_2; \;\; ; \;_k\}$ be a basis for W:

Thus, the set $\{\;_1; \;_2; \;\; ; \;_k\}$ is a set of linearly independent vectors in W.

Since W is a subspace of V, and hence this linearly independent set in W can be extended to form a basis of V.

$\Rightarrow$ we can choose vectors $\;_{k+1}; \;_{k+2}; \;\; ; \;_n\}$ in V such that $\{\;_1; \;_2; \;\; ; \;_n\}$ is a basis for V.

Let $\{f_1; f_2; \;\; ; f_n\}$ be the basis for V, which is dual to the basis $\{\;_1; \;_2; \;\; ; \;_n\}$ of V:

Now, our wish is to prove that $\dim W + \dim W^0 = \dim V$:

i.e., to prove that $k + \dim W^0 = n$.

i.e. to prove that $\dim W^0 = n - k$.

i.e., to prove that there exists a basis of the annihilator W^0 consisting of $(n-k)$ elements.

i.e., to prove that $\{f_{k+1}, f_{k+2}, \cdots, f_n\}$ is a basis for the annihilator W^0.

Certainly f_i belongs to W^0 for $i \geq k+1$.

Since $f_i(\alpha_i) = \delta_{ij}$ and $\delta_{ij} = 0$ if $i \geq k+1$ and $j \leq k$.

i.e., $f_i(\alpha) = 0$; $\forall\ \alpha$ = a linear combination of $\alpha_1, \alpha_2, \cdots, \alpha_k$ for all $i \geq k+1$

Let $c_{k+1} f_{k+1} + c_{k+2} f_{k+2} + \cdots + c_n f_n = 0$

Since $f_{k+1}, f_{k+2}, \cdots, f_n$ are linearly independent.

$$\begin{aligned} \Rightarrow \quad (c_{k+1} f_{k+1} + \cdots + c_n f_n)(\alpha) &= 0(\alpha) \\ c_{k+1} f_{k+1}(\alpha) + \cdots + c_n f_n(\alpha) &= 0 \cdot f_{k+1}(\alpha) + \cdots + 0 \cdot f_n(\alpha) \\ \Rightarrow \quad c_{k+1} = c_{k+2} = \cdots = c_n &= 0 \end{aligned}$$

Therefore, the functionals $\{f_{k+1}, \cdots, f_n\}$ are linearly independent.

Now, it remains to prove that $\{f_{k+1}, \cdots, f_n\}$ span W^0.

Suppose $f \in V^*$, then we have

$$f = \sum_{i=1}^{n} f(\alpha_i) f_i \tag{2.13}$$

Also, if $f \in W^0$, then $f(\alpha_i) = 0$ for $i \leq k$

Therefore from (2.13), we have

$$\begin{aligned} f &= f(\alpha_1) f_1 + f(\alpha_2) f_2 + \cdots + f(\alpha_k) f_k + f(\alpha_{k+1}) f_{k+1} + \cdots + f(\alpha_n) f_n \\ \Rightarrow \quad f &= 0 + 0 + \cdots + 0 + f(\alpha_{k+1}) f_{k+1} + f(\alpha_{k+2}) f_{k+2} + \cdots + f(\alpha_n) f_n \\ f &= \sum_{i=k+1}^{n} f(\alpha_i) f_i; \quad \text{where } f \in W^0 \end{aligned}$$

$\Rightarrow \{f_{k+1}, \cdots, f_n\}$ spans W^0.

Thus, we have $\dim W^0 = n - k$

i.e., we have $\dim W + \dim W^0 = \dim V$.

This completes the proof of the theorem.

Example 2.7. Find the dual basis of the basis

$\mathcal{B} = \{(1, -1, 3), (0, 1, -1), (0, 3, -2)\}$ for V.

Solution. Let $\alpha_1 = \{(1, -1, 3);\ \alpha_2 = (0, 1, -1);\ \alpha_3 = (0, 3, -2)\}$.

Then $\mathcal{B} = \{\alpha_1, \alpha_2, \alpha_3\}$.

If $\mathcal{B}' = \{f_1, f_2, f_3\}$ is a dual basis for $\mathcal{B}$, then

$$\begin{aligned} f_1(\alpha_1) &= 1;\ f_1(\alpha_2) = 0;\ f_1(\alpha_3) = 0 \\ f_2(\alpha_1) &= 0;\ f_2(\alpha_2) = 1;\ f_2(\alpha_3) = 0 \\ f_3(\alpha_1) &= 1;\ f_3(\alpha_2) = 0;\ f_3(\alpha_3) = 0 \end{aligned}$$

Now to find an explicit expression for f_1, f_2 and f_3.

Let $a, b, c \in V$, then

$$\begin{aligned} \text{Let } (a, b, c) &= x(1, -1, 3) + y(0, 1, -1) + z(0, 3, -2) \qquad (2.14) \\ &= x\alpha_1 + y\alpha_2 + z\alpha_3 \\ f_1(a, b, c) &= x;\quad f_2(a, b, c) = y;\quad f_3(a, b, c) = z \end{aligned}$$

Now, to find the values of x, y, z.

From (2.14), we have

$$x = a;\quad -x + y + 3z = b;\quad 3x - y - 2z = c$$

Solving these equations, we get $x = a; y = 7a - 2b - 3c;\quad z = b + c - 2a$

$$\begin{aligned} \text{Hence}\quad f_1(a, b, c) &= a \\ f_2(a, b, c) &= 7a - 2b - 3c \\ f_3(a, b, c) &= b + c - 2a \end{aligned}$$

Let us Sum Up:

In this unit, the students acquired knowledge to

- the representation of transformation by matrices.
- the concepts of linear functionals and dual space.

Check Your Progress:

1. Find the dual basis of the basis

$\mathsf{B} = \{(1; \quad 2; 3); (1; \quad 1; 1); (2; \quad 4; 7)\}$ for V.

2. Let W_1 and W_2 be subspaces of a finite-dimensional vector space V.

   (a) Prove that $(W_1 + W_2)^0 = W_1^0 + W_2^0$.

   (b) Prove that $(W_1 \cap W_2)^0 = W_1^0 + W_2^0$.

Suggested Readings:

1. M. Artin, Algebra , Prentice Hall of India Pvt. Ltd., 2005.

2. S.H. Friedberg, A.J. Insel and L.E Spence, Linear Algebra , 4th Edition, Prentice-Hall of India Pvt. Ltd., 2009.

3. I.N. Herstein, Topics in Algebra , 2nd Edition, Wiley Eastern Ltd, New Delhi, 2013.

4. J.J. Rotman, Advanced Modern Algebra , 2nd Edition, Graduate Studies in Mathematics, Vol. 114, AMS, Providence, Rhode Island, 2010.

5. G. Strang, Introduction to Linear Algebra , 2nd Edition, Prentice Hall of India Pvt. Ltd, 2013.

BLOCK - II

Unit - 3: Polynomials

Unit - 4: Polynomials and Commutative Rings.

Block-II

UNIT-3

POLYNOMIALS

Structure

Overview

In this unit, we will illustrate the basic properties of the algebra of polynomials over the field.

Objectives

After successful completion of this lesson, students will be able to

understand the concepts of algebra over a field F.

define a polynomial over the field F.

3.1. Algebras

Definition 3.1. Let F be a field. A linear algebra over the field F is a vector space $\mathcal{A}$ over F with an additional operation called multiplication of vectors which associates with each pair of vectors ; in $\mathcal{A}$ a vector in $\mathcal{A}$ called the product of and in such a way that

(a) multiplication is associative,

$$(\quad) = (\quad)$$

(b) multiplication is distributive with respect to addition,

$$(\ + \ ) = \quad + \quad \text{ and } (\ + \ ) = \quad +$$

(c) for each scalar c in F,

$$c(\quad) = (c \ ) + (c \ )$$

If there is an element 1 in $\mathcal{A}$ such that $1 \quad = \quad 1 =$ for each in $\mathcal{A}$; we call $\mathcal{A}$ a linear algebra with identity over F, and call 1 the identity of $\mathcal{A}$. The algebra $\mathcal{A}$ is called commutative if $=$ for all and in $\mathcal{A}$.

Example 3.1. The set of $n \quad n$ matrices over a field, with the usual operations, is a linear algebra with identity. In particular, the field itself is an algebra with identity.

This algebra is not commutative if $n \quad 2$.

Example 3.2. The space of all linear operators on a vector space, with composition as the product, is a linear algebra with identity. It is commutative if and only if the space is one-dimensional.

Example 3.3. Let F be any field and let S be any non-empty set. Let V be the set of all functions from set S into F.

Define addition in V as:

$$(f+g)(s) = f(s)+g(s); \quad \forall f, g \in V;\ s \in S$$

Define a scalar multiplication in V as:

$$(cf)(s) = c(f(s)); \quad \forall \text{ scalar and } f \in V.$$

with respect to these two operations, the set V becomes a vector spce over F, called the space of functions from a set into field. We shall denote this vector space by $\mathbf{F}^{\infty}$.

Thus, the vectors in F^{∞} are infinite sequences $f = (f_0, f_1, f_2, \dots)$ of scalars f_i in F.

Let a and b be scalars in F.

Let $f = (f_0, f_1, f_2, \dots)$ and $g = (g_0, g_1, g_2, \dots) \in F^{\infty}$.

TThen $af+bg$ is an infinite sequence given by

$$af+bg = (af_0+bg_0, af_1+bg_1, \dots)$$

Define a product in $\mathbf{F}^{\infty}$ by

$$(fg)_n = \sum_{i=1}^{n} f_i g_{n-i} \qquad (n = 0, 1, 2, \dots)$$

$$\text{Thus, } fg = (f_0g_0+f_0g_1+f_1g_0, f_0g_2+f_1g_1+f_2g_0, \dots)$$

and as

$$(gf)_n = \sum_{i=0} g_i f_{n-i} = \sum_{i=0} f_i g_{n-i} = (fg)_n; \quad \text{for } n = 1, 2, \dots;$$
$$\Rightarrow fg = gf$$

Thus, the multiplication is commutative.

Next, we shall prove that the product is associative.

Let $f, g, h \in \mathbf{F}^{\infty}$; then

$$\begin{aligned}(fg)h_n &= \sum_{i=1}^{n} (fg)_i h_{n-i} \\ &= \sum_{i=1}^{n} \left(\sum_{j=0}^{i} f_i g_{i-j} \right) h_{n-i} \\ &= \sum_{i=1}^{n} \sum_{j=0}^{i} f_i g_{i-j} h_{n-i}\end{aligned}$$

$$= \underset{j=0}{\overset{n}{X}} f_j \underset{i=0}{\overset{n\ j}{X}} g_i h_{n\ i\ j}$$

$$= \underset{j=0}{\overset{n}{X}} f_j(gh)_{n\ j} = [f(gh)]_n$$

$$(fg)h_n = [f(gh)]_n \quad \text{for } n = 0;1;2;$$

$$) \quad (fg)h = f(gh)$$

Thus, the multiplication is commutative.

We can easily verify that this operation satisfies

$$(+) = +$$

$$(+) = +$$

$$c( ) = (c\ ) = (c\ ) \quad 8c\,2\,F$$

The vector $1 = (1;0;0;\ ) \,2\, F^1$ serves as an identity for F^1:

) F^1 with the operation defined above is a commutative linear algebra with identity over the field F:

Remark 3.1. The vector (0; 1; 0; 0;) plays a distinguished role in the following discussions and we shall consistenly denote it by x . Throughout this chapter x will never be used to denote an element of the eld F .

Define

$$x^0 = 1$$

$$\text{i:e:;}\ x = (0;1;0;\ \ ;0;\ \ ;)$$

$$x\ \ x = x^2 = (0;0;1;0;\ \ ;0;\ \ )$$

$$x\ \ x\ \ x = x^3 = (0;0;0;1;0;\ \ ;0;\ \ )$$

$$.$$

In general, for each integer $k\quad 0$,

$$(x^k)_k = 1 \quad \text{and} \quad (x^k)_n = o$$

for all non-negative integers $n\, 6\!= k$.

The set $1; x; x^2;$ is both linearly independent and infinite.

Thus the algebra F^1 is not finite-dimensional.

Note 3.1.

1. The algebra $\mathbf{F}^{\infty}$ is sometimes called the algebra of formal power series over F.

2. The element $f = (f_0; f_1; f_2; \cdots)$ is frequently written as

$$f = \sum_{n=0}^{\infty} f_n x^n$$

3.2. The Algebra of Polynomials

In this section, we define a polynomial over the field F.

Definition 3.2. Let $F[x]$ be the subspace of $\mathbf{F}^{\infty}$ spanned by the vectors $1; x; x^2; \cdots$ An element of $F[x]$ is called a polynomial over F.

$F[x]$ consists of all (finite) linear combinations of x and its powers.

$i.e.;$ A non-zero vector f in $\mathbf{F}^{\infty}$ is a polynomial if and only if there is an integer $n \geq 0$ such that $f_n \neq 0$ and such that $f_k = 0$ for all integers $k > n$:

If this integer exists, then it is obviously uique and is called the degree of f and it is denoted by deg f.

Note that, we do not assign a degree to the 0-polynomial.

Note 3.2. If f is a non-zero polynomial of degree n, it follows that

$$f(x) = f_0 x^0 + f_1 x^1 + f_2 x^2 + \cdots + f_n x^n \quad (f_n \neq 0)$$

1. The scalars $f_0; f_1; f_2; \cdots; f_n$ are called coefficients of f and hence we may say that f is a polyomial with coefficents in F:

2. Polynomial of the form cx^0 are called scalar polynomial and frequently we use c for cx^0.

3. A non-zero polynomial f of degree n such that $f_n = 1$ is called a monic polynomial.

Theorem 3.1. Let f and g be non-zero polynomials over F: Then

(i) fg is a non-zero polynomial;

(ii) $\deg (fg) = \deg f + \deg g$;

(iii) fg is a monic polynomial if both f and g are monic polynomials;

(iv) fg is a scalar polynomial if and only if both f and g are scalar polynomials;

(v) if $f + g \neq 0;$ $\deg (f + g) \quad \max(\deg f; \deg g)$:

Proof. Let $\deg f = m$ and $\deg g = n$. If k is a non-negative integer,then

$$(fg)_{m+n+k} = \sum_{i=0}^{m+n+k} f_i g_{m+n+k \;\; i} \tag{3.1}$$

If $f_i g_{(m+n+k) \;\; i} \neq 0;$ then we have

$$\begin{aligned} & i \quad m \quad \text{and} \quad m + n + k \; i \quad n \\ \text{i:e:;} \quad & i \quad m \quad \text{and} \quad m + k \; i \; 0 \\ \text{i:e:;} \quad & m + k < i \quad \text{and} \quad i \quad m \\ \text{i:e:;} \quad & m + k < i \quad \text{and} \quad m \Rightarrow k = 0 \quad (\text{* } k \text{ is non-negative}) \\ \Rightarrow \quad & m + 0 < i \quad \text{and} \quad i \quad m \Rightarrow i = m \end{aligned}$$

If $k = 0$ and $i = m$, then (3.1), becomes

$$\begin{aligned} (fg)_{m+n+0} &= \sum_{i=0}^{m+n} f_m g_{m+n+0 \;\; m} \\ (fg)_{m+n} &= \sum_{i=0}^{m+n} f_m g_n \\ \Rightarrow \quad (fg)_{m+n} &= f_m g_n \quad \text{if } k = 0 && (3.2) \\ \text{and } (fg)_{m+n+k} &= 0 \quad \text{if } k > 0 && (3.3) \end{aligned}$$

(i) If f and g are non-zero polynomials, then from (3.2), we have

$$(fg)_{m+n} = f_m g_n$$

Therefore, fg is a non-zero polynomial.

(ii) If $\deg f = m$ and $\deg g = n$, then from (3.2), we have

$$\begin{aligned} \deg (fg) &= m + n \\ &= \deg f + \deg g \end{aligned}$$

(iii) If f and g are monic polynomials, then (3.2), we have fg is monic polynomial.

(iv) Clearly from (3.2) and (3.3) we have, f and g are scalar polynomials if and only fg are scalar polynomial.

(v) We can easily verify that if $f+g\neq 0;$ $\deg(f+g)\quad \max(\deg f;\deg g):$

Hence the proof.

Corollary 3.1. The set of all polynomials over a given field F equipped with the operations defined by

$$\begin{aligned} af+bg &= (af_0+bg_0;af_1+bg_1;\quad) \\ \text{and}\quad (fg)_n &= \sum_{i=0}^{n} f_i g_{n\,i} \qquad (n=0;1;2;\quad) \end{aligned}$$

is a commutative linear algebra with identity over $F:$

Proof. The set of all polynomials over a given field F is denoted by $F[x]:$

We know that F^1 is a commutative linear algebra with identity over F.

Also, we know that $F[x]$ is a subspace of F^1.

Now, our aim is to prove that $F[x]$ is a commutative linear algebra with identity over $F:$

It is enough to prove that product of two polynomials is again a polynomial.

Let f and g be any two polynomials.

Case 1: Let either $f=0$ or $g=0:$ Then

$$(fg)_n = \sum_{i=0}^{n} f_i g_{n\,i}$$

$\Rightarrow$ product fg is zero:

Case 2: Let either $f\quad 0$ and $g\neq 0:$ Then by part (i) of the above theorem, we have $fg\neq 0:$

Corollary 3.2. Suppose $f;g$ and h are polynomials over the field F such that $f\neq 0$ and $fg=fh:$ Then $g=h:$

Proof. Given that $fg=gh$ and $g\neq 0$.

$$
\begin{aligned}
fg &= fh\\
\Rightarrow\quad fg - fh &= 0\\
\Rightarrow\quad f(g-h) &= 0\\
\Rightarrow\quad g-h &= 0 \quad (\because f \neq 0)\\
\Rightarrow\quad g &= h
\end{aligned}
$$

Note 3.3. Let $f = \sum_{i=0}^{m} f_i x^i$ and $g = \sum_{j=0}^{n} g_j x^j$; then

$$fg = \sum_{s=0}^{m+n} \left(\sum_{r=0}^{s} f_r g_{s-r} \right) x^s \tag{3.4}$$

The above product fg is also given by

$$fg = \sum_{i;j} f_i g_j x^{i+j} \tag{3.5}$$

where the sum is extended over all integers pairs $i; j$ such that $0 \leq i \leq m$; and $0 \leq j \leq n$:

Definition 3.3. Let $\mathcal{A}$ be a linear algebra with identity over the field F: We shall denote the identity of $\mathcal{A}$ by 1 and make the convection that $\alpha^0 = 1$ for each α in $\mathcal{A}$. Then to each polynomial $f = \sum_{i=0}^{n} f_i x_i$ over F and α in $\mathcal{A}$ we associate an element $f(\alpha)$ in $\mathcal{A}$ by the rule

$$f(\alpha) = \sum_{i=0}^{n} f_i \alpha_i$$

Example 3.4. Let C be the field of complex numbers and let $f = x^2 + 2$.

(a) If $\mathcal{A} = C$ and z belongs to C; $f(z) = z^2 + 2$; in particular $f(2) = 6$ and

$$f\left(\frac{1+i}{1-i}\right) = 1:$$

(b) If $\mathcal{A}$ is the algebra of all 2×2 matrices over C and if

$$B = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$$

then

$$f(B) = 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}^2 = \begin{bmatrix} 3 & 0 \\ -3 & 6 \end{bmatrix}$$

(c) If $\mathcal{A}$ is the algebra of all linear operators on C^3 and T is the elements of $\mathcal{A}$ given by

$$T(c_1; c_2; c_3) = \left(i\sqrt{2}c_1; c_2; i\sqrt{2}c_3\right)$$

then $f(T)$ is the linear operator on C^3 defined by

$$f(T)(c_1; c_2; c_3) = (0; 3c_2; 0)$$

(d) If $\mathcal{A}$ is the algebra of all polynomials over C and $g = x^4 + 3i;$ then $f(g)$ is the polynomial in $\mathcal{A}$ given by

$$\begin{aligned} f(g) &= \left(x^4 + 3i\right)^2 + 2 \\ &= 7 + 6ix^4 + x^8 \end{aligned}$$

Theorem 3.2. Let F be a field and $\mathcal{A}$ be a linear algebra with identity over $F:$ Suppose f and g are polynomials over F, that α is an element of $\mathcal{A}$, and that c belongs to F. Then

(i) $(cf + g)(\alpha) = cf(\alpha) + g(\alpha)$

(ii) $(fg)(\alpha) = f(\alpha)g(\alpha)$

Proof. (i) Let $f = \sum_{i=0}^{n} f_i x^i$ and $g = \sum_{j=0}^{m} g_j x^j$, then

$$\begin{aligned} fg &= \sum_{i;j} f_i g_j x^{i+j} \\ (cf + g) &= \sum_{i=0}^{n} (cf_i)x^i + \sum_{j=0}^{m} g_j x^j \\ (cf + g)(\alpha) &= \sum_{i=0}^{n} (cf_i)\alpha^i + \sum_{j=0}^{m} g_j\alpha^j \\ &= c\left(\sum_{i=0}^{n} f_i\alpha^i\right) + \left(\sum_{j=0}^{m} g_j\alpha^j\right) \\ &= cf(\alpha) + g(\alpha) \end{aligned}$$

This proves (i).

(ii) Let $f = \sum_{i=0}^{n} f_i x^i$ and $g = \sum_{j=0}^{m} g_j x^j$, then

$$\begin{aligned} fg &= \sum_{i;j} f_i g_j x^{i+j} \\ \Rightarrow \quad fg(\alpha) &= \sum_{i;j} f_i g_j \alpha^{i+j} = f(\alpha)g(\alpha) \end{aligned}$$

This completes the proof of (ii)

3.3. Polynomial Ideals

In this section we are concerned with results which depend primarily on the multiplicative structure of the algebra of polynomials over the field.

Lemma 3.1. Suppose f and d are non-zero polynomials over a field V such that $\deg d \ \deg f$. Then there exists a polynomial g in $F[x]$ such that either

$$f \quad dg = 0 \quad \text{or} \quad \deg(f \quad dg) < \deg f$$

Proof. suppose

$$f = a_m x^m + \sum_{i=0}^{m\ 1} a_i x^i; \quad a_m \neq 0 \tag{3.6}$$

$$\text{and } d = b_n x^n + \sum_{i=0}^{n\ 1} b_i x^i; \quad b_n \neq 0 \tag{3.7}$$

Given that $\deg d \leq \deg f \Rightarrow n \quad m$ (or) $m \quad n$ and $\geq$

$$f \quad \frac{a_m}{b_n} x^{m\ n} d = 0 \quad \text{or} \quad \deg \ f \quad \frac{a_m}{b_n} x^{m\ n} d < \deg f$$

We may take $g = \frac{a_m}{b_n} x^{m\ n}$.

Using this lemma, we can show that the familar process of long divison of polynomials with real or complex coefficients is possible over any field.

Theorem 3.3. If $f; d$ are polynomials over a field F and d is different from 0 then there exists a polynomial $q; r$ in $F[x]$ such that

(i) $f = dq + r$:

(ii) either $r = 0$ or $\deg r < \deg d$:

The polynomials $q; r$ satisfying (i) and (ii) are unique.

Proof. Case 1: Let $f = 0$ (or) $\deg f < \deg d$:

In this case, let us take $q = 0$ and $r = f:$

Then, both the conditions (i) and (ii) are true.

Case 2: Let $f \neq 0$ and deg $f \quad$ deg $d:$

Then by lemma (3.1), there exists a polynomial g such that

$$f \quad dg = 0 \text{ or } \deg (f \quad dg) < \deg f \tag{3.8}$$

If $f \quad dg \neq 0$ and $\deg (f \quad dg) \quad \deg f$, then taking $f = f \quad dg$ and $g =$ h in(3.8), then there exists a polynomial h such that

$$(f \quad dg) \quad dh = 0 \text{ or } \deg (f \quad dg) \quad dh < \deg (f \quad dg)$$

i:e:; there exists a polynomial h such that

$$(f \quad dg) \quad dh = 0 \text{ or } \deg (f \quad d(g + h) < \deg (f \quad dg)$$

Continuing this process as long as necessary, we ultimately obtain polynomials q and r such that either $r = 0$ (or) $\deg r < \deg d$ and $f = dq + r$.

Now our claim is such polynomials q and r are unique.

If poslsible, let $f = dq_1 + r_1$ where $r_1 = 0$ (or) $\deg r_1 < \deg d:$

$$\begin{aligned} \Rightarrow \quad dq + r &= dq_1 + r_1 \\ \Rightarrow \quad d(q \quad q_1) &= r_1 \quad r \end{aligned}$$

If $q \quad q_1 \quad 0$ then $d(q \quad q_1) \neq 0$ and

$$\deg d + \deg (q \quad q_1) = \deg (r_1 \quad r)$$

But this is a contradiction, since $\deg r_1 \quad r < \deg d$, this is impossible.

Hence $q \quad q_1 = 0$ and also $r_1 \quad r = 0:$

This completes the proof of the theorem.

Definition 3.4. Let d be a non-zero polynomial over the field F. If f is in $F[x]$, the preceding theorem shows there is at most one polynomial q in $F[x]$ such that $f = dq:$ If such a q exists we say that d divides f, that f is divisible by d, that f is a multiple of d and call q the quotient of f and d. We also write $q = f{=}d:$

Corollary 3.3. Let f be a polynomial over the eld $F;$ and let c be an element of F. Then f is divisible by $x \quad c$ if and only if $f(c) = 0:$

Proof. Given a polynomial f, then there exists a polynomial q and r such that

$$\begin{aligned} f(x) &= (x \quad c)q(x) + r(x); \quad \text{where r is a scalar polynomial} \\ f(c) &= (c \quad c)q(c) + r(c) \\ f(c) &= 0 \ q(c) + r(c) \\ ) \quad f(c) &= r(c) \end{aligned}$$

Hence $r = 0$ if and only if $f(c) = 0$:

i:e:; $f = (x \quad c)q$ if and only if $f(c) = 0$.

i:e:; f is divisible by $(x \quad c)$ if and only if $f(c) = 0$.

Hence the proof.

Definition 3.5. Let F be a field. An element c in F is said to be a root or a zero of a given polynomial f over F; if $f(c) = 0$:

Corollary 3.4. A polynomial f of degree n over a field F has at most n roots in F:

Proof. If $\deg f = 0$, which implies that f is a constant, then there is nothing to prove.

If $\deg f = 1$, which implies that f is a monic polynomial, then obviously f has atmost one root.

So, we assume that the theorem is true for polynomials of degree $(n \quad 1)$:

Let f be the polynomial of degree n.

Let a be a root of f:

$$\begin{aligned}) \quad f &= (x \quad a)q \quad \text{where degree of } q = n \quad 1: \\ ) \quad f(b) &= (b \quad a)q(b) \\ ) \quad f(b) &= 0 \ \text{ if and only if } a = b \text{ or } q(b) = 0 \end{aligned}$$

where $q(x)$ is a polynomial of degree $(n\ 1)$ and hence by assumption $q(x)$ has atmost $(n \ 1)$ roots.

Thus, $f(x)$ has atmost n roots.

Hence the proof of the theorem.

Definition 3.6. Let $f = c_0 + c_1 x + c_2 X^2 + \qquad + c_n x^n$: Then the derivative of f is the polynomial given by

$$f'(x) = c_1 + 2c_2x + \cdots + nc_nx^{n-1}$$

Notation:

$$\begin{aligned} f' &= Df \\ f'' &= D^2 f \\ f^{(3)} &= D^3 f \quad \text{and so on} \end{aligned}$$

Remark 3.2. Differentiation is linear, that is D is a linear operator on $F[x]$.

Theorem 3.4 (Taylor's Formulas).
Let F be a field of characteristic zero, c an element of F , and n a positive integer. If f is a polynomial over F with deg $f \leq n$, then

$$f = \sum_{k=0}^{n} \frac{(D^k f)}{k!}(c)(x-c)^k$$

Proof. Taylor's formula is a consequence of the binomial theorem and the linearity of the operators $D, D^2, \ldots, D^n$.

Using Binomial theorem, we get

$$(a+b)^m = \sum_{k=0}^{m} \binom{m}{k} a^{m-k} b^k \tag{3.9}$$

where

$$\begin{aligned} \binom{m}{k} &= \frac{m!}{k!\,(m-k)!} \\ &= \frac{1\cdot 2\cdot 3 \cdots (m-k)(m-k+1)(m-k+2)\cdots(m-1)m}{k!\,(1\cdot 2\cdot 3\cdots(m-k))} \\ &= \frac{(m-k+1)(m-k+2)\cdots(m-1)m}{1\cdot 2\cdot 3\cdots k} \\ &= \frac{(m-0)(m-1)\cdots(m-(k-1))}{1\cdot 2\cdot 3\cdots k} \end{aligned}$$

Consider

$$\begin{aligned} x^m &= [c+(x-c)]^m \\ &= \sum_{k=0}^{m} \binom{m}{k} c^{m-k}(x-c)^k \\ &= \binom{m}{0} c^{m-0}(x-c)^0 + \binom{m}{1} c^{m-1}(x-c)^1 + \cdots + \binom{m}{m} c^{m-m}(x-c)^m \\ x^m &= c^m + mc^{m-1}(x-c) + \cdots + (x-c)^m \end{aligned}$$

when $f = x^m$, the requirement of the theorem is satisfied.

Now, let

$$
\begin{aligned}
f &= \sum_{m=0}^{n} a_m x^m \\
\Rightarrow \quad D^k f(c) &= \sum_{m=0}^{n} a_m (D^k x^m)(c) \\
\Rightarrow \quad \frac{D^k f(c)}{k!} &= \sum_{m=0}^{n} a_m \frac{D^k x^m}{k!}(c) \\
\Rightarrow \quad \frac{D^k f(c)}{k!}(x-c)^k &= \sum_{m=0}^{n} a_m \frac{D^k x^m}{k!}(c)(x-c)^k \\
\Rightarrow \quad \sum_{k=0}^{n} \frac{D^k f(c)}{k!}(x-c)^k &= \sum_{k=0}^{n}\left(\sum_{m=0}^{n} a_m \frac{D^k x^m}{k!}(c)(x-c)^k\right) \\
&= \sum_{m=0}^{n} a_m \left(\sum_{k=0}^{n} \frac{D^k x^m}{k!}(c)(x-c)^k\right) \\
&= \sum_{m=0}^{n} a_m x^m \\
&= f
\end{aligned}
$$

Definition 3.7. If c is a root of the polynomial f; the multiplicity of c as a root of f; is the largest positive integer r such that $(x-c)^r$ divides f:

Note 3.4.

1. Clearly, the multiplicity of a root is less than the degree of f:
2. If f is a polynomial over a field of charcateristic zero, the multiplicity of c; as a root of f; is related to the number of f which are zero at c .

Theorem 3.5. Let F be a field of characteristic zero and f a polynomial over F with $\deg f \leq n$: Then the scalar c is a root of f multiplicity r if and only if

$$
\begin{aligned}
(D^k f)(c) &= 0; \qquad 0 \leq k \leq r-1 \\
(D^r f)(c) &\neq 0
\end{aligned}
$$

Proof. Necessary Part: Let r be the multiplicity of c as a root of f which implies that $(x-c)^r$ divides f .

i:e:; there exists a polynomial g such that

$$
f = (x-c)^r g \quad \text{and} \quad g(c) \neq 0 \tag{3.10}
$$

Now applying Taylor's formula to the function `g' we get

$$g = \sum_{m=0}^{n\ r} \frac{(D^m g)}{m!}(c)(x \quad c)^m \tag{3.11}$$

Using (3.11) in (3.10), we get

$$f = (x \quad c)^r \left[\sum_{m=0}^{n\ r} \frac{(D^m g)}{m!}(c)(x \quad c)^m \right]$$

$$= \sum_{m=0}^{n\ r} \frac{(D^m g)}{m!}(c)(x \quad c)^{m+r}$$

Di erentiating both sides n times, we get

$$D^k(f) = \sum_{m=0}^{n\ r} \frac{(D^{k+m} g)}{m!}(c)(x \quad c)^{m+r}$$

$$\frac{D^k(f)}{k!} = \sum_{m=0}^{n\ r} \frac{(D^{k+m} g)}{m!\,k!}(c)(x \quad c)^{m+r}$$

Thus, we have

$$\frac{D^k(f)(c)}{k!} = \begin{cases} 0 & \text{if } 0 \quad k \quad r \quad 1 \\ \sum_{m=0}^{r} \dfrac{(D^{k\ r} g)(c)}{(k \quad r)!} & \text{if } r \quad k \quad n \end{cases}$$

$$\text{i:e:;} \quad \frac{D^k(f)(c)}{k!} = 0 \quad \text{for} \quad 0 \quad k \quad r \quad 1$$

$$\text{i:e:;} \quad D^k f(c) = 0 \quad \text{for} \quad 0 \quad k \quad r \quad 1 \tag{3.12}$$

When $k = r$, we have

$$\frac{D^k(f)(c)}{k!} = \frac{D^{k\ r}(g)(c)}{(k \quad r)!}$$

$$\frac{D^r(f)(c)}{k!} = \frac{D^{r\ r}(g)(c)}{(r \quad r)!} = \frac{1\ g(c)}{1}$$

$$\Rightarrow D^r(f)(c) = r!\,g(c) \neq 0$$

$$\Rightarrow D^r(f)(c) \neq 0 \tag{3.13}$$

Thus the conditions (3.12) and (3.13) proves the necessary part of the theorem.

Su cient Part: Assume that the conditions (3.12) and (3.13) are true.

Now, our aim is to prove that the scalar c is a root of f of multiplicity r.

i:e:; to prove that there exists the largest positive integer r such that $(x \ c)^r$ divides f.

If possible, assume that, r is not the largest positive integer such that $(x-c)^r$ divides f.

$\Rightarrow$ there exists a polynomial h such that

$$f = (x-c)^{r+1}h \tag{3.14}$$

Note that when the conditions (3.12) and (3.13) are true, then there exists a polynomial g such that

$$f = (x-c)^r g \quad \text{and } g(c) \neq 0 \tag{3.15}$$

From (3.14) and (3.15), we have

$$\begin{aligned} (x-c)^{r+1}h &= (x-c)^r g \\ \Rightarrow g &= (x-c)h \\ \Rightarrow g(c) &= 0 \end{aligned}$$

which is a contradiction to the condition that $g(c) \neq 0$.

$\Rightarrow$ There exists a largest positive integer r such that $(x-c^r)$ divides f.

Hence the proof of the theorem.

Definition 3.8. Let F be a field. An ideal in $F[x]$ is a subspace M of $F[x]$ such that fg belongs to M whenever f is in $F[x]$ and g is in M:

Example 3.5. If F is a field and d is a polynomial over F.

Let $M = dF[x] = \{df = f \in F[x]\}$

M is the set of all multiples df of d by arbitrary f in $F[x]$.

Now, our wish is to prove that M is an ideal.

Since $1 \in F[x]$; $d \cdot 1 \in M \Rightarrow d \in M$:

Thus, M is non-empty.

Next, our claim is that M is a subspace.

For this, let $f; g \in F[x]$ so that $df; dg \in M$.

Let c be a scalar.

$$\text{Consider } c(df) - dg = d(cf - g) \tag{3.16}$$

$$= dh \in M \quad \text{Where } h \in F[x] \tag{3.17}$$

Thus, M is a subspace.

Let $f \in F[x]$ and $g \in F[x]$.

$$f \in F[x] \Rightarrow \exists d \text{ such that } fd \in F[x] \tag{3.18}$$

$$\text{Also } g \in F[x] \Rightarrow \exists d \text{ such that } (fd)g \in F[x] \tag{3.19}$$

which implies $fg \in M.$

Thus, M is an ideal.

Note 3.5. $M = dF[x] = \{df / f \in F[x]\}$ is called the principal ideal generated by $\mathbf{d}$.

Example 3.6. Let $d_1, d_2, \ldots, d_n$ be a finite number of polynomials over F. Then the sum M of the subspaces $d_iF[x]$ is a subspace.

i.e., $M = d_1F[x] + d_2F[x] + \cdots + d_nF[x]$ is also a subspace.

Also, M is an ideal.

For this, let $p \in M.$

Then by definiton, there exists $f_1, f_2, \ldots, f_n \in F[x]$ such that

$$p = d_1 f_1 + d_2 f_2 + \cdots + d_n f_n \tag{3.20}$$

Let g be any arbitrary polynomial over $F.$ Then,

$$\begin{aligned} pg &= d_1(f_1g) + d_2(f_2g) + \cdots + d_n(f_ng) \\ \Rightarrow \quad & pg \in M \quad \forall p \in M \text{ and } g \in F \\ & \Rightarrow \quad M \text{ is an ideal.} \end{aligned}$$

This ideal M is called the principal ideal generated by the polynomial $d_1, d_2, \ldots, d_n$

Theorem 3.6. If F is a field, and M is any non-zero ideal in $F[x]$, there is a unique monic polynomial d in F[x] such that M is the principal ideal generated by d.

Proof. Given that M is a non-zero ideal in $F[x]$.

$\Rightarrow$ M contains atleast one non-zero polynomial. Among all the non-zero polynomials in M, let d be one polynomial with minimal degree.

Without loss of generality, we may assume that d is monic.

Even if not, we can multiply d by a scalar to make it monic.

If $f \in M$ then $f = dq + r$ where either $r = 0$ (or) $\deg r < \deg d$.

Note that d is a monic polynomial in M.

$\Rightarrow dq \in M$ $\quad$ ($\because M$ is an ideal)

and

$f \in M, \; dq \in M \Rightarrow f - dq \in M$ $\quad$ ($\because M$ is a subspace)

Thus $r \in M$ where $\deg r < \deg d$ in M.

This contradiction to the assumption that d is the minimal polynomial.

$\Rightarrow$ The only possibility is that $r = 0$

Thus $f = dq =$ a multiple of d

[$\because f \in M$] is an arbitrary element of M, it follows that every element of M is a multiple of d.

$\Rightarrow$ $M = dF[x]$ (or) M is the principle ideal generated by d.

It remains to prove that d is unique.

If possible, let g be another monic polynomial such that $M = gF[x]$ where $d \in M$.

$\Rightarrow$ there exists non-zero polynomials $p, q \in F[x]$ such that

$$d = gp \tag{3.21}$$

$$\text{and } g = dq \tag{3.22}$$

Now.

$$\begin{aligned}
d &= gp \\
\Rightarrow d &= dqp \\
\deg d &= \deg d + \deg p + \deg q \\
\Rightarrow \deg p + \deg q &= 0 \\
\Rightarrow \deg p &= \deg q = 0 \\
\Rightarrow p &= q = 1 \\
\Rightarrow d &= g
\end{aligned}$$

Thus, d is unique.

This completes the proof of the theorem.

Corollary 3.5. If $p_1, p_2, \ldots, p_n$ are polynomials over a field F, not all of which are 0, there is a unique polynomial d in $F[x]$ such that

(a) d is in the ideal generated by $p_1; p_2; \ldots; p_n$.

(b) d divides each of the polynomials p_i.

Any polynomials satisfying (a) and (b) necessarily satisifes

(c) d is divisible by every polynomial which divides each of the polynomials $p_1; p_2; \ldots; p_n$.

Proof. Let d be the monic generator of the ideal

$$p_1F[x] + p_2F[x] + \cdots + p_nF[x] \qquad (3.23)$$

i:e:; Every member of this ideal is Dividible by d.

i:e:; each of the polynomials p_i is divisible by d.

Now, suppose f is a polynomial which divides each of the polynomials $p_1; p_2; \ldots; p_n$

$$
\begin{aligned}
p{=}f_1 &\Rightarrow \exists \text{ a polynomial } g_1 \text{ such that } p_1 = f g_1\\
p{=}f_2 &\Rightarrow \exists \text{ a polynomial } g_2 \text{ such that } p_2 = f g_2\\
&\vdots\\
p{=}f_n &\Rightarrow \exists \text{ a polynomial } g_n \text{ such that } p_n = f g_n
\end{aligned}
$$

Also (3.23) $\Rightarrow$ $d \in$ the ideal $p_1F[x] + p_2F[x] + \cdots + p_nF[x]$

$\Rightarrow$ $\exists$ polynomials $q_1; q_2; \ldots; q_n \in F[x]$ such that

$$
\begin{aligned}
d &= p_1q_1 + p_2q_2 + \cdots + p_nq_n\\
&= (f g_1)q_1 + (f g_2)q_2 + \cdots + (f g_n)q_n\\
&= f\ g_1q_1 + g_2q_2 + \cdots + g_nq_n
\end{aligned}
$$

$\Rightarrow$ d is divisible by f

$\Rightarrow$ d is divisible by every polynomial which dividies each of the polynomials $p_1; p_2; \ldots; p_n$

Thus, so far we have shown that d is the monic polynomial satisfying the given conditions (a); (b) and (c).

It remains to prove that the uniqueness of d.

If possible, assume that d^0 be any other monic polynomial satisfying conditions (a) and (b).

i:e:; d^0 is the ideal generated by $p_1; p_2; \ldots; p_n$

and

d^0 divides each of the polynomial p_i .

) d^0= a scalar multiple of d .

Also, here d^0 is monic which implies that $d^0 = d$:

Hence the proof.

Definition 3.9. Let $p_1; p_2; \quad ; p_n$ be polynomials over a field F , not all of which are 0 , the monic operator d of the ideal

$$p_1F[x] + \quad + p_nF[x]$$

is called the greatest common divisor (g.c.d) of $p_1; p_2; \quad ; p_n$. This terminology is justified by the preceding corollary. We say that the polynomials $p_1; p_2; \quad ; p_n$ are relatively prime if their greatest common divisor is 1 , or equivalently if the ideal they generate is all of $F[x]$:

Example 3.7. Let F be a subfield of the complex numbers and consider the ideal

$$M \quad = \quad (x + 2)F[x] + (x^2 + 8x + 16)F[x] \tag{3.24}$$

We assert that $M = F[x]$. For M contains

$$x^2 + 8x + 16 \quad x(x + 2) \quad = \quad 6x + 16$$

and hence M contains $6x + 16 \quad 6(x + 2) = 4$:

Thus the scalar polynomial 1 belongs to M as well as its multiplies.

Example 3.8. Let C kbe the field of complex numbers. Then

(a) g.c.d. $(x + 2; x^2 + 8x + 16) = 1$ (See above example)

(b) g.c.d. $((x\ 2)^2(x + i); (x\ 2)(x^2 + 1)) = (x\ 2)(x + i)$:

For the ideal

$$(x \quad 2)^2(x + i)F[x] + (x \quad 2)(x^2 + i) \quad (x \quad 2)(x^2 + 1)F[x]$$

contains

$$(x \quad 2)^2(x + i) \quad (x \quad 2)(x^2 + 1) \quad = \quad (x \quad 2)(x + i)(i \quad 2)$$

Hence it contains $(x\ 2)(x + i)$, which is monic and dividies both $(x\ 2)^2(x + i)$ and $(x\ 2)(x^2 + 1)$.

Example 3.9. Let F be the field of rational numbers and in $F[x]$ let M be the ideal generated by $(x \quad 1)(x+2)^2);\ (x+2)^2(x \quad 3);$ and $(x\ 3)$:

Then M contains

$$\frac{1}{2}\left[(x+2)^2(x \quad 1) \quad (x \quad 2)\right] = (x+2)^2$$

and since

$$(x+2)^2 = (x \quad 3)(x+7)+25$$

M contains the scalar polynomial 1. Thus $M = F[x]$ and the polynomials $(x \quad 1)(x+2)^2);\ (x+2)^2(x \quad 3);$ and $(x\ 3)$ are relatively prime.

Let us Sum Up:

In this unit, the students acquired knowledge to

the principal ideal generated by d.

the concepts of algebra of polynomials.

Check Your Progress:

1. Let F be a subfield of the complex numbers and let A be the following $2\ 2$ matrix over F

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

For each of the following polynomials f over F, compute $f(A)$:

(a) $f = x^2 \quad x+2;$

(b) $f = x^3 \quad 1;$

(c) $x^2 \quad 5x+7$

2. Find the g.c.d of each of the following pairs of polynomials

(a) $2x^5 \quad x^3 \quad 3x^2 \quad 6x+4;\ \ x^4+x^3 \quad 2x \quad 2$.

(b) $3x^4+8x^2\ 3;\ x^3+2x^2+3x+6$.

Suggested Readings:

1. M. Artin, Algebra , Prentice Hall of India Pvt. Ltd., 2005.

2. S.H. Friedberg, A.J. Insel and L.E Spence, Linear Algebra , 4th Edition, Prentice-Hall of India Pvt. Ltd., 2009.

3. I.N. Herstein, Topics in Algebra , 2nd Edition, Wiley Eastern Ltd, New Delhi, 2013.

4. J.J. Rotman, Advanced Modern Algebra , 2nd Edition, Graduate Studies in Mathematics, Vol. 114, AMS, Providence, Rhode Island, 2010.

5. G. Strang, Introduction to Linear Algebra , 2nd Edition, Prentice Hall of India Pvt. Ltd, 2013.

Block-II

UNIT-4

POLYNOMIALS AND COMMUTATIVE RINGS

Structure

Overview

In this unit, we shall prove that each polynomial over the field F can be written as a product of `prime' polynomials.

Objectives

After successful completion of this lesson, students will be able to

understand the concept of monic polynomial in $F[x]$.

understand the concept of Determinant functions.

4.1. The Prime Factorization of a polynomial

Definition 4.1. Let F be a field. A polynomial f in $F[x]$ is said to be reducible over F if there exist polynomials $g; h$ in $F[x]$ of degree 1 such that $f \; gh$ and if not, f is said to be irreducible over F. A non-scalar irreducible polynomial over F is called a prime polynomial over F, and we sometimes say it is a prime in $F[x]$.

Example 4.1. The polynomial $x^2 + 1$ is reducible over the field C of complex numbers.

For if,

$$x^2 + 1 \quad = \quad (x + i)(x \quad i)$$

and the polynomials $x + i; x \quad i$ belongs to $C[x]$.

On the otherhand, $x^2 + 1$ is irreducible over the field R of real numbers.

For if,

$$x^2 + 1 \quad = \quad (ax + b)(a^0x + b^0)$$

with $a; a^0; b; b^0$ in R, then

$$aa^0 \quad = \quad 1; \quad ab^0 + ba^0 = 0; \quad bb^0 = 1 \tag{4.1}$$

On simplification, we get $a^2 + b^2 = 1$, which is impossible with real numbers a and b, unless $a = b = 0$:

Theorem 4.1. Let $p; f$; and g be polynomials over the field F. Suppose that p is a prime polynomial and that p divides the product $f g$: then either p divides f or p divides g.

Proof. Given that p is a prime polynomial and p divides $f g$:

Without loss of generality, we may assume that p is a monic prime polynomial.

Thus, the only divisors of p are 1 and p.

Let $d =$ The g.c.d. of f and p implies that $d=f$ and $d=p$.

) $d = 1$ (or) $d = p$: (* The only divisor of p are 1 and p)

If $d = p$, then p divides f also. Then the theorem is obviously true.

If $d = 1$, then $1 = \text{g.c.d}(f; p)$.

Thus, f and p are relatively prime.

Claim: $p=g$.

g.c.d $(f; p) = 1$) $1 =$ a linear combination of f and p

i:e:; 9 polynomials f_0 and p_0 such that

$$1 = f_0 f + p_0 p \tag{4.2}$$

$$) \quad g = f_0 f g + p_0 p g \tag{4.3}$$

$$= (f g) p_0 + p(p_0 g) \tag{4.4}$$

) p dividies both $(f g)p_0$ and $p(p_0 g)$.

) p dividies $(f g)p_0 + p(p_0 g)$.

i:e:; p dividies g.

Hence the theorem.

Corollary 4.1. If p is a prime and divides a product $f_1; f_2; \quad ; f_n$ then p divides one of the polynomials $f_1; f_2; \quad ; f_n$.

Proof. Here n denotes the number of polynomials in the product.

Now, we shall prove the result by induction on n.

If $n = 2$; then by hypothesis p is a prime and $p=f_1 f_2$.

Then by above theorem, either p divides f_1 or f_2.

Hence the result is true for $n = 2$:

Now, we shall assume that the result is true for $n = k$.

i:e:; $p=(f_1 f_2 \quad f_k)$) $p=f_1$ or $p=f_2 \quad p=f_k$

Now, we shall prove the theorem for $n = k + 1$:

Assume that p is a prime and $p=(f_1 f_2 \quad f_{k+1})$.

Let $g = f_1 f_2 \quad f_n$

Then p divides $g f_{k+1}$.

) p divides g or p divides f_{k+1} (by assumption).

i:e:; p divides $f_1 f_2 \quad f_n$ or p divides f_{k+1}:

Then by assumption p divides f_1 or p divides f_2 or $\quad$ p divides f_{k+1}.

By induction, the theorem is true for all n

Hence the proof.

Theorem 4.2. If F is a field , a non-scalar monic polynomial in F[x] can be factored as a product of monic primes in F[x] in one and except for order, only one way.

Proof. Suppose f is a non-scalar monic polynomial in F[x]: i:e:; over a field F .

Let $\deg f = n$.

Now, we shall prove the result by induction on n.

If $\deg f = 1$ then f is irreducible.

Then there is nothing to prove.

) The theorem is true for $n = 1$:

Let us assume that the theorem is true for all non-scalar monic polynomial f in F[x] of degree $< n$:

Now we shall prove that the theorem is true for any polynomial of degree n:

Case (i): If f is irreducible.

Then f is factored as a product of monic primes and the theorem is complete.

Case (ii): If f is reducible, then by definition $f = gh$, where both f and g are non-scalar monic polynomials of degree $< n$:

Now, g and h are polynomials of degree $< n$:

By using induction hypothesis both g and h can be factored as a product of monic primes in F[x]:

) The product gh can be monic primes in F[x]

$\Rightarrow$ f can be factored as a product of monic primes in $F[x]$.

$\Rightarrow$ It remains to prove that such a product is unique.

If possible assume that f has two such products $p_1, p_2, \cdots, p_m$ and $q_1, q_2, \cdots, q_n$.

i.e., $f = p_1 p_2 \cdots p_m = q_1 q_2 \cdots q_n$

where $p_1, p_2, \cdots, p_m, q_1, q_2, \cdots, q_n$ are monic primes in $F[x]$.

$\Rightarrow$ $p_m = q_1 q_2 \cdots q_n$

By the above corollary,

$\Rightarrow$ $p_m =$ either q_1 (or) q_2 (or) $\cdots$ (or) q_n.

$\Rightarrow$ $p_m = q_i, \ \forall i = 1, 2, \cdots, n$

Now $\deg f = \sum_{i=1} \deg p_i = \sum_{i=1} \deg q_j$

If $m = 1$ and $n = 1$, then there is nothing to prove.

$\Rightarrow$ Let us assume that $m > 1, n > 1$.

By rearranging the numbers $q_1, q_2, \cdots, q_n$, we can have $p_m = q_n$.

$$\Rightarrow p_1 p_2 \cdots p_{m-1} p_m = q_1 q_2 \cdots q_{n-1} p_m$$

$$\Rightarrow p_1 p_2 \cdots p_{m-1} = q_1 q_2 \cdots q_{n-1}$$

Here the polynomial $p_1 p_2 \cdots p_{m-1}$ is of degree less than n.

By using inductive hypothesis $p_1 p_2 \cdots p_{m-1}$ can be factored as a product of monic primes in $F[x]$.

$\Rightarrow$ The product $q_1 q_2 \cdots q_{n-1}$ can only be a rearrangement of the product $p_1 p_2 \cdots p_{m-1}$.

This along with the fact that $q_i = p_m$ implies that the factorisaton of f as a product of monic primes is unique, upto the order of the factors.

This completes the proof of the theorem.

Note 4.1. Let $p_1, p_2, \cdots, p_r$ be distinct monic primes and $n_1, n_2, \cdots, n_r$ denote positive integers such that

$$f = p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$$

Then this decomposition is also unique and is called the primary decomposition of f.

It is easily verified that every monic divisor of f has the form

$$p_1^{m_1} p_2^{m_2} \quad p_r^{m_r}; \quad 0 \quad m_i \quad n_i$$

Example 4.2. Suppose F is a field, and let $a; b; c$ be distinct elements of F. Then the polynomials $x \quad a; x \quad b; x \quad c$ are distinct monic primes in F[x] . If $m; n;$ and s are positive integers, $(x \; c)^s$ is the g.c.d. of the polynomials

$$(x \quad b)^n (x \quad c)^s \quad \text{and} \quad (x \quad a)^m (x \quad c)^s$$

whereas the three polynomials

$$(x \quad b)^n (x \quad c)^s; \qquad (x \quad a)^m (x \quad c)^s \text{and} \quad (x \quad a)^m (x \quad b)^n$$

are relatively prime.

Theorem 4.3. Let f be a non-scalar monic polynomial over the field F and let

$$f \quad = \quad p_1^{n_1} p_2^{n_2} \quad p_k^{n_k}$$

be the prime factorization of f . For each j , 1 j k . let

$$f_j \quad = \quad f = p_j^{n_j} = \quad \underset{i \quad j}{Y} \quad p_1^{n_i}$$

Then $f_1; f_2; \quad ; f_k$ are relatively prime.

Proof. We leave the proof of this theorem to the reader.

Theorem 4.4. Let f be a polynomial over the field F with derivative f^0 . Then f is a product of distinct irreducible polynomials over F if and only if f and f^0 are relatively prime.

Proof. Assume that in the prime factorisation of f over the field F; some (non-scalar) prime polynomial p is repeated.

$$\begin{aligned} \text{i:e:; Assume } f &= p^2 h \qquad \text{where } h \; 2 \; F[x] \\ ) \quad f^0 &= 2pp^0 h \\ &= p(2p^0 h) \end{aligned}$$

) P is a divisor of f^0

) p divides both f and f^0 when p is non-scalar.

Hence f and f^0 are not relatively prime.

Thus, f is not a product of distinct irreducible polynomial over F which implies that f and f^0 are not relatively prime.

Hence the necessary part.

Sufficient Part: Now let

$$f = p_1 p_2 \quad p_k \tag{4.5}$$

where $p_1; p_2; \quad ; p_k$ are distinct, non-scalar irreducible polynomials over F.

) each p_j is a divisor of f.

$$\text{Let } \frac{f}{p_j} = f_j \tag{4.6}$$

$$\text{Then } f^0 = p_1^0 f_1 + p_2^0 f_2 + \quad + p_k^0 f_k \tag{4.7}$$

Let p be a prime polynomial which divides both f and f^0

Then (4.5) implies that $p = p_i$ for some $i = 1; 2; \quad ; k:$

Also p_i divides f_j for $j \quad i$.

Thus, p divides f^0 and we have $p = p_i$.

) p_i divides f^0.

i:e:; p_i divides $p_1^0 f_1 + \quad + p_k^0 f_k$

) p_i divides $\underset{j=1}{X} p^{0^j} f_j$

) p_i must divide each $p_j^0 f_j$ $(j = 1; 2; \quad ; k)$

(or) p_i must divide $p_i^0 f_i$

) p_i divides either f_i (or) p_i divides p_i^0.

But, p_i cannot divide f_i and also p_i cannot divide p_i^0.

Since degree of p_i^0 is one less than the degree of p_i.

Also these imply that no prime polynomial can divide both f and f^0 and hence our assumption is wrong.

Hence f and f^0 are relatively prime.

This completes the proof of the theorem.

Definition 4.2. The field F is called algebrically closed if every prime polynomial over F has degree 1.

4.2. Commutative Rings

In this section we shall prove the essential facts about determinants of square matrices.

Definition 4.3. A ring is a set K, together with two operations $(x; y)$! $x + y$ and $(x; y)$! xy satisfying

(a) K is a commutative group under the operation $(x; y)$! $x + y$ (K is a commutative group under addition);

(b) $(xy)z = x(yz)$ (multipllication is associative);

(c) $x(y + z) = xy + xz; \ (y + z)x = yx + zx$ (the two distributive laws hold)

If $xy = yx$ for all x and y in K, we say that the ring is commutative. If there is an element 1 in K such that $1x = x1 = x$ for each x; K is said to be a ring with identity, and 1 is called the identity for K.

Note 4.2. A field is a commutative ring with non-zero identity such that to each non-zero x there corresponds an element x^{1} with $xx^{1} = 1$:

For example, the set of integers, with the usual addition and multiplication is a commutative ring with identity, but it is not a field (since the multiplicative inverse of any integer is the reciprocal of the integer, which is not in the set of integers).

4.3. Determinant Functions

Let K be a commutative ring with identity. We define an $m \ n$ matrix over K, as a function $A :$ set of integers $(i; j) \ [1 \ i \ m; \ 1 \ j \ n]$) K.

As usual, we represent such a matrix by a rectangular array having m row and n columns.

The sum and product of matrices are defined as

$$\begin{aligned}(A + B)_{ij} &= A_{ij} + B_{ij}\\ (AB)_{ij} &= \sum_{k} A_{ik} B_{kj}\end{aligned}$$

Sum of two matrices A and B is defined when A and B have same number of rows and columns.

Product of two matrices A and B when the number of columns of A is equal to the number of rows of B.

We wish to aassign to each $n\ n$ (square matrix) over K, a scalar (an element of K) known as the determinant of the matrix. It is possible, to define the determinant of a square matrix A by simply writing down a formula for this determinant in terms of entries of A. However such a formula is rather complicated.

We shall define a determinant function on $K^{n\ n}$ as a function which assigns to each $n\ n$ matrix over K -a scalar, where these functions satisfy some special properties.

(i) It is linear as a function of each of rows of the matrix.

(ii) Its value is 0 on any matrix having two equal rows.

(iii) Its value on the $n \quad n$ identity matrix is 1.

Definition 4.4. Let K be a commutative ring with identity, a positive integer, and let D be a function which assigns to each $n\ \ n$ matrix A over K a scalar $D(A)$ in K. We say that D is n-linear if for each $i;\ 1 \quad i \quad n; D$ is a linear function of the ith row when either the other $(n\ 1)$ rows are held fixed.

This definition requires some explaination.

Explanation: If $D : K^{n\ n}$! K is an into function and if ${}_1;\ {}_2;\ \ ;\ {}_n$ denote the n rows of the matrix A 2 $K^{n\ n}$; we also

$$D(A) = D({}_1;\ {}_2;\ \ ;\ {}_n)$$

i:e:; we think of D, as the function of the rows of A.

The statement that D is n-linear means

$$D({}_1;\ {}_2;\ \ ;\ {}_{i\ 1}; c\ {}_i + {}^{0}_{i};\ \ ;\ {}_n) = cD({}_1;\ {}_2;\ \ ;\ {}_i;\ \ ;\ {}_n) + D({}_1;\ {}_2;\ \ ;\ {}_{i\ 1};\ {}^{0}_{i};\ \ ;\ {}_n) \quad (4.8)$$

Note 4.3. If we fix all rows, except the ith row, and then regard D as a function of the i^{th} row, it is often convenient to write $D({}_i)$ instead of $D(A)$.

) (4.8) can be written conveninently as

$$D(c\ _i + {}^0_i) = cD(\ _i) + D(\ {}^0_i)$$

Example 4.3. Let $k_1; k_2; \quad ; k_n$ be positive integers, $1\ k_i\ n$; and let a be an element of K. For each $n\ n$ matrix A over K, define

$$D(A) = aA(1;\ k_1)A(2;\ k_2) \quad A(n; k_n)$$

Then the function D defined above is n-linear.

For if, let us regard D as a function of ith row of A, while the other rows of A are fixed.

Let $D(\ _i) = A(i; k_i)b$ where b is some fixed element of K.

Let ${}^0_i = A^0_{i1}; A^0_{i2}; \quad ; A^0_{in}$.

$$\begin{aligned} D(c\ _i + {}^0_i) &= cA(i; k_i) + A^0(i; k_i)\ b \\ &= cA(i; k_i)b + A^0(i; k_i)b \\ &= cD(\ _i) + D(\ {}^0_i) \quad 8i = 1; 2; \quad ; n: \end{aligned}$$

Thus D is a linear function of each of the rows of A.

Note 4.4. A partciular n-linear function of this type is

$$D(A) = A_{11}A_{22} \quad A_{nn}$$

In otherwords, the product of the diagonal entries is an n-linear function on $K^{n\ n}$.

Example 4.4. Let us find all 2-linear functions on $2\ 2$ matrices over K. Let D be such a function. If we denote the rows of the $2\ 2$ identity matrix by $_1;\ _2$.

i:e:; $\ _1 = (1; 0)$ and $\ _2 = (0; 1)$.

Then we have

$$\begin{aligned} D(A) &= D(A_{11}\ _1 + A_{12}\ _2; A_{21}\ _1 + A_{22}\ _2) \\ &= D(A_{11}\ _1; A_{21}\ _1 + A_{22}\ _2) + D(A_{12}\ _2; A_{21}\ _1 + A_{22}\ _2) \\ &= A_{11}D(\ _1; A_{21}\ _1 + A_{22}\ _2) + A_{12}D(\ _2; A_{21}\ _1 + A_{22}\ _2) \\ &= A_{11}[D(\ _1; A_{21}\ _1) + D(\ _1; A_{22}\ _1)] \\ &\quad +A_{12}[D(\ _2; A_{21}\ _1) + D(\ _2; A_{22}\ _2)] \\ &= A_{11}[A_{21}D(\ _1;\ _1) + A_{22}D(\ _1;\ _2)] \\ &\quad A_{12}[A_{21}D(\ _2;\ _1) + A_{22}D(\ _2:\ _2)] \end{aligned}$$

$$\begin{aligned}
&= A_{11}A_{21}D(\ {}_1;\ {}_1) + A_{11}A_{22}D(\ {}_1;\ {}_2)\\
&\quad A_{12}A_{21}D(\ {}_2;\ {}_1) + A_{12}A_{22}D(\ {}_2;\ {}_2)\\
&= A_{11}A_{21}a + A_{11}A_{22}b + A_{12}A_{21}c + A_{12}A_{22}d
\end{aligned}$$

where $a = D(\ {}_1;\ {}_1)$; $b = D(\ {}_1;\ {}_2)$; $c = D(\ {}_2;\ {}_1)$; $d = D(\ {}_2;\ {}_2)$ are any four scalars in K

Thus, D is a 2-linear function on $2\ \ 2$ matrices over K.

Lemma 4.1. A linear combination of n-llinear functions is n-linear.

Proof. It suffices to prove that a linear combination of two n-linear functions is also n-linear.

Let D and E be n-linear functions. If $a; b\ 2\ K$, the linear combination of D and E are defined by

$$(aD + bE)(A) = aD(A) + bE(A)$$

Let us fix all rows of A except its ith row. Then

$$\begin{aligned}
(aD + bE)(c\ {}_i + {}^0_i) &= aD(c\ {}_i + {}^0_i) + bE(c\ {}_i + {}^0_i)\\
&= a\ \ cD(\ {}_i) + D(\ {}^0_i)\ \ + b\ \ cE(\ {}_i) + E(\ {}^0_i)\\
&= acD(\ {}_i) + aD(\ {}^0_i) + bcE(\ {}_i) + bE(\ {}^0_i)\\
&= [acD(\ {}_i) + bcE(\ {}_i)] +\ \ aD(\ {}^0_i) + bE(\ {}^0_i)\\
&= c\,[aD(\ {}_i) + bE(\ {}_i)] +\ \ aD(\ {}^0_i) + bE(\ {}^0_i)\\
&= c\,[aD + bE]\,(\ {}_i) + [aD + bE]\,(\ {}^0_i)
\end{aligned}$$

) $aD + bE$ is n-linear.

This completes the proof of the lemma.

Note 4.5. If K is a field and V is the set of $n\ \ n$ matrices over K, the above lemma says that the set of all n-linear functions on V is a subspace of the space all functions from V into K.

Example 4.5. Let D be the function defined on $2\ \ 2$ matrices over K by

$$D(A) = A_{11}A_{22}\quad A_{12}A_{21} \tag{4.9}$$

If $D_1(A) = A_{11}A_{22}$ and $D_2(A) = A_{12}A_{21}$, then

$D = D_1 + D_2$.

$i{:}e{:};$ D is a linear combination of two 2-linear functions D_1 and D_2.

Thus, D is a 2-linear function.

Note 4.6.

1. If I is the identity matrix of order 2.
   i:e:; $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ then $D(I) = 1(1) \quad 0(0) = 1$:

   i:e:; If $_1 = (1; 0)$; $_2 = (0; 1)$ then $D(_1; _2) = 1$:

2. If the two rows of A are equal i:e:; $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{11} & A_{12} \end{bmatrix}$

   then $D(A) = A_{11}A_{12} \quad A_{11}A_{12} = 0$:

3. If A^0 is the matrix obtained from $2 \; 2$ matrix A; by interchanging its rows.
   i:e:; $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$; $A^0 = \begin{bmatrix} A_{21} & A_{22} \\ A_{11} & A_{12} \end{bmatrix}$

$$\begin{aligned} \text{then } D(A^0) &= A_{21}A_{12} \quad A_{11}A_{22} \\ &= \quad [A_{11}A_{22} \quad A_{21}A_{12}] \\ &= \quad D(A) \end{aligned}$$

Definition 4.5. Let D be an n-linear function. We say D is alternating (or) (alternate) if the following conditions are satisified:

(a) $D(A) = 0$ whenever two rows of A are equal.

(b) If A^0 is a matrix obtained from A by interchanging two rows of A then $D(A^0) = D(A)$.

Definition 4.6. Let K be a commutative ring with identity, and let n be a positive integer. Suppose D is a function from $n \quad n$ matrices over K into K. We say that D is a determinant function if D is n-linear, alternating and $D(A) = 1$:

Lemma 4.2. let D be a 2-linear function with the property that $D(A) = 0$ for all $2 \; 2$ matrices A over K having equal rows. Then D is alternating.

Proof. Our wish is to prove that if A is a $2 \; 2$ matrix and A^0 is obtained by interchanging the rows of A, then $D(A^0) = D(A)$:

If the rows of A are and , it suffices to show that

$D(\; ; \;) = D(\; ; \;)$.

Given that D is 2-linear.

$$D(\ + \ ; \ \ + \ ) \ = \ D(\ ; \ ) + D(\ ; \ ) + D(\ ; \ ) + D(\ ; \ ) \qquad (4.10)$$

Also, given that $D(A) = 0$; i:e:; $D(\ + \ ; + \ ) = 0$; $D(\ ; \ ) = 0$; $D(\ ; \ ) = 0$

Thus, the equation (4.10) reduces to

$$0 \ = \ D(\ ; \ ) + D(\ ; \ )$$

i:e:; $D(\ ; \ ) \ = \ D(\ ; \ )$

Lemma 4.3. Let D be an n-linear function on n n matrices over K. Suppose D has the property that $D(A) = 0$ whenever two adjacent rows of A are equal. Then D is alternating.

Proof. Now, our aim is to prove that D is alternating.

i:e:; it is enough to prove that

(i) $D(A) = 0$ if any two rows of A are equal.

(ii) $D(A^0) = D(A)$, if A^0 is obtained from A by interchanging any two rows of A.

First, let us assume that, A^0 is obtained from A by interchanging two adjacent rows of A.

Thus, by above lemma, we have) $D(A^0) = D(A)$:

This proves (ii).

Let B be obtained from A, by interchanging the i^{th} and j^{th} rows of A, where $i < j$:

This process can be done as follows:

We begin by interchanging i^{th} row with the $(i + 1)$th rows. We continue, this process, until the rows are in the following order:

$$_1; \quad ; \ _{i \ 1}; \ _{i+1}; \quad ; \ _j; \ _i; \ _{j+1}; \quad ; \ _n \qquad (4.11)$$

The above requires $k = j \ \ 1$ successive interchange of adjacent rows.

In the above order (4.3), let us move $_j$ to the ith position by using $(k \ 1)$ interchange of adjacent rows.

At the end of this, we obtained B from A by performing $K + (k\ \ 1) = 2k\ \ 1$ successive interchanges of adjacent rows.

$$\text{Thus, } D(B) = (\ \ 1)^{2k\ \ 1} D(A) = (\ \ 1)D(A) = \ \ D(A) \qquad (4.12)$$

Suppose A is any $n\ \ n$ matrix with two equal rows says $\ _i = \ _j$ where $i < j$:

If $j = i + 1$; then $\ _i = \ _j$ implies the matrix A has two equal adjacent rows.

Then by (4.12), we have $D(A) = 0$:

If $i > j + 1$, then we interchange $\ _{i+1}$ and $\ _j$ which implies the resulting matrix B has two equal adjacent rows.

Thus from (4.12), we have $D(B) = 0$.

) $D(A) = 0$:

i:e:; $D(A) = 0$.

This proves (i).

Definition 4.7. If $n > 1$ and A is an $n\ \ n$ matrix over K, we let $A(i|j)$ denote the $(n\ \ 1)\ \ (n\ \ 1)$ matrix obtained by deleting the i^{th} row and j^{th} column of A. If D is an $(n\ \ 1)$ linear function and A is an $n\ \ n$ matrix, we put $D_{ij}(A) = D\ \ A(i|j)\ \ $:

Theorem 4.5. Let $n > 1$ and let D be an alternating $(n\ \ 1)$-linear function on $(n\ \ 1)\ \ (n\ \ 1)$ matrices over K. For each j; $1\ \ j\ \ n$, the function E_j defined by

$$E_j(A) = \overset{n}{\underset{i=1}{\Lambda}} (\ \ 1)^{i+j} A_{ij} D_{ij}(A)$$

is an alternating n-linear function on $n\ n$ matrices A. If D is a determinant function, so is each E_j:

Proof. Let A be an $n\ \ n$ matrix.

Then by above definition, we have

$$D_{ij}(A) = D\ \ A(i|j)$$

) $D_{ij}(A)$ is independent of the ith row of A. Since D is $(n\ 1)$ linear.

Thus, D_{ij} is a linear as a function of any row of A, except its ith row.

) $A_{ij}D_{ij}$ is an n-linear function of A.

We know that a linear combination of n-linear functions is also n-linear.

Given that

$$E_j(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} D_{ij}(A) \qquad (4.13)$$

Thus, $E_j(A)$ is a linear combination of n-linear functions.

i:e:; $E_j(A)$ is n-linear.

Now, we shall prove that E_i is alternating.

It is enough to show that $E_j(A) = 0$ whenever A has two equal and adjacent rows.

For this purpose, let the two adjacent rows α_k and α_{k+1} be equal.

(i:e:;) $\alpha_k = \alpha_{k+1}$.

If $i \neq k$ and $i \neq k+1$, the matrix $(A(i|j)$ has two equal rows and thus $D_{ij}(A) = D\ A(i|j) = 0$

(or) $D_{ij}(A) = 0$ for $i \neq k$ and $i \neq k+1$.

In the summation for $E_j(A)$, the only surviving terms are when $i = k$ and $i = k+1$:

$\Rightarrow$ Equation (4.13) we have

$$E_j(A) = (-1)^{k+j} A_{kj} D_{kj}(A) + (-1)^{k+1+j} A_{(k+1)\,j} D_{(k+1)\,j}(A) \qquad (4.14)$$

Here $\alpha_k = \alpha_{k+1}$ (or) The kth and $(k+1)$th rows are equal.

$$\begin{aligned}
\Rightarrow \quad A_{kj} &= A_{(k+1)\,j} && (4.15)\\
\text{and } A(k|j) &= A(k+1|j) && (4.16)\\
\Rightarrow \quad D_{kj}(A) &= D\ A(k|j)\ (A) &&\\
&= D\ A(k+1)|j)\ (A) &&\\
&= D_{(k+1)\,j}(A) &&\\
\text{(or) } D_{kj}(A) &= D_{(k+1)\,j}(A) && (4.17)
\end{aligned}$$

Using (4.15) and (4.17) in (4.14), we get

$$E_j(A) = 0 \qquad (4.18)$$

Thus, E_j is alternating.

Hence proof of part (i) is completed.

Next, we shall prove that E_j is a determinant function, if D is a determinant function.

i:e:; to show that E_j is n-linear, alternating and $E_j(I) = 1$:

From part (i), we have E_j is n-linear and alternating.

Hence, it remains to prove that $E_j(I) = 1$:

Let $I^{(n)}$ denote the $n \quad n$ identity matrix.

) $I^n(j|j)$ is the matrix obtained from $I^{(n)}$ by deleting its jth row and jth column.

) $I^{(n)}(j|j) =$ The $(n \quad 1) \quad (n \quad 1)$ identity matrix $I^{(n \ 1)}$.

Also, note that $I^{(n)}_{ij} = \quad_{ij}$

Now Putting $A = I^{(n)}$ in (4.13), we get

$$\begin{aligned} E_j \ I^{(n)} &= \sum_{i=1}^{n} (\ 1)^{i+j} I^{(n)}_{ij} D_{ij} \ I^{(n)} \\ &= \sum_{i=1}^{n} (\ 1)^{i+j} \ _{ij} D \ I^n(i|j) \\ &= (\ 1)^{2i} \ (1) D \ I_n(j|j) \\ ) \quad E_j \ I^{(n)} &= D \ I^{(n \ 1)} \end{aligned}$$

But $D(I^{(n \ 1)}) = 1$

) $E_j(I^{(n)}) = 1$

Thus, E_j is a determinant function.

This completes the proof of the theorem.

Corollary 4.2. Let K be a commutative ring with identity and let n be a positive integer. Then there exists at least one determinant function on $K^{n \ n}$.

Proof. Let us prove the result by the principle of induction on n.

We know that there exists determinant funciton on $1 \quad 1$ matrices over K and on $2 \ 2$ matrices over K.

Thus D is a determinant function.

Hence the result is true for $n = 1$ (or) $n = 2$:

By the principle of induction, let us assume that the result is true for all $(n \ 1) \ (n \ 1)$ matrices over K.

i:e:; Assume that, there exists determinant function on $K^{(n\ 1)\ (n\ 1)}$

Theorem 4.5 tells us explicitly how to construct a determinant function on $n\ \ n$ matrices.

$)\ \ 9$ a determinant function on $K^{n\ \ n}$.

Thus, the result is true for all n:

This completes the proof of the corollary.

Example 4.6. If B is a $2\ 2$ matrix over K ,we let

$$|B| = B_{11}B_{22} \quad B_{12}B_{21}$$

Then $|B| = D(B)$, where D is the determinant function on $2 \quad 2$ matrices.

Now, we show that this function is unique on $K^{2\ \ 2}$. Let

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \tag{4.19}$$

be a $3 \quad 3$ matrix over K .

If we define $E_1; E_2$ and E_3 as in Theorem 4.5, then

$$\begin{aligned}
E_1(A) &= A_{11} \begin{matrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{matrix} \quad A_{21} \begin{matrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{matrix} + A_{31} \begin{matrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{matrix} \\
E_2(A) &= \quad A_{12} \begin{matrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{matrix} + A_{22} \begin{matrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{matrix} \quad A_{32} \begin{matrix} A_{11} & A_{13} \\ A_{21} & A_{23} \end{matrix} \\
E_3(A) &= A_{13} \begin{matrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{matrix} \quad A_{23} \begin{matrix} A_{11} & A_{12} \\ A_{31} & A_{32} \end{matrix} + A_{33} \begin{matrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{matrix}
\end{aligned}$$

From Theorem 4.5, we conclude that $E_1; E_2$ and E_3 are determinant functions.

Actually, we have to show that $E_1 = E_2 = E_3$

By expanding the each of the above expressions, we can easily verified. Instead of doing this, we give some specific examples.

(a) Let $K = R[x]$ and

$$A = \begin{bmatrix} x\ \ 1 & x^2 & x^3 \\ 0 & x\ \ 2 & 1 \\ 0 & 0 & x\ \ 3 \end{bmatrix}$$

Then,

$$E_1(A) = (x \quad 1) \begin{matrix} x \quad 2 & 1 \\ 0 & x \quad 3 \end{matrix} = (x \quad 1)(x \quad 2)(x \quad 3)$$

$$E_2(A) = x^2 \begin{matrix} 0 & 1 \\ 0 & x \quad 3 \end{matrix} + (x \quad 2) \begin{matrix} x \quad 1 & x^3 \\ 0 & x \quad 3 \end{matrix} = (x \quad 1)(x \quad 2)(x \quad 3)$$

$$\text{and } E_3(A) = x^3 \begin{matrix} 0 & x \quad 2 \\ 0 & 0 \end{matrix} \begin{matrix} x \quad 1 & x^2 \\ 0 & 0 \end{matrix} + (x \quad 3) \begin{matrix} x \quad 1 & x^2 \\ 0 & x \quad 2 \end{matrix}$$

$$= (x \quad 1)(x \quad 2)(x \quad 3)$$

(b) Let $K = R$ and

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Then

$$E_1(A) = \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} = 1$$

$$E_2(A) = \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} = 1$$

$$E_3(A) = \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} = 1$$

Let us Sum Up:

In this unit, the students acquired knowledge to

the prime polynomials.

the properties of determinant functions.

Check Your Progress:

1. Each of the following expression define a function D on the set 3 3 matrices over the field of real numbers. In which of these cases is D a 3-linear functions?
   (a) $D(A) = A_{11} + A_{22} + A_{33}$;

(b) $D(A) = (A_{11})^2 + 3A_{11}A_{22}$;

(c) $D(A) = A_{11}A_{12}A_{13}$;

(d) $D(A) = 0$;

(e) $D(A) = 1$.

2. Let K be a commutative ring with identity. If A is a 2 2 matrix over K , the classical adjoint of A is the 2 2 matrix adj A defined by

$$\mathrm{ad\,j}(A) = \begin{bmatrix} A_{22} & A_{12} \\ A_{21} & A_{11} \end{bmatrix}$$

If det denotes the unique determinant function on 2 2 matrices over K , show that

(a) $(\mathrm{ad\,j}\ A)A = A(\mathrm{ad\,j}\ A) = (\det\ A) = I$;

(b) $\det(\mathrm{ad\,j}\ A) = \det(A)$;

(c) $\mathrm{ad\,j}(A^t) = (\mathrm{ad\,j}\ A)^t$

(A^t denotes the transpose of A)

Suggested Readings:

1. M. Artin, Algebra , Prentice Hall of India Pvt. Ltd., 2005.

2. S.H. Friedberg, A.J. Insel and L.E Spence, Linear Algebra , 4th Edition, Prentice-Hall of India Pvt. Ltd., 2009.

3. I.N. Herstein, Topics in Algebra , 2nd Edition, Wiley Eastern Ltd, New Delhi, 2013.

4. J.J. Rotman, Advanced Modern Algebra , 2nd Edition, Graduate Studies in Mathematics, Vol. 114, AMS, Providence, Rhode Island, 2010.

5. G. Strang, Introduction to Linear Algebra , 2nd Edition, Prentice Hall of India Pvt. Ltd, 2013.

BLOCK - III

Unit – 5: Determinants

Unit – 6: Elementary Canonical Forms-I.

Block-III

UNIT-5

DETERMINANTS

Structure

Overview

In this unit, we shall discuss the uniqueness of the determinant function

Objectives

After successful completion of this lesson, students will be able to

understand the concept of permutation of determinant.

understand the additional properties of determinants.

5.1. Permutations and the Uniqueness of Determinants

In this section, we prove the uniqueness of the determinant function on n n matrices over K. The proof will lead us quite naturally to consider permutations and some of their basic properties.

Definition 5.1. A sequence $(k_1; k_2; \quad ; k_n)$ of positive integers not exceeding n with the property that no two of the k_i are equal is called a permutation of degree n.

Note 5.1. If is a permutation of degree n, one can pass from $(1; 2; \quad ; n)$ to $(\ 1;\ 2;\ ;\ n)$ by a succession of interchanges of pairs, which can be done in several ways.

No matter how it is done, the number of such interchange of pairs, will be always either even or odd. The permutation is then called even or odd respectively.

Theorem 5.1. Let K be a commutative ring with identity and let n be a positive integer. There is precisely one determinant function on the set of n n matrices over K, and it is the function det defined by

$$\det(A) = \sum (\operatorname{sgn}\ )A(1;\ 1) \quad A(n;\ n) \tag{5.1}$$

where sgn is the sign of the permutation :

If D is any alternating n-linear function on $K^{n\ n}$, then for each n n matrix A,

$$D(A) = (\det A)D(I) \tag{5.2}$$

Proof. Suppose D is an alternating n-linear function on n n matrices over K.

Let A be an n n matrix over K whose rows are $_1;\ _2; \quad :\ _n:$

Let $_1:\ _2; \quad ;\ _n$ denote the rows of Identity matrix of order n n; over K.

In this case, we know that

$$ {}_i = \sum_{j=1}^{n} A(i;j)\ {}_j \quad (1 \quad i \quad n) \qquad (5.3)$$

$$\begin{aligned}
\text{Now } D(A) &= D(\ \ ;\ {}_2;\ \ ;\ {}_n) \\
&= D\left(\sum_j A(1;j)\ {}_j;\ {}_2;\ \ ;\ {}_n\right) \quad \text{(taking } i = 1 \text{ in (5.3))} \\
&= \sum_j A(1;j) D(\ {}_j;\ {}_2;\ \ ;\ {}_n) \quad (* \ D \text{ is linear}) \\
&= \sum_j A(1;j) D\left(\ {}_j;\ \sum_k A(2;k)\ {}_k;\ {}_3;\ \ ;\ {}_n\right) \\
&= \sum_j A(1;j) \sum_k A(2;k) D(\ {}_j;\ {}_k;\ \ ;\ {}_n)
\end{aligned}$$

Continuing the process in the same way after a finite number of steps say n, then we have

$$D(A) = \sum_{k_1;k_2;\ ;k_n} A(1;k_1)A(2;k_2) \quad A(n;k_n)D(\ {}_{k_1};\ {}_{k_2};\ \ ;\ {}_{k_n}) \qquad (5.4)$$

In (5.4), the sum is extended over all sequences $(k_1;k_2;\ \ ;k_n)$ of positive integers, whose number does not excee n.

Thus, D is a finite sum of functions, given by $D(A) = aA(1;k_1)A(2;k_2) \quad A(n;k_n)$

Since D is alternating,

$$\Rightarrow \ D(\ {}_{k_1};\ {}_{k_2};\ \ ;\ {}_{k_n}) = 0$$

whenever two of the indices k_i are equal.

i:e:; $D(\ {}_{k_1};\ {}_{k_2};\ \ ;\ {}_{k_n}) = 0$; if the sequence is not a permutation.

In (5.4), it is enough, if we perform the summation only over those sequences which are permutation of degree n.

Note that a finite sequence (or) an n-tuple, is a function defined on the first n positive integers.

$\Rightarrow$ A permutation of degree n may be defined as a $1 - 1$ function from $\{1;2;\ \ ;n\}$ onto $\{1;2;\ \ ;n\}$:

Such a function corresponds to the n-tuple $(\ 1;\ 2;\ \ ;\ n)$ and hence this functions is simply a rule for ordering $1;2;\ \ ;n$ in some well-defined manner.

$\Rightarrow$ if D is an alternating n-linear function and A is an $n\ n$ matrix over K, we then have

$$D(A) \ = \sum A(1;\ 1) \quad A(n;\ n)D(\ _{1};\ \ ;\ _{n}) \tag{5.5}$$

where the sum is extended over distinct permutations of degree n.

Next, we shall prove that

$$D(\ _{1};\ \ ;\ _{n}) \ = \ D(\ _{1};\ \ ;\ _{n}) \tag{5.6}$$

where the sign depends only on the permutation :

The reason for this as follows:

The sequence $(\ 1;\ 2;\ \ ;\ n)$ can be obtained from the sequence $(1; 2;\ \ ; n)$ by a finite number of interchanges of pairs of elements.

For example, if $\ 1 \neq \ 1$, we can transpose 1 and $\ 1$, obtaing $(\ 1;\ \ ; 1;\ \ )$. Proceeding in this way we shall arrive at the sequence $(\ 1;\ \ ;\ n)$ after n or less such interchanges of pairs.

Since D is alternating, the sign of its value changes each time that we interchange two of the rows $_{i}$ and $_{j}$.

Thus, if we pass from $(1; 2;\ \ ; n)$ to $(\ 1;\ 2;\ \ ;\ n)$ through m interchange of pairs $(i; j)$ we then have

$$D(\ _{1};\ \ ;\ _{n}) \ = \ (\ 1)^{m} D(\ _{1};\ \ ;\ _{n}) \tag{5.7}$$

In particular, if D is a determinant function, then we have

$$D(\ _{1};\ \ ;\ _{n}) \ = \ (\ 1)^{m} \tag{5.8}$$

where m depends only upon not upon D.

Thus all determinant functions assign the same value to the matrix with rows epsilon $_{1};\ \ ;\ _{n}$ and this value is either $+1$ or 1:

We define the sign of a permutation by

$$\text{sgn} \ = \begin{cases} +1 & \text{if} \quad \text{is even} \\ 1 & \text{if} \quad \text{is odd} \end{cases}$$

This basic property of permutations canbe deduced from what we understand by determinant function. The integer m occuring in (5.7) is even, if is an even permutation and m is odd, if is an odd permutation.

) (5.7) becomes

$$D(\ _1;\quad ;\ _n) = (\text{sgn}\quad)D(\ _1;\quad ;\ _n)$$

Thus, equation (5.5) becomes

$$\begin{aligned} D(A) &= \sum A(1;\ \ 1)\quad A(n;\ \ n)(\text{sgn}\quad)D(\ _1;\ _2;\quad ;\ _n) \\ &= \quad (\text{sgn}\quad)A(1;\ \ 1)\quad A(n;\ \ n)(\text{sgn}\quad)\ D(I) \end{aligned}$$

where I is the identity matrix of order $n \quad n$ whose rows are $\ _1;\ _2;\quad ;\ _n$.

This implies that, there is precisely one determinant function on $n\ n$ matrices over K.

We call this function by det A and it follows that

$$\det(A) = \sum(\text{sgn}\quad)A(1;\ \ 1)\quad A(n;\ \ n) \qquad (5.9)$$

Thus, we have

$$D(A) = \det(A)D(I)$$

Hence the theorem.

Important Observations:

Now, we have an explicit formula for determinant of an $n\ n$ matrix (5.9) and since this formula involves permuations of degree n, let us conclude this section, by making the following observations about permuations:

1. There are exactly $n! = 1\quad 2\quad n$ permutation of degree n.

   If is such a permutation, there are n possible choices for $\ 1$.

   Once this choice is completed, there are $(n\quad 1)$ choices for $\ 2$; $(n\quad 2)$ choices for $\ 3$;

   $)$ There are $n(n\quad 1)(n\quad 2)\quad 2\ \ 1 = n!$ permuations $\quad :$

2. Since there are $n!$ such permutation $\quad ;$ (5.9) gives $\det(A)$ as a sum of $n!$ terms, one for each permutation of degree n.

3. A given term is a product $A(1;\ 1)\quad A(n;\quad n)$ of n entries of A, one entry from each row and one from each column, and is prefixed by either $+$ or sign according as the permutations is even (or) odd.

4. When permutations are regarded as $1 - 1$ function from the set $\{1; 2; \quad ; n\}$ onto itself, we can define a product of permutations.

   The product of two permutations and will simply be the composed function defined by

$$(\quad)(i) = (\ (i))$$

5. If denotes the identity permutation, then

$$(i) = i$$

6. If is the identity permutation, then each has an inverse ^{1} such that

$$^{1} = \quad ^{1} =$$

From these observations, we can say that the operation of composition, the set of permutations of degree n is a group. This group is usually called the symmetric group of degree n.

Remark 5.1.

$$\text{sgn}\,(\quad) = (\text{sgn}\ )(\text{sgn}\ )$$

In other words is even if both and are either both are even (or) when both are odd and is odd if one of the permutations is even and the other is odd.

Theorem 5.2. Let K be a commutative ring with identity, and let A and B be an $n\ n$ matrices over K. Then

$$\det(AB) = (\det A)(\det B)$$

Proof. Let B be a fixed $n\ n$ matrix K.

For each $n\ n$ matrix A, define

$$D(A) = \det(AB)$$

Denote the rows of A by $_1;\ _2;\quad ;\ _n$:

$$D(\ _1;\ _2;\quad ;\ _n) = \det(\ _1B;\quad ;\ _nB)$$

Here $_jB$ denotes 1 n matrix which is the product of the 1 n matrix $_j$ and the n n matrix B . So that $_jB$ is a matrix of order 1 n (or) $_jB$ is a row matrix.

Also, c $_i$ + ^0_i B = c $_i$B + ^0_iB and det is n -linear.

Thus D is linear.

Next, we shall prove that D is alternating.

i:e:; to p[rove that D($_1$; $_2$; ; $_n$) = 0 if any two rows are equal.

) If $_i$ = $_j$ which implies $_i$B = $_j$B .

Thus, two rows of det ($_1$B; $_2$B; ; $_n$B) are equal.

Hence det ($_1$B; $_2$B; ; $_n$B) is alternating.

) D is alternating.

Thus, D is n -linear and alternating and by Theorem 5.1,

$$D(A) = (\det A)D(I) \tag{5.10}$$

$$\text{But } D(I) = \det (IB)$$

$$) \quad D(I) = \det B \tag{5.11}$$

Substitute (5.11) in (5.10), we get

$$D(A) = (\det A)(\det B)$$

This completes the proof of the theorem.

5.2. Additional Properties of Determinants

In this section, we shall relate some of the useful properties of the determinant function on n n matrices.

Result 1: If A^t denotes the Transpose of A , then prove that

det (A^t) = det (A):

Proof. Let be a permutation of degree n , then

$$\begin{aligned}
A^t(i; \quad i) &= A(\; i; i) \\
\det(A^t) &= \sum (\mathrm{sgn} \quad)A(\quad 1; 1) \quad A(\quad n; n) \\
\text{when } i &= \quad {}^1(j) \\
; A(\quad i; i) &= A(j; \quad {}^1(j)) \\
) \quad A(\quad 1; 1) &= A(1; \quad {}^1 \, 1) \\
&. \\
A(\quad n; n) &= A(n; \quad {}^1 \, n) \\
) \quad A(\quad 1; 1)A(\quad 2; 2) \quad A(\quad n; n) &= A(1; \quad {}^1 \, 1)A(2; \quad {}^1 \, 2) \quad A(n; \quad {}^1 \, n) \\
\text{Also} \quad {}^1 &= I \\
) \quad (\mathrm{sgn} \quad) \; \mathrm{sgn} \quad {}^1 &= 1
\end{aligned}$$

(or) both sgn and sgn ^{1} are either $+1$ (or) they both are either 1:

) in either case, sgn ( ) = sgn ^{1} :

Further, as varies over all permutations of degree n.

^{1} also varies over all permutations of degree n.

$$\begin{aligned}
\det(A^t) &= \sum_{1} \mathrm{sgn} \quad {}^1 \; A(\quad 1; 1) \quad A(\quad n; n) \\
&= \det(A)
\end{aligned}$$

Result 2: If B is obtained from A, b y adding a multiple of one row of A to another (or by adding a multiple of one column of A to another), then

$$\det B = \det A$$

Proof. Let us prove the result for the case of rows. (A similar proof will hold for the case of columns).

Let us assume that B is obtained from A by adding a multiple of row $_j$ to the row $_i$ where $i < j$.

i:e:; B is obtained from A , by adding c $_j + _i$ (where $i < j$).

Since the function det is linear, as a function of ith row, we have

$$\begin{aligned}
\det B &= \det(\alpha_1; \alpha_2; \ldots; \alpha_{i-1}; c\alpha_j + \alpha_i; \ldots; \alpha_j; \ldots; \alpha_n) \quad (* \; i < j)\\
&= \det(\alpha_1; \alpha_2; \ldots; \alpha_{i-1}; \alpha_i; \ldots; \alpha_j; \ldots; \alpha_n)\\
&\quad +(\alpha_1; \alpha_2; \ldots; \alpha_{i-1}; c\alpha_j; \ldots; \alpha_j; \ldots; \alpha_n)\\
&= \det(\alpha_1; \alpha_2; \ldots; \alpha_{i-1}; \alpha_i; \ldots; \alpha_j; \ldots; \alpha_n)\\
&\quad +c(\alpha_1; \alpha_2; \ldots; \alpha_{i-1}; \alpha_j; \ldots; \alpha_j; \ldots; \alpha_n)\\
&= \det A + 0\\
&= \det A
\end{aligned}$$

Hence the result.

Result 3: Suppose we have an $n \times n$ matrix of the Block Form

$$\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

where A is an $r \times r$ matrix, C is an $s \times s$ matrix, B is $r \times s$, and 0 denote the $s \times r$ zero matrix. Then

$$\det \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = (\det A)(\det C)$$

Proof. Let us define

$$D(A; B; C) = \det \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} \tag{5.12}$$

Now our claim is to prove that $D(A; B; C) = (\det A)(\det C)$.

Let us fix A and B, and allow C to vary.

(We know that D is alternating and C is an $s \times s$ matrix).

$\therefore$ D is alternating and s-linear function of the rows of C.

Hence by theorem we have

$$D(A; B; C) = (\det C)D(A; B; I) \tag{5.13}$$

where I is the identity matrix of order $s \times s$.

Now, consider $D(A; B; I)$:

$$D(A; B; I) = D(A; 0; I) \tag{5.14}$$

$$\Rightarrow D(A; 0; I) = (\det A)D(I; 0; I) \tag{5.15}$$

But,

$$D(I; 0; I) = \det \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} = 1 \qquad (5.16)$$

$$\Rightarrow \quad D(A; 0; I) = \det A \qquad (5.17)$$

$$\Rightarrow \quad D(A; B; I) = \det A \qquad (5.18)$$

Thus, from (5.13), we have

$$D(A; B; C) = (\det C)(\det A)$$

Hence the problem.

Example 5.1. Suppose K is the field of rational numbers and we wish to compute the determinant of the 4 4 matrix

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

Solution. Given that

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

By subtracting suitable multiples of row 1 from rows 2, 3 and 4, we obtain the matrix

$$= \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 4 & 4 & 4 \\ 0 & 5 & 9 & 13 \\ 0 & 3 & 1 & 3 \end{bmatrix}$$

If we subtract $\frac{5}{4}$ of row 2 from row 3 and then subtract $\frac{3}{4}$ of row 2 from row 4, we obtain

$$B = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 4 & 4 & 4 \\ 0 & 0 & 4 & 0 \end{bmatrix}$$

and again $\det B = \det A$. The block form of B tells us that

$$\det A \;=\; \det B = \begin{array}{cccc} 1 & 1 & 4 & 8 \\ & & & \\ 0 & 4 & 4 & 0 \end{array} = 4(32) = 128$$

Definition 5.2. The $n\ n$ matrix $ad\,j\ A$; which is the transpose of the matrix of cofactors of A, is called the classical adjoijnt of A.

Definition 5.3. An $n\ n$ matrix A over K is said to be invertible over K if there is an $n\ n$ matrix $A^{\ 1}$ with entries in K, such that

$$AA^{\ 1} \quad A^{\ 1}A = I \tag{5.19}$$

Theorem 5.3. Let A be an $n \quad n$ matrix over K. Then A is invertible over K if and only if $\det A$ is invertible in K. When A is invertible, the unique inverse for A is

$$A^{\ 1} \;=\; (\det A)^{\ 1}\, ad\,j\ A$$

In particular, an $n\ n$ matrix over a field is invertible if and only if its determinant is different from zero.

Proof. Now, let $n > 1$ and let A be an $n \quad n$ matrix over K. We have already seen that we can construct a determinant function on $n\ n$ matrices, if we are given a $(n\ 1)\ \ (n\ 1)$ matrix. We also know that a determinant function is unique.

Then

$$E_j(A) \;=\; \sum_{i=1}^{n} (\quad)\,1^{\,i+j}\, A_{ij} D_{ij}(A)$$

is an alternating n-linear function on $n \quad n$ matrices.

If we fix any jth column,

$$\det A \;=\; \sum_{i=1}^{n} (\ \ 1)^{i+j} A^{ij} \det A(i|j) \tag{5.20}$$

Here the scalars $(\ \ 1)^{i+j}\det A(i|j)$ is usually called the $i;\ j$ cofactor of A (or) the cofactor of the $(i;\ j)$th entry of A.

Let $C_{ij} = (\ \ 1)^{i+j}\det A(i|j)$ then we have

$$\det A \;=\; \sum_{i=1}^{n} A_{ij}C_{ij} \tag{5.21}$$

where the cofactor C_{ij} is $(-1)^{i+j}$ times the determinant of the $(n-1)\times(n-1)$ matrix obtained by deleting the ith row and jth column of A.

Next, out claim is to prove that $\sum_{i=1}^{n} A_{ik}C_{ij} = 0$ if $j \neq k$:

For, replace the jth column of A by the kth column of A and call the resulting matrix B.

i:e:; The matrix B has two identical column in jth and kth columns.

$\Rightarrow \det B = 0$

Since $B(i|j) = A(i|j)$,we have

$$
\begin{aligned}
0 &= \det B \\
&= \sum_{i=1}^{n} (-1)^{i+j} B_{ij} \det B(i|j) \\
&= \sum_{i=1}^{n} (-1)^{i+j} A_{ik} \det A(i|j) \\
&= \sum_{i=1}^{n} A_{ik}C_{ij}
\end{aligned}
$$

$$\Rightarrow \quad \sum_{i=1}^{n} A_{ik}C_{ij} = 0 \text{ if } j \neq k:$$

These properties of the cofactors can be summarized by

$$\sum_{i=1}^{n} A_{ik}C_{ij} = \delta_{jk} \det A \qquad (5.22)$$

By the definition of classical adjoint of A, we have

$$
\begin{aligned}
(\text{adj}\, A)_{ij} &= C_{ji} && (5.23) \\
&= (-1)^{i+j} \det A(j|i) && (5.24)
\end{aligned}
$$

The formulas (5.22) can be summarised in matrix equation

$$(\text{adj}\, A)A = (\det A)I \qquad (5.25)$$

It can also be proved that $A(\text{adj}\, A) = (\det A)I$.

Since $A^t(i|j) = A(i|j)^t$, we have

$$(-1)^{i+j} \det A^t(i|j) = (-1)^{j+i} \det A^t(j|i) \qquad (5.26)$$

which simply says that the i,j cofactor of A^t is the j,i cofactor of A.

Thus, we have

$$\mathrm{adj}\,(A^t) = (\mathrm{adj}\,A)^t \tag{5.27}$$

By applying (5.25) to A^t, we have

$$(\mathrm{adj}\,A^t)A^t = (\det A^t)I = (\det A)I$$

Taking tranpose on both sides, we get

$$A(\mathrm{adj}\,A^t)^t = (\det A)I$$

Using (5.27), we have

$$A(\mathrm{adj}\,A) = (\det A)I$$

The facts $(\mathrm{adj}\,A)A = (\det A) = I$ and $A(\mathrm{adj}\,A) = (\det A)I$ tells us the following fact about the invertibility of matrices over K.

If the element $\det A$ has a multiplicative inverse in K, then A is invertible and

$$A^{\;1} = (\det A)^{\;1}\,\mathrm{adj}\,A \tag{5.28}$$

is the unique inverse of A.

Conversely, if A is invertible over K, the element $\det A$ is invertible in K.

Note 5.2. Similar matrices have the same determinant, that is if P is invertible over K and $B = P^{\;1}AP$, then $\det B = \det A$:

Cramers Rule:

Now, we shall discuss for solving systems of linear equations.

Suppose A is an $n\;n$ matrix over the field F and we wish to solve the system of linear equations $AX = Y$ for some given n-tuple $(y_1; y_2; \quad ; y_n)$.

If $AX = Y$, then we have

$$\begin{aligned}
(\mathrm{adj}\,A)AX &= (\mathrm{adj}\,A)Y \\
\Rightarrow\quad (\det A)X &= (\mathrm{adj}\,A)Y \\
\Rightarrow\quad (\det A)x_j &= \sum_{i=1}^{n} (\mathrm{adj}\,A)_{ji}y_i \\
&= \sum_{i=1}^{n} (\;1)^{i+j} y_i \det A(i|j)
\end{aligned}$$

This last expression is the determinant of the $n \times n$ matrix obtained by replacing the jth column of A by Y.

If $\det A = 0$, then there is nothing to discuss.

So, $\det A \neq 0$. Let A be an $n \times n$ matrix over the field F such that $\det A \neq 0$. If $y_1; y_2; \ldots; y_n$ are any scalars in F, the unique solution $X = A^{-1}Y$ of the system of equation $AX = Y$ is given by

$$x_j = \frac{\det B_j}{\det A}; \quad j = 1; 2; \ldots; n$$

where B_j is the $n \times n$ matrix obtained from A by replacing the jth column of A by Y.

Let us Sum Up:

In this unit, the students acquired knowledge to

find the value of the determinant.

the inverse of the matrices.

Check Your Progress:

1. If K is a commutative ring with identity and A is the matrix over K given by

$$A = \begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}$$

Show that $\det A = 0$.

2. Prove that the determinant of the Vandermonde matrix

$$\begin{bmatrix} 1 & a & a^2 \\ 1 & c & c^2 \end{bmatrix}$$

is $(b-a)(c-a)(a-b)$:

3. Use the classical adjoint formula to compute the inverse of each of the following 3×3 real matrices.

(a)

$$\begin{bmatrix} 2 & 3 & 2 \\ 6 & 0 & 3 \\ 4 & 1 & 1 \end{bmatrix}$$

(b)

$$\begin{bmatrix} \cos & 0 & \sin \\ 0 & 1 & 0 \\ \sin & 0 & \cos \end{bmatrix}$$

4. Use Cramer's rule to solve each of the following systems of linear equations over the eld of rational numbers.

(a)

$$\begin{aligned} x + y + z &= 11 \\ 2x \quad 6y \quad z &= 0 \\ 3x + 4y + 2z &= 0 \end{aligned}$$

(b)

$$\begin{aligned} 3x \quad 2y &= 7 \\ 3y \quad 2z &= 6 \\ 3z \quad 2x &= 1 \end{aligned}$$

Suggested Readings:

1. M. Artin, Algebra , Prentice Hall of India Pvt. Ltd., 2005.

2. S.H. Friedberg, A.J. Insel and L.E Spence, Linear Algebra , 4th Edition, Prentice-Hall of India Pvt. Ltd., 2009.

3. I.N. Herstein, Topics in Algebra , 2nd Edition, Wiley Eastern Ltd, New Delhi, 2013.

4. J.J. Rotman, Advanced Modern Algebra , 2nd Edition, Graduate Studies in Mathematics, Vol. 114, AMS, Providence, Rhode Island,

2010.

5. G. Strang, Introduction to Linear Algebra , 2nd Edition, Prentice Hall of India Pvt. Ltd, 2013.

Block-III

UNIT-6

ELEMENTARY CANONICAL FORMS-I

Overview

In this unit, we shall discuss the characteristic value and characteristic vector of a linear transformation.

Objectives

After successful completion of this lesson, students will be able to

- understand the concept of minimal polynomial.
- explain the concept of diagonalization.

6.1. Introduction

Our principal aim is to study linear transformation on finite-dimensional vector spaces. On this front, we have seen many specific examples of linear transformation and proved some few theorem about the general linear transformations.

In the finite-dimensional case,we have used ordered bases to represent Linear transformation by matrices. We have explored the vector space $L(V; W)$ consisiting of linear transformation from V to W and then we studied $L(V; V)$, consisiting of linear transformations of V into itself.

Given the linear operator on an n-dimensional space V. If we could find an ordered basis $B = \{ {}_1; {}_2; \; ; {}_n \}$ of V in which T can be represented by a diagonal matrix DS of the form.

$$D = \begin{bmatrix} c_1 & 0 & 0 & & 0 \\ 0 & c_2 & 0 & & 0 \\ 0 & 0 & c_3 & & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & c_n \end{bmatrix}$$

We can gain considerable information about T. For example, the numbers like rank of T and determinantn of T can be determined by simply looking at D:

Now, few questions are raised now?

1. Can each linear operator T be represented by a diagonal matrix in some ordered basis? If not, for which operators T does such a basis exists?

2. If there is such a basis, how to find it?

3. If there is no such basis, what is the simplest type of matrix, by which we can represent T ?

6.2. Characteristic Values

Note that we can explicitly describe the range space and null space of T by using D. Since $[T]_B = D$ if and only if $T(\ _k) = c_k\ _k$ 8k = 1; 2; ; n the range will be nothing but the subspace spanned by those ^0_ks whose coefficient c_k = neq0 and the null space will be the subspace spanned by those ^0_ks whose coefficient $c_k = 0$.

In otherwords, we can study vectors which are sent by T into scalar multiples of themselves.

Definition 6.1. Let V be a vector space over the field F and let T be a linear operator on V. A characteristic value of T is a scalar c in F such that there is a non-zero vector in V with $T(\ ) = c$: If c is a characteristic value of T, then

1. any such that $T(\ ) = c$ is called a characteristic vector of T associated with the characteristic value c;

2. the collection of all such that $T(\ ) = c$ is called the characteristic space associated with c.

Note 6.1. Characteristic values are often called characteristic roots, latent roots, eigen values, proper values or spectral values. In this book we shall use only the name characteristic value .

Remark 6.1. If T is any linear operator and c is any scalar, the set of vectors such that $T(\ )c$ is a subspace of V.

1. It is the null space of the linear transformation $(T\ \ cI)$.

2. Let the subspace $T(\ ) = c$ 6= 0

   i:e:; $(T\ \ cI)(\ )$ 6= 0( ) where 6= (

   i:e:; $(T\ \ cI)(\ ) = 0$ where 6= 0:

) $(T\ \ cI)$ is not 1 -1 :

3. If the underlying space V be finite dimensional, $(T - cI)$ fails to be $1 - 1$ if $\det(T - cI) \neq 0$:

Theorem 6.1. Let T be a linear operator on a finite-dimensional space V and let c be a scalar. The following are equivalent.

1. c is a characteristic value of T.

2. The operator $(T - cI)$ is singular (not invertible).

3. $\det(T - cI) = 0$:

Proof. (i) $\Rightarrow$ (ii) :

Assume that c is a characteristic value of T.

$\Rightarrow$ The ooperator $T - cI$ is not $1 - 1$.

$\Rightarrow$ $T - cI$ is singular (or not invertible).

$\Rightarrow$ (i) $\Rightarrow$ (ii).

(ii) $\Rightarrow$ (iii) :

Assume that the operator $T - cI$ is singular or invertible.

$\Rightarrow$ The null space of $T - cI = \{0\}$:

$\Rightarrow$ $\{\alpha \mid T\alpha = c\alpha\} = \{0\}$:

$\Rightarrow$ $\{\alpha \mid (T - cI)\alpha\} = \{0\}$:

$\Rightarrow$ $\det(T - cI) = 0$:

(iii) $\Rightarrow$ (i) :

Assume that $\det(T - cI) = 0$.

Note that the expansion of det(T cI) will be a polynomial of degree n in the variable c.

The characteristic values are nothing but the roots of this polynomial.

$\Rightarrow$ 'c' is a characteristic value of T.

$\Rightarrow$ (iii) $\Rightarrow$ (i).

This completes the proof of the theorem.

Note 6.2. If $\mathcal{B}$ is any ordered basis for V and if $[T]_{\mathcal{B}} = A$; then $T - cI$ is invertible if and only if the matrix $A - cI$ is invertible. Accordinly we make the following definition.

Definition 6.2. If A is an $n \quad n$ matix over the field F, a characteristic value of A in F is a scalar c in F such that the matrix $(A \ cI)$ is singular (not invertible).

Remark 6.2. Let c be a characteristic value of A.

) $\det(A \ cI) = 0$

) $\det(cI \ A) = 0$.

Note that, if $f = \det(xI \ A)$

Then $\det(cI \ A) = 0$) $f(c) = 0$

) The characteristic value of A in F are nothing but the scalars c in F for which $f(c) = 0$.

Hence $f = \det(xI \ A)$ is called the characteristic polynomial of the matrix A.

Note that $f = \det(xI \ A)$ is a monic polynomial of degree n.

Lemma 6.1. Similar matrices have the same characteristic polynomial.

Proof. Assume that the two matrices A and B are similar.

Then by definition, $B = P^{1}AP$.

Now our aim is to prove that the characteristic polynomial of A and B are same.

ie:; to prove that $\det(xI A) = \det(xI B)$.

Consider

$$\begin{aligned}
\det(xI \quad B) &= \det(xI \quad P^{1}AP) \\
&= \det(P^{1}xIP \quad P^{1}AP) \\
&= \det[(P^{1}xI \quad P^{1}A)P] \\
&= \det[P^{1}(xI \quad A)P] \\
&= \det p^{1} \det(xI \quad A)\det(P) \\
&= \det(xI \quad A)
\end{aligned}$$

Note 6.3. This lemma enables us to define the characteristic polynomial of the operator T as the characteristic polynomial of any $n \quad n$ matrix, which represent T is same ordered basis of V.

Just in the case of matrices, the characteristic values of T will be the roots of the characteristic polynomial for T.

$\Rightarrow$ T cannot have more than n distinct characteristic values.

It is important to point out that T may not have any characteristic values.

Example 6.1. Let T be a linear operator on R^2 which is represented in the standard ordered basis by the matrix

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

The characteristic polynomial for T (or for A) is

$$\det(xI - A) = \begin{vmatrix} x & 1 \\ -1 & x \end{vmatrix} = x^2 + 1$$

$\Rightarrow$ $\det(xI - A) = 0 \Rightarrow x^2 + 1 = 0 \Rightarrow x = \pm i$ which are not real.

Thus, the operator T has no characteristic values.

However, if U is any linear operator on C^2 which is represented by

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

then U has two characteristic value i and $-i$.

$\Rightarrow$ In discussing the characteristic values of a matrix A, we must specify the field involved. The matrix A above has no characteristic value in R, but has the two characteristic value i and $-i$ in C.

Example 6.2. Let A be the real 3×3.

$$A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}$$

Find the characteristic values and characteristic roots associated with the characteristic values.

Solution. The characteristic polynomial for A is

$$\det(xI \quad A) = \begin{matrix} x \quad 3 & 1 & 1 \\ 2 & x \quad 2 & 1 \\ 2 & 2 & x \end{matrix}$$

$$= x^3 \quad 5x^2 + 8x \quad 4$$

$$= (x \quad 1)(x \quad 2)^2$$

Thus, the characteristic values of A are 1 and 2.

Let T be a linear operator on R^3 which is represented by the above matrix A, in the ordered basis.

Next, we shall find the characteristic vectors associated with the characteristic values 1 and 2.

If the characteristic value is 1:

$$A \quad I = \begin{bmatrix} 3 & 1 & 1 \\ 2 & 2 & 1 \\ 2 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

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

Here $\det(A \quad I) = 0$ which implies that rank of $(A \quad I) \neq 3$:

Consider any $2 \; 2$ minor of $A \; I$ and find its determinent.

For instance, $\begin{matrix} 2 & 1 \\ 2 & 1 \end{matrix} = 0$ but For instance, $\begin{matrix} 2 & 1 \\ 2 & 2 \end{matrix} = 2 \quad 0$

) rank of $(A \quad I) = 2$ and hence $T \quad I$ has nullity equal to 1.

So the space of characteristic vectors associated with the characteristic value is 1-dimensional.

The vector $_1 = (1; 0; 2)$ span the null space of $T \quad I$.

) $T(\;) = \;$ if and only if is a scalar multiple of $_1$.

If the characteristic value is 2:

$$\text{Consider } A \quad 2I = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 2 & 2 & 2 \end{bmatrix}$$

Clearly $A \quad 2I$ has rank 2.

$\Rightarrow T - 2I$ has nullity $= 1$.

Thus, the space of characteristic vectors associated with the characteristic value 2 has dimension $= 1$.

$\Rightarrow T(\alpha) = 2\alpha$ if and only if α is a scalar multiple of $\alpha_2 = (1, 1, 2)$.

Definition 6.3. Let T be a linear operator on a finite-dimensional space V. We say that T is diagonalizable if there is a basis for V each vector of which is a characteristic vector of T.

Note 6.4. The reason for the name "Diagonalizable".

If there is an ordered basis $\mathcal{B} = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ for V, in which each α_i is a charcteristic vector of T, then the matrix of T, in the ordered basis $\mathcal{B}$ is a Diagonal matrix.

i.e., If $T(\alpha_i) = c_i\alpha_i$, then the matrix of T in the ordered basis $\mathcal{B}$ is

$$[T]_{\mathcal{B}} = \begin{bmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & 0 & \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & c_n \end{bmatrix}$$

Note 6.5. 1. The scalars $c_1, c_2, \ldots, c_n$ need not be distinct.

2. Infact, they all will be same, if T is a scalar multiple of the identity operator (or) $T = cI$.

3. T is also diagonalizable, when the characteristic vectors of T span V.

4. In Example 6.1, we have a lienar operator on R^2 whhich is not diagonlizable, because it has no charcteristic value.

   In Example 6.2, the operator T has characteristic values. In fact the characteristic polynomial is $f = (x-1)(x-2)^2$. But still T fails to be diagonalizable.

5. Suppose that T is diagonlizable linear operator. Let $c_1, c_2, \ldots, c_k$ be the distinct characteristic values of T.

   Assume that c_1 is repeated d_1 times.

   c_2 is repeated d_2 times.

   $\vdots$

   c_n is repeated d_n times.

Then there exists an ordered basis $\mathcal{B}$ in which the matrix of T is represented by a diagonal matrix, whose diagonal entries are the scalars c_i; which are each repeated d_i times. In fact, in this way the matrix of T has the Block form

$$[T]_{\mathcal{B}} = \begin{bmatrix} c_1 I_1 & 0 & 0 \\ 0 & c_2 I_2 & 0 \\ \cdot & \cdot & \cdot \\ 0 & 0 & c_n I_n \end{bmatrix}$$

where I_j os the $d_j \quad d_j$ identity matrix.

6. The number d_i is equal to the number of times the scalar c_i is repeated, as a root of f is equal to the dimension of the space of characteristic vectors associated with the characteristic value c_i: This is because the nullity of the diagonal matrix is equal to the number of zeros which it has on its main diagonal.

Lemma 6.2. Suppose that $T \quad = c \quad$. If f is any polynomial, then $f(T) \quad = f(c)$

Lemma 6.3. Let T be a linear operator on the finite-dimensional space V. Let $c_1; c_2; \quad ; c_k$ be the distinct characterist values of T and let W_i be the space of characteristic vectors associated with the characteristic values c_i. If $W = W_1 + W_2 + \quad + W_k$, then

$$\dim W = \dim W_1 + \dim W_2 + \quad + \dim W_k$$

In fact, if $\mathcal{B}_i$ is an ordered basis for W_i, then $\mathcal{B} \quad \{\mathcal{B}_1; \quad ; \mathcal{B}_k\}$ is an ordered basis for W.

Proof. Given that W_1 is the space of characteristic vectors associated with the characteristic value c_1 etc.,

Similarly, W_k is the space of characteristic vectors associated with the characteristic value c_k.

) The space $W = W_1 + W_2 +$ cdots $+ W_k$ is the subspace spanned by all the characteristic vectors of T.

Note that when $W = W_1 + W_2 +$ cdots $+ W_k$; then we expect that

$$\dim W < \dim W_1 + \quad + \dim W_k;$$

because of linear relations which may exist between vectors in the various spaces.

This lemma states that the characteristic space associated with different characteristic values are independent of one another.

$$\begin{aligned}
\text{Let } {}_i\ 2\ W_i &= \text{The space of characteristic vectors associated with the characteristic}\\
&\ \text{value } c_i \quad (i = 1; 2; \quad ; k)\\
)\quad {}_1\ 2\ W_1 \quad) &\quad T\ {}_1 = c_1\ {}_1\\
&:\\
{}_k\ 2\ W_k \quad) &\quad T\ {}_k = c_1\ {}_k
\end{aligned}$$

Suppose that (for each i) we have a vector ${}_i$ in W_i, and assume that ${}_1 + {}_2 +$ $+ {}_n = 0$. Now, we shall prove that ${}_i = 0$ for each i.

Let f be any polynomial.

$$\begin{aligned}
T\ {}_i &= c_i\ {}_i\\
)\quad f(T)\ {}_i &= f(c_i)\ {}_i\\
\text{i:e:;}\quad f(T)\ {}_1 + \quad + f(T)\ {}_k &= f(c_1)\ {}_1 + \quad + f(c_k)\ {}_k
\end{aligned}$$

Since $T\ {}_i = c_i\ {}_i$, then by above lemma we have

$$\begin{aligned}
f(T)0 &= f(c_i)0\\
0 &= f(T)0\\
&= f(T)({}_1 + {}_2 + \quad + {}_k)\\
&= f(T)\ {}_1 + f(T)\ {}_2 + \quad + f(T)\ {}_k\\
&= f(c_1)\ {}_1 + \quad + f(c_k)\ {}_k
\end{aligned}$$

Choose polynomials $f_1; f_2; \quad ; f_k$ such that

$$f_i(c_j) = {}_{ij} = \begin{cases} 1; & \text{if } i = j\\ 0; & \text{if } i\ 6= \end{cases} \tag{6.1}$$

Then

$$\begin{aligned}
0 &= f_i(T)0\\
&= f_i(c_1)\ {}_1 + c_i(c_2)\ {}_2 + \quad + f_i(c_k)\ {}_k\\
&= {}_{i1}\ {}_1 + {}_{i2}\ {}_2 + \quad + {}_{ik}\ {}_k\\
&= \sum_j {}_{ij}\ {}_j\\
&= {}_{ii}\ {}_i = {}_i
\end{aligned}$$

Now, let $_i$ be an ordered basis for W_i and $B = (B_1; B_2; \quad ; B_k)$:

i:e:; B spans the subspace $W = W_1 + W_2 + \quad + W_k$.

Now, we shall prove that B is a linearly independent sequence of vectors.

Note that any linear relation between the vectors in B has the form

$$_1 + \quad + {}_k = 0$$

where every $_i$ is some linear combinations of the vectors in the respective ordered basis B_i .

Since $_i = 0$ for each i .

i:e:; $_1 + {}_2 + \quad + {}_k = 0$) each $_i = 0$.

Thus, each $_i$ is linearly independent.

Therefore, there exists only the trivial relation between the vectors in B .

Thus, $B = (B_1; B_2; \quad ; B_k)$ is an ordered basis for V .

This completes the proof of the lemma.

Theorem 6.2. Let T be a linear operator on a nite-dimensional space V . Let $c_1; c_2; \quad ; c_k$ be the distinct characteristic values of T and let W_i be the null space of $(T \quad c_iI)$. The following are equivalent.

(i) T is diagonalizable.

(ii) The characteristic polynomial for T is

$$f = (x \quad c_1)^{d_1} \quad (x \quad c_k)^{d_k}$$

and $\dim W_i = d_i; \ i = q; w; \quad ; k$.

(iii) $\dim W_1 + \dim W_2 + \quad + \dim W_k = \dim V$:

Proof. (i)) (ii) :

Given that T is diagonlizable.

Let $c_1; c_2; \quad :c_k$ be the distinct characteristic value of T .

Then we know that there exists an ordered basis B in which T is represetned by a diagonal matrix whose diagonal entries are the scalars c_i which are respectively repeated d_i times. Then the matrix has the Block form

$$[T]_B = \begin{bmatrix} c_1I_1 & 0 & 0 & 0 \\ 0 & c_2I_2 & & \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & c_nI_n \end{bmatrix} \qquad (6.2)$$

where I_j is an identity matrix of order $d_j\ d_j$. This implies that the characteristic polynomial is as follows:

If the above $[T]_B$ is matrix A , then the characteristic polynomial of A isdet $(xI\ A)$.

$$\text{i:e:;}\quad f(x) = \begin{bmatrix} x & 0 & & 0 \\ 0 & x & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & x \end{bmatrix} \begin{bmatrix} c_1I_1 & 0 & 0 & 0 \\ 0 & c_2I_2 & & \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & c_nI_n \end{bmatrix} \qquad (6.3)$$

where c_i is repeated d_i times each c_jI_j is a Block.

This implies that $f = (x\quad c_1)^{d_1}\quad (x\quad c_k)^{d_k}$.

Also, (6.2)) d_i is equal to the number of times which c_i is repeated as a root of f .

i:e:; d_i is equal to the dimension of the space W_i of characteristic vectors associated with the characteristic values c_i: $(i = 1; 2;\quad ; k)$

Hence (i)) (ii) .

(ii)) (iii) :

Given that the characteristic polynomial for T is

$$f = (x\quad c_1)^{d_1}\quad (x\quad c_k)^{d_k} \qquad (6.4)$$

and $\dim W_i = d_i$.

Note that the degree of the characteristic polynomial f is $d_1 + d_2 + \quad + d_k$.

and also $\dim W_1 + \dim W_2 + \quad + \dim W_k = \dim V$.

Hence (ii)) (iii)

(iii)) (i) :

Given that $\dim W_1 + \dim W_2 + \quad + \dim W_k = \dim V$.

This is possible only when $V = W_1 + W_2 + \quad + W_k$.

i:e:; The charcteristic vectors T span V .

i:e:; T is diagonalizable.

Hence (iii) $\Rightarrow$ (i) .

This completes the proof of the theorem.

Example 6.3. Let T be the linear operator on R^3 which is represented in the standard ordered basis by the matrix

$$A = \begin{bmatrix} 5 & 6 & 6 \\ 1 & 4 & 2 \\ 3 & 6 & 4 \end{bmatrix}$$

Solution. Let us first find the characteristic polynomial for A .

The characteristic polynomial of

$$\begin{aligned} A &= \det(xI \ A) \\ &= \begin{bmatrix} x \ 5 & 6 & 6 \\ 1 & x \ 4 & 2 \\ 3 & 6 & x \ 4 \end{bmatrix} \\ &= (x \ 2)^2(x \ 1) \quad \text{(on expanding the determinant)} \end{aligned}$$

is the characteristic polynomial of A .

Therefore, the characteristic value of A are 1 and 2 .

Now, let us find the dimensions of the spaces of characteristic vectors associated with the characteristic values 1 and 2 .

When $c_1 = 1$:

$$\begin{aligned} A \ c_1I &= A \ 1I \\ &= A \ I \\ &= \begin{bmatrix} 4 & 6 & 6 \\ 1 & 3 & 2 \\ 3 & 6 & 5 \end{bmatrix} \end{aligned}$$

Now,

$$\det(A \ I) = 0 \Rightarrow \text{rank of } A \ I \neq 3: \tag{6.5}$$

Consider any 2 2 matrix of A I and find its determinant.

For instance, $\begin{vmatrix} 4 & \\ 1 & 3 \end{vmatrix} \neq 0$:

$\Rightarrow$ rank of $A \quad I = 2$:

If W_1 is the space of characteristic vector associated with the characteristic value 1, then we know that $\dim W_1 = 1$ ($\because$ $c_1 = 1$ is repeated only once).

If W_2 is the space of characteristic vector associated with the characteristic value 2, then we know that $\dim W_1 = 2$ ($\because$ $c_2 = 2$ is repeated twice).

When $c_2 = 2$;

$$
\begin{aligned}
A \quad c_2 I &= A \quad 2I \\
&= \begin{bmatrix} 3 & 6 & 6 \\ 1 & 2 & 2 \\ 3 & 6 & 6 \end{bmatrix}
\end{aligned}
$$

Note that $\det(A \quad 2I) = 0 \Rightarrow$ rank of $(A \quad 2I) \neq 3$:

Also, note that the determinant of all $2 \quad 2$ minors are zero.

$\Rightarrow$ rank of $(A \quad 2I) = 1$:

Here $\dim W_1 = 1$; $\dim W_2 = 2$ and $\dim V = 3$:

$\Rightarrow$ $\dim V = \dim W_1 + \dim W_2$:

Hence by theorem, T is diagonalizable.

The null space of $(T \quad I)$ is spanned by the vecotrs ${}_1 = (3; \quad 1; 3)$ and so $\{ {}_1 \}$ is a basis for W_1:

The null space of $(T \; 2I)$ (i:e:; the space W_2) consists of the vectors $(x_1; x_2; x_3)$ with $x_1 = 2x_2 + 2x_3$:

Thus, one example of a basis for W_2 is

$$
\begin{aligned}
{}_2 &= (2; 1; 0) \\
{}_3 &= (2; 0; 1)
\end{aligned}
$$

If $\mathcal{B} = \{ {}_1; \; {}_2; \; {}_3 \}$ then $[T]_{\mathcal{B}}$ is the diagonal matrix.

$$
D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}
$$

The fact that T is diagonalizable means that the original matrix A is similar to the diagonal matrix D.

The matrix P which enables us to change coordinate from the basis B to the standard basis is the matrix which has the transpose of $_1; _2; _3$ as its column vectors;

$$P = \begin{bmatrix} 3 & 2 & 2 \\ 1 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$$

Furthermore, $AP = PD;$ so that

$$P\ ^{1}AP = D:$$

6.3. Annihilating Polynmials

Suppose T is a linear operator on a vector space V, over a field F. Let p be a polynomial over F. Then $p(T)$ is again a linear operator on V.

Let q be any other polynomial over F. Then

$$\begin{aligned} (p+q)(T) &= p(T)+q(T) \\ (pq)(T) &= p(T)q(T) \end{aligned}$$

We say that the polynomial p annihilates the operator T if $p(T) = 0:$

Thus, the collection of polynomials p which anniilate T is an ideal in the polynomial algebra $F[x]$.

It may be that T is not annihilated by any non-zero polynomial. But, that cannot happen, if the space V is finite-dimensional.

Let $\dim V = n$. Suppose T is a linear operator on V. Note that $I; T; T^2; \quad ; T^{n^2}$ is a sequence of n^2+1 operators in $L(V;V)$ - the space of linear operators on V.

We know that $\dim L(V;V) = \dim_F V \ \dim_F V = n \ \ n = n^2:$

) The maximal linearly independent set in $L(V;V)$ contains n^2 elements.

) The above n^2+1 elements must be linearly dependent.

i:e:; $\exists$ scalars $c_0; c_1;$ $; c_{n^2}$ not all zero such that

$$c_0 I + c_1 T + \quad + c_{n^2} T^{n^2} = 0$$

for some scalars c_i not all zero. So the ideal of polynomials which annihilate T contains a non-zero polynomial of degree n^2 or less.

According to Taylor formula, every polynomial ideal consists of all multiples of some fixed monic polynomial, the generator of the ideal.

Thus, there corresponds to the operator T a monic polynomial p with this property:

If f is a polynomial over F, then $f(T) = 0$ if and only if $f = pg$, where g is some polynomial over F.

Definition 6.4. Let T be a linear operator on a finite-dimensional vector space V over a field F. The minimal polynomial for T is the (unique) monic generator of the ideal of polynomials over F which annihilates T.

Note 6.6. The name minimal polynomial stems from the fact that the generator of a polynomial ideal is characterized by being the monic polynomial of minimum degree in the ideal. That means that the minimal polynomial p for the linear operator T is uniquely determined by these three properties:

1. p is a monic polynomial over the scalar field F.
2. $p(T) = 0$:
3. No polynomial over F which annihilates T has smaller degree than p.

Facts About Minimal Polynomials:

1. If A is an n n matrix over F, we define the minimal polynomial of A, in an analogous way, as the unique monic generator of the ideal of all polynomials over F, which annihilate A.

   If the operator T is represented, in some ordered basis, by the matrix A, then both T and A have the same minimal polynomial. That is because $f(T)$ is represented in the basis by the matrix $f(A)$ $\Rightarrow$ $f(T) = 0$ if and only if $f(A) = 0$.

2. Since $f\ P^{1}AP = P^{1} f(A)P;$ it follows that any two similar matrices have the same minimal polynomial.

3. Suppose A is an n n matrix with entries in the field F. Let F_1 be a field which contains F as a subfield.

   For example:

   (1) F is a field of rational numbers and F_1 is a field of real numbers and

   (2) F a field of real numbers and F_1 a field fof complex numbers.

) We may regard A as an n n matrix over either F or F_1.

) It may appear that, there will be two different minimal polynomial for A. But the fact is both the minimal polynomials must be the same.

4. We have observed that, if dim $V = n$ and T is any linear operator on V, then the degree of the minimal polynomial of T does not exceed n2. The fact, however, is that it cannot exceed n.

5. We shall see shortly, that every operator is annihilated by its own characteristic polynomial.

Theorem 6.3. Let T be a linear operator on an n-dimensional vector space (or let A be an n n matrix). The characteristic and minimal polynomials for T [for A] have the same roots, except for multiplicities.

Proof. Let p be the minimal polynomial for T.

) p is a monic polynomial over F.

$p(T) = 0$ and no polynomial over F, which annihilates T, has smaller degree than that of p.

Let c be a scalar. Now, our aim is to prove that p(c) = 0 if and only if c is a characteristic value of T.

First assume that $p(c) = 0$. i:e:; if c is a root of the minimal polynomial for T, then c is also the root of the characteristic polynomial of T and vice versea.

i:e:; to prove that $f(c) = 0$ if and only if c is a characteristic value of T.

Necessary Part: Let $p(c) = 0$:

) c is a root of the polynomial p.

) x c is a factor of the polynomial p.

) 9 some polynomial say q such that $p = (x\ c)q$

Thus deg p = deg of $(x\ c)$ + deg of q

$\Rightarrow$ $\deg q < \deg p$.

$\Rightarrow$ $q(T) \neq 0$.

$\Rightarrow$ $\exists$ a vector such that $q(T) \neq 0$.

$\Rightarrow$ if $= q(T)$ then 0.

Also, $p(T) = 0$.

$$\begin{aligned} \Rightarrow p(T) &= 0 \\ 0 &= (T \quad cI)q(T) \\ &= (T \quad cI) \\ &= T \quad c \\ \Rightarrow T &= c \quad \text{where} \quad \neq (\end{aligned}$$

$\Rightarrow$ c is a characteristic value of T.

Sufficient Part: Assume that c is a characteristic value of T.

$\Rightarrow$ $\exists \quad \neq 0$ such that $T \quad = c$.

Hence by theorem, we have $p(T) \quad = p(c)$.

$\Rightarrow$ $p(T) = p(c)$.

But, $p(T) = 0$.

Hence $p(c) = 0$.

Thus, c is a root of the minimal polynomimal p.

This completes the proof of the theorem.

Example 6.4. If T is a diagonalizable linear operator, then the minimal polynomial for T is a product of distinct linear factors.

Solution. Let T be a diagonisable operator.

Let $c_1, c_2, \quad, c_k$ be the distinct characteristic value of T.

Then the minimal polynomial for T is the polynomial

$$p = (x \quad c_1)(x \quad c_2) \quad (x \quad c_k)$$

If is a characteristic vector of T, then one of the operators

$T \quad c_1I, T \quad c_2I, \quad, T \quad c_kI$ sends into 0.

$\Rightarrow$ $(T \quad c_1I)(T \quad c_2I) \quad (T \quad c_kI) = 0$, for every characteristic vector .

) $p(T) = (T \quad c_1I)(T \quad c_2I) \quad (T \quad c_kI) = 0$

Example 6.5. Let us try to find the minimal polynomials for the operators in Examples 6:1; 6:2 and 6:3 . We shall discuss them in reverse order.

The operator in Example 6:3 was found to be diagaonlizable with characteristic polynomial

$$f = (x \quad 1)(x \quad 2)^2$$

) By the previous example, we see that minimal polynomial for T is $p = (x\ 1)(x\ 2)$.

$$\text{Compute } (A \quad I)(A \quad 2I) = \begin{bmatrix} 4 & 6 & 6 \\ 1 & 3 & 2 \\ 3 & 6 & 5 \end{bmatrix} \begin{bmatrix} 3 & 6 & 6 \\ 1 & 2 & 2 \\ 3 & 6 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

i:e:; $(A \quad I)(A \quad 2I) = 0$.

In Example 6:2 , the operator T also had the characteristic polynomial

$$f = (x \quad 1)(x \quad 2)^2$$

But, this T is not diagonalizable.

) We cannot conclude that the minimal polynomial of T is $(x\ 1)(x\ 2)$:

Then, what do we know about the minimal polynomial?

Here $x = 1$ (with multiplicity 1) and $x = 2$ (with multiplicity 2) are the roots of characteristic polynomial of A .

) The minimal polynomial for T will be of the form

$$(x \quad 1)^k(x \quad 2)^l \quad (k \quad 1; \quad l \quad 1) \tag{6.6}$$

Now, our aim is to find integers k and l in such a way that (6.6) becomes a minimal polynomial for T .

(a) Let us try $(x \quad 1)(x \quad 2)$:

$$\text{Consider } (A \quad I)(A \quad 2I) = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 2 & 2 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & 1 \\ 2 & 0 & 1 \\ 4 & 0 & 2 \end{bmatrix}$$

) We can conclude that the degree of the minimal polynomial for T is greater than or equal to 3.

(b) Let us try for $(x \quad 1)^2(x \quad 2)$ or $(x \quad 1)(x \quad 2)^2$.

Note that $(x \ 1)(x \ 2)^2$ is the characteristic polynomial, would seem a less random choice.

$$(A \quad I)(A \quad 2I) = (A \quad I)(A \quad 2I)(A \quad 2I)$$

$$= \begin{bmatrix} 2 & 0 & 1 \\ 2 & 0 & 1 \\ 4 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 2 & 2 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

) The minimal polynomial for T is the characteristic polynomial.

In Example $6:1$, we discussed the linear operator T on R^2 which is represented in the standard basis by the matrix

$$A = \begin{bmatrix} & 1 \\ 1 & 0 \end{bmatrix}$$

Here the characteristic polynomial is $x^2 + 1$, which has no real roots (i and i).

However, to determine the minimal polynomial, we can forget about T and can concentrate on A.

When considered as $2 \ 2$ complex matrix, A has characteristic values $+i$ and i. Both the roots must appear in the minimal polynomial.

i:e:; The minimal polynomial is divisible by $x^2 + 1$:

Let us compute $A^2 + I$:

$$A^2 + I = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\text{i:e:;} \quad A^2 + I = 0$$

Thus, the minimal polynomial is $x^2 + 1$.

Theorem 6.4. (Cayley-Hamilton). Let T be a linear operator on a finite dimensional vector space V. If f is the characteristic polynomial for T, then $f(T) = 0$; in other words, the minimal polynomial divides the characteristic polynomial for T.

Proof. Let K be the commutative ring with identity, consisting of all polynomials in T. In fact, K is actually a commutative algebra with identity over the scalar field.

Let $\{ _1; \ _2; \ ; \ _n\}$ be an ordered basis for V.

Let A be the matrix which represents T in this basis. Then

$$T(_i) = \sum_{j=1}^{n} A_{ji} \ _j \quad (i = 1; 2; \quad ; n)$$

$$\Rightarrow \sum_{j=1}^{n} \ _{ij}T(_j) = \sum_{j=1}^{n} A_{ji} \ _j \quad (1 \quad i \quad n)$$

$$\Rightarrow \sum_{j=1}^{n} \ _{ij}T \quad A_{ji}I \ _j = 0$$

Let B denote the elements of $K^{n \ n}$ with entries are

$$B_{ij} = \ _{ij}T \quad A_{ji}I \tag{6.7}$$

When $n = 2$: Then

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

$$= \begin{bmatrix} _{11}T \quad A_{11}I & _{12}T \quad A_{21}I \\ _{21}T \quad A_{11}I & _{22}T \quad A_{11}I \end{bmatrix}$$

$$\Rightarrow B = \begin{bmatrix} T \quad A_{11}I & A_{21}I \\ A_{11}I & T \quad A_{11}I \end{bmatrix}$$

$$
\begin{aligned}
\Rightarrow \det B &= (T \quad A_{11}I)(T \quad A_{22}I) \quad A_{12}A_{21}I \\
&= T^2 \quad (A_{11}+A_{22})T + (A_{11}A_{22} \quad A_{12}A_{21})I \\
\Rightarrow \det B &= f(T)
\end{aligned}
$$

where $f(T)$ is the characteristic polynomial:

$$f = x^2 \quad (\text{trace } A)x + \det A$$

For the case $n > 2$; it is also clear that

$$\det B = f(T)$$

since f is the determinant of the matrix $xI \quad A$ whose entries are the polynomials

$$(xI \quad A)_{ij} = {}_{ij}\,x\,A_{ji}$$

Now, our wish is to prove that $f(T) = 0$.

i:e:; to prove that $f(T)$ is a zero operator.

i:e:; to prove that $(\det B\,{}_k = 0; \quad k = 1; 2; \quad ; n$.

By the de nition of B, the vectors $\{\ _1; \ _2; \quad ; \ _n\}$ satisfy the equations.

$$\sum_{j=1}^{n} B_{ij}\ _j = 0 \quad (1 \quad i \quad n) \tag{6.8}$$

When $n = 2$, it is suggestive to write the equation (6.8) in the form

$$\begin{bmatrix} T \quad A_{11}I & A_{21}I \\ A_{12}I & T \quad A_{22}I \end{bmatrix} \begin{bmatrix} \ _1 \\ \ _2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{6.9}$$

In this case, the classical adjoint, $\operatorname{adj} B$ is the matrix

$$B = \begin{bmatrix} T \quad A_{22}I & A_{21}I \\ A_{12}I & T \quad A_{11}I \end{bmatrix}$$

and

$$
\begin{aligned}
BB &= \begin{bmatrix} \det B & 0 \\ 0 & \det B \end{bmatrix} \\
&= \det B \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \det B \quad I
\end{aligned}
$$

$$\Rightarrow \quad BB \begin{bmatrix} \ _1 \\ \ _2 \end{bmatrix} = (\det B) \begin{bmatrix} \ _1 \\ \ _2 \end{bmatrix}$$

$$\text{i:e:;}\quad (\det B)\begin{bmatrix}\epsilon_1\\ \epsilon_2\end{bmatrix} = \tilde{B}B\begin{bmatrix}\epsilon_1\\ \epsilon_2\end{bmatrix} = \tilde{B}\left(B\begin{bmatrix}\epsilon_1\\ \epsilon_2\end{bmatrix}\right) \quad (*\ (6.9))$$

$$\Rightarrow (\det B)\begin{bmatrix}\epsilon_1\\ \epsilon_2\end{bmatrix} = \tilde{B}\begin{bmatrix}0\\ 0\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$$

$$\Rightarrow (\det B)\epsilon_k = 0$$

In the general case, let $\tilde{B} = \text{adj}\,B$. Then by (6.8), we have

$$\sum_{j=1}^{n} \tilde{B}_{ki}B_{ij}\alpha_j = 0$$

For each pair k, i, and summing on i, we have

$$\begin{aligned} 0 &= \sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{B}_{ki}B_{ij}\alpha_j \\ &= \sum_{j=1}^{n}\left(\sum_{i=1}^{n}\tilde{B}_{ki}B_{ij}\right)\alpha_j \\ &= \sum_{j=1}^{n}\delta_{kj}(\det B)\alpha_j \\ &= (\det B)\sum_{j=1}^{n}\delta_{kj}\alpha_j = (\det B)\delta_{kk}\alpha_k \\ &= (\det B)\alpha_k \quad 1 \le k \le n \end{aligned}$$

Thus,we proved the theorem for all the cases.

Hence the theorem.

Note 6.7. The Cayley-Hamilton theorem is useful to us at this point primarily because it narrows down the search for the minimial polynomials of various operators. If we know the matrix A which represents T in some ordered basis, then we can compute the characteristic polynomial f.

Let us Sum Up:

In this unit, the students acquired knowledge to

find the value of the characteristic values and characteristic vectors.

find the diagonalizaton of the matrices.

Check Your Progress:

1. Let A be an n n triangular matrix over the field F. Prove that the characteristic values of A are the diagonal entries of A, i:e; the scalars A_{ii}.

2. Let T be the linear operator on R^3 which is represented in the standard ordered basis by the matrix

$$A = \begin{bmatrix} 9 & 4 & 4 \\ 8 & 3 & 4 \\ 16 & 8 & 7 \end{bmatrix}$$

Prove that T is diagonalizable by exhibiting a basis for $R3$, each vector of which is a characteristic vector of T.

3. Let

$$A = \begin{bmatrix} 6 & 3 & \\ 4 & 1 & \\ 10 & 5 & 2 \; 3 \end{bmatrix}$$

Is A similar over the field R to a diagonal matrix? Is A similar over the field C to a diagonal matrix?

4. Let $a; b$ and c be elements of a field F, and let A be the following 3 3 matrix over F;

$$A = \begin{bmatrix} 0 & 0 & c \\ 1 & 0 & b \\ 0 & 1 & a \end{bmatrix}$$

Prove that the characteristic polynomial for A is x^3 ax^2 bx c and that this is also the minimal polynomial for A.

5. Let A be the 4 4 real matrix.

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

Suggested Readings:

1. M. Artin, Algebra , Prentice Hall of India Pvt. Ltd., 2005.
2. S.H. Friedberg, A.J. Insel and L.E Spence, Linear Algebra , 4th Edition, Prentice-Hall of India Pvt. Ltd., 2009.
3. I.N. Herstein, Topics in Algebra , 2nd Edition, Wiley Eastern Ltd, New Delhi, 2013.
4. J.J. Rotman, Advanced Modern Algebra , 2nd Edition, Graduate Studies in Mathematics, Vol. 114, AMS, Providence, Rhode Island, 2010.
5. G. Strang, Introduction to Linear Algebra , 2nd Edition, Prentice Hall of India Pvt. Ltd, 2013.

BLOCK - IV

Unit – 7: Elementary Canonical Forms-II

Unit – 8: Decompositions

Block-IV

UNIT-7

ELEMENTARY CANONICAL FORMS-II

Structure

Overview

In this unit, we shall introduce a few concepts which are useful in attempting to analyze a linear operator.

Objectives

After successful completion of this lesson, students will be able to

understand the concept of invariant subspaces.

understand the concept of simultaneous triangulation.

7.1. Invariant Subspaces

Definition 7.1. Let V be a vector space and T a linear operator on V. If W is a subspace of V, we say that W is invariant under T if for each vector in W the vector T( ) is in W; i:e:; if T(W) is contained in W.

Example 7.1. Let T be any linear operator on V, then

1. V is invariant under T.
2. The zero subspace of V, f0g is invariant under T.
3. The range of T is invariant under T.
4. The nulls space of T is invariant under T.

Example 7.2. Let F be a field and let D be the differentiation operator on the space F[x] of polynomials over F. Let n be a positive integer and let W be the subspace of polynomials of degree not greater than n. Then W is invariant under D.

Note 7.1. Simply we can say that D is `degree decreasing'.

Example 7.3. A very useful generalization of Example 7.1.

Let T be a linear operator on V Let U be any other linear operator on V, which commutes with T (i:e:;) TU = UT.

Let W be the range of the linear operator U.

Let N be the null space of the linear operator U.

Now, we shall prove that both W and N are invariant under T.

For if, let 2 The range of U.

$$)\qquad = \ U$$

$$)\quad T(\ ) \ = \ T(U(\ )) = (TU)(\ ) = U(T\ \ )$$

$$\text{(or)}\ \ T(\ ) \ = \ U(T\ )\ \text{where}\ T\ 2\ \ V\text{(or)}$$

$$T(\ ) \qquad 2 \quad \text{Range of}\ U$$

) Range of $U = W$ is invariant under T .

Let 2 N) 2 The null space of U which implies that $U \ \ = 0$:

) $U(T(\ )) = (UT)\ = (TU)\ = T(U(\ \ )) = T(0) = 0$:

i:e:; $U(T(\ )) = 0$

i:e:; $T(\ )$ 2 The null space of U .

i:e:; $T(\ )$ 2 N

) The null space of $U = N$ is invariant under T .

Note 7.2. When the subspace W is invariant under the operator T , then T induces a linear operator T_W on the space W . The linear operator T_W is defined by $_W(\ ) = T(\ )$ for in W , but T_W is quite different object from T since its domain is W not V .

Note 7.3. Let V be a finite-dimensional. Then the invariance of W under T has the following matrix representation:

Let $\mathsf{B} = \{\ _1;\ _2;\ \ ;\ _n\}$ be an ordered basis for V and $\mathsf{B}^0 = \{\ _1;\ _2;\ \ ;\ _r\}$ is an ordered basis for W $(r = \dim W)$. Let A be the matrix, whhich represents the transformation T in the basis B .

i:e:; $A = [T]_{\mathsf{B}}$.

In this case, we know that,

$$T(\ _j) \ = \ \sum_{i=1}^{n} A_{ij\ \ i}$$

$$)\quad T(\ _j) \ = \ 0 \ \text{ if } j \ \ r \text{ and } i > r$$

Schematically, if A has the block form

$$A \ = \ \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

Where B is an r r matrix, C is an r (n r) matrix, and D is an (n r) (n r) matrix.

The matrix B is the matrix of the induced operator T_W in the ordered basis $\mathcal{B}^0$.

Lemma 7.1. Let W be an invariant subspace for T. The characateristic polynomial for the restriction operator T_W divides the characteristic polynomial for T. The minimal polynomial for T_W divides the minimal polynomial for T.

Proof. Let $\mathcal{B} = \{\alpha_1; \alpha_2; \ldots; \alpha_n\}$ be an ordered basis for V and $\mathcal{B}^0 = \{\alpha_1; \alpha_2; \ldots; \alpha_r\}$ is an ordered basis for W $(r = \dim W)$. Let A be the matrix, whhich represents the transformation T in the basis $\mathcal{B}$.

i:e:; $A = [T]_{\mathcal{B}}$.

In this case, we know that,

$$T(\alpha_j) = \sum_{i=1}^{n} A_{ij}\alpha_i$$
$$\Rightarrow T(\alpha_j) = 0 \text{ if } j \le r \text{ and } i > r$$

Thus, A has the block form

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

where $A = [T]_{\mathcal{B}}$ and $B = [T_w]_{\mathcal{B}^0}$.

Because of the block form of the matrix

$$\det(xI - A) = \det(xI - B)\det(xI - D)$$

i:e:; The characteristic polynomial for T = (The characteristic polynomial for T_W)det $(xI - D)$.

$$\Rightarrow \frac{\text{The characteristic polynomial for } T}{\text{The characteristic polynomial for} T_W} = \det(xI - D).$$

Thus, the characteristic polynomial for T_W divides the characteristic polynomial for T.

Hence, we proved theorem about characteristic polynomial.

The kth power of the matrix A has the block form

$$A^k = \begin{bmatrix} B^k & C^k \\ 0 & D^k \end{bmatrix}$$

where C^k is some $r \times (n - r)$ matrix.

Thus any polynomial which is satisifed by A, will also be satisfied by B (as well as D)

(or) Any polynomial which annihilates A, also annihilates B as well as D.

Thus the minimial polynomial for B divides the minimal polynomial for A.

i:e:; The minimal polynomial for T_W divides the minimal polynomial for T.

Example 7.4. Let T be any linear operator on a finite-dimensional space V. Let W be the subspace spanned by all the characteristic vectors of T. Let $c_1; c_2; \cdots; c_k$ be the distinct characteristic values of T.

For each i, let W_i be the space of characteristic vector associated with the characteristic value c_i.

Let B_i be an ordered basis for W_i.

Then $\mathsf{B}' = (\mathsf{B}_1; \mathsf{B}_2; \cdots; \mathsf{B}_k)$ is an ordered basis for W and

$$\dim W = \dim W_1 + \cdots + \dim W_k$$

Let us take $\mathsf{B}' = \{\beta_1; \beta_2; \cdots; \beta_r\}$ in which the first few elements β's from the basis B_1, the next few elements β's from the basis B_2 and so on.

Here T is an linear operator on V (i:e:;) $T : V \to V$.

Then $T(\beta_i) = t_i\beta_i \quad (i = 1; 2; \cdots; r)$

where $(t_1; t_2; \cdots; t_r) = (c_1; c_1; \cdots; c_1; c_2; c_2; \cdots; c_2; \cdots; c_k; c_k; \cdots; c_k)$ where each c_i is repeated $\dim W_i$ times.

Now, W is invariant under T (i:e:; $T(W) \subseteq W$) :

$$\begin{aligned}
\text{Let } \beta &\in W \\
\Rightarrow \beta &= x_1\beta_1 + \cdots + x_r\beta_r \\
T(\beta) &= T(x_1\beta_1 + \cdots + x_r\beta_r) \\
&= x_1T(\beta_1) + \cdots + x_rT(\beta_r) \\
&= x_1(t_1\beta_1) + \cdots + x_rt(t_r\beta_r) \\
\Rightarrow T(\beta) &= t_1x_1(\beta) + \cdots + t_rx_r \\
\Rightarrow T(\beta) &= t_1x_1(\beta_1) + \cdots + t_rx_r(\beta_r)
\end{aligned}$$

Here $\{\beta_1; \beta_2; \cdots; \beta_r\}$ is a basis for W.

$\Rightarrow \{\beta_1; \beta_2; \cdots; \beta_r\}$ is a linearly independent set in V.

If $\dim V = n$; this linearly independent set in V can be extended to form a basis of V .

Choose any other vectors $_{r+1}; \; _{r+2}; \quad ; \; _r$ in V such that $\{ \; _1; \; _2; \quad ; \; _n\}$ is a basis for V .

Then we know that, the matrix of T , relative to $\mathcal{B}$ has the block form

$$[T]_{\mathcal{B}} = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

and that the matrix of the restriction operator T_W relative to the basis $\mathcal{B}^0$ is

$$B = \begin{bmatrix} t_1 & 0 & & 0 \\ 0 & t_2 & & 0 \\ . & . & & . \\ 0 & 0 & & t_r \end{bmatrix}$$

The characteristic polynomial of B (i:e:; of T_W) is given by

$$g = (x \quad c_1)^{e_1}(x \quad c_2)^{e_2} \quad (x \quad c_k)^{e_k}$$

where $e_i = \dim W_i$.

) If f is the characteristic polynomial for T , then g divides f .

) The multiplicity of c_i as a root of f is at least $\dim W_i$.

Definition 7.2. Let W be an invariant subspace for T and let be a vector in V . The T-conductor of into W is the set $S_T(\; ; W)$, which consists of all polynomials g (over the scalar field) such that $g(T)$ is in W .

Remark 7.1.

1. Unless specified, the T in the suffix can be dropped, and we can denote by $S(\; ; W)$.

2. In the special case, when the subspace $W = \{0\}$, $S_T(\; ; W)$ is called the T-annihilator of :

Lemma 7.2. If W is an invariant subspace for T , then W is invariant under every polynomial in T . Thus, for each in V , the conductor $S(\; ; W)$ is an ideal in the polynomial algebra $F[x]$:

Proof. Given that W is invariant for T .

) $T(W) \quad W$.

If $\ \in W$, then $T(\ ) \in W$.

$T(\ ) \in W \Rightarrow T(T(\ )) \in W \Rightarrow T^2\ \ \in W$

Proceeding like this, we get $T^k(\ ) \in W;\ \forall k$:

$\Rightarrow\ f(T)\ \ \in W$; for every polynomial f.

Note that, the definition of $S(\ ; W)$ makes sense if W is any arbitrary subset of V.

In fact, if W further happends to be a subspace of V, then $S(\ ; W)$ becomes a subspace of $F[x]$; because

$$(cf + g)(T)\ \ = \ \ cf(T)\ + g(T)$$

Now, if W is invariant under T, let g be any polynomial in $S(\ ; W)$.

$g \in S(\ ; W) \Rightarrow g(T)\ \in W$

Take $\ = g(T)\ $.

If $\ \in W$, then for every polynomial f, $f(T)\ \ \in W$.

$\Rightarrow$ If f is any polynomial, then

$$\begin{aligned} f(T)\ g(T)\ \ &\in W \\ \Rightarrow\ f(T)g(T)\ \ &\in W \\ \Rightarrow\ (fg)(T)\ \ &\in W \\ \Rightarrow\ fg &\in S(\ ; W) \end{aligned}$$

Hence the lemma.

Remark 7.2.

1. The unique monic generator of the ideal $S(\ ; W)$ is also called the T-conductor of $\ $ into W (the T-annihilator in case $W = \{0\}$).
2. The T-conductor of $\ $ into W is the monic polynomial g of least degree such that $g(T)\ \in W$:
3. If g is the T-conductor of $\ $ into W, then an arbitrary polynomial $f \in S(\ ; W)$ if and only if g divides f.
4. The conductor $S(\ ; W)$ always contains the minimal polynomial for T; hence every T-conductor divides the minimal polynomial for T.

Definition 7.3. A linear operator T is said to be triangulable if there is an ordered basis in which T is represented by a triangular matrix.

Lemma 7.3. Let V be a finite-dimensional vector space over the field F . Let T be a linear operator on V such that the minimal polynomial for T is a product of linear factors

$$p = (x \quad c_1)^{r_1} \quad (x \quad c_k)^{r_k}$$

Let W be a proper $(W \neq V)$ subspace of V which is invariant under T . There exists a vector in V such that

(a) is not in W ;

(b) (T cI) is in W , for some characteristic value c of the operator T .

Proof. The condition (a) and (b) say is that T -conductor of into W is a linear polynomial.

Let $\in V$ such that $\notin W$.

Let g be a T -conductor of into W and let p denote the minimal polynomial for T .

Then g divides p , the minimal polynomial for T .

Suppose g is a constant polynomial.

$$\text{Let } g = c_0 \tag{7.1}$$
$$\Rightarrow \quad g(T) = c_0 \tag{7.2}$$
$$\Rightarrow \quad g(T) = c_0 \notin W \tag{7.3}$$

Hence g is not a constant.

$$g = (x \quad c_1)^{e_1} \quad (x \quad c_k)^{e_k}$$

where atleast one of the e_i is positive.

Let one such factor be $(x \quad c_j)^{e_j}$

$\Rightarrow$ $(x \quad c_j)^{e_j}$ divides g .

$\Rightarrow$ $(x \quad c_j)$ dividies g .

$\Rightarrow$ $\frac{g}{(x \quad c_j)} = h:$

$\Rightarrow$ $g = (x \quad c_j)h:$

Since $\notin W \Rightarrow h(T) \notin W$:

If $= h(T)$ then $\notin W$.

i:e:; $\exists$ a vector in V such that $\notin W$.

This proves (a) .

consider $(T \quad c_j I) = (T \quad c_j I)h(T) = g(T) \in W$:

This proves (b):

Theorem 7.1. Let V be a finit-dimensional vector space over the field F and let T be a linear operator on V . Then T is triangulable if and only if the minimal polynomial for T is a product of linear polynomials over F .

Proof. Necessary Part: Assume that the minimal polynomial for T is a product of linear polynomials over F .

i:e:; let $p = (x \quad c_1)^{r_1}(x \quad c_2)^{r_2} \quad (x \quad c_k)^{r_k}$

Thus, the hypothesis lemma 7.3 are satisifed.

i:e:; If $W \neq V$ is a subspace of V is invariant under T , then there exists an $\in V$ such that $\notin W$ and $(T \quad cI) \in W$, for some characteristic value c of the operator T .

By the repeated application of lemma 7.3, we shall arrive at an ordered basis $B = \{\ _1;\ _2; \quad ;\ _n\}$ in which the matrix representing T is upper triangular.

$$[T]_B = \begin{bmatrix} a_{11} & a_{12} & a_{13} & & a_{1n} \\ 0 & a_{22} & a_{23} & & a_{2n} \\ 0 & 0 & a_{33} & & a_{3n} \\ & : & : & & \\ 0 & 0 & 0 & & 0 \end{bmatrix} \tag{7.4}$$

Now (7.4) merely says that

$$T(\ _j) = a_{1j}\ _1 + \quad + a_{jj}\ _j \quad (1 \quad j \quad n) \tag{7.5}$$

i:e:; $T(\ _j) \in$ subspace spanned by $\ _1;\ _2; \quad ;\ _j$.

To Find $\ _1;\ _2; \quad ;\ _n$:

Apply the lemma 7.3 to $W = \{0\}$ and obtain a vector $\ _1$.

Let W_1 is the space spanned by $\ _1$.

Apply the lemma 7.3 to W_1 and obtain $_2$.

Let W_2 is the space spanned by both $_1$ and $_2$.

Apply the above lemma to W_2 and obtain $_3$.

Continue in this way, we find $_1; \ _2; \quad ; \ _n$.

In fact, it is the triangular type relations (7.5), which ensure that (for $j = 1; 2; \quad ; i$) the subspace spanned by $_1; \ _2; \quad ; \ _i$ is invariant under T.

Thus T is triangularable.

Sufficient Part: Let T be triangulable.

) The Characteristic polynomial for T has the form

$$f = (x \quad c_1)^{d_1} \quad (x \quad c_k)^{d_k} \quad \text{where } c_i \ 2 \ F \tag{7.6}$$

Just look at the triangular matrix (7.4).

The diagonal entries $a_{11}; a_{22}; \quad ; a_{nn}$ are the characteristic values, where each c_i is repeated d_i times.

If f can be factorised as in (7.6), this means that the minimal polynomial p can be factorised in the same manner. (* $p = f$).

Thus, the minimal polynomial for T can be factored as a product of linear polynomials over F.

This proves the sufficient part.

Theorem 7.2. Let V be a finite dimensional vector space over the field F and let T be a linear operator on V. Then T is diagonlisable if and only if the minimal polynomial for T has the form

$$p = (x \quad c_1) \quad (x \quad c_k)$$

where $c_1; c_2; \quad ; c_k$ are distinct elements of F.

Proof. Necessary Part: As we already discussed that, if T is diagonalizable, its minimal polynomial is a product of distinct linear factors.

Sufficient Part:Assume that the minimal polynomial for T is a product of distinct linear factors.

Now, we shall prove that T is diagonalizable.

i:e:; To prove that there exists a basis of V, in which each vector is a characteristic vector of T.

Let W be the subspace spanned by all the characteristic vectors of T in such a way that $W\ 6{=}\ V$.

i:e:; There exists a vector $2\ V(\ 2\ W)$ and a characteristic vector c_j of T such that the vector

$$= \quad (T \quad c_j I)$$

lies in W.

$$
\begin{aligned}
\text{Now} \quad & 2 \quad W \\
) \quad & = \quad {}_1 + {}_2 + \quad + {}_k \\
\text{Where } T({}_i) \quad & = \quad c_i\,{}_i \quad (1 \quad i \quad k) \\
\text{and } \ f(T)\,{}_i \quad & = \quad f(c_i)\,{}_i \\
) \quad h(T) \quad & = \quad h(c_1)\,{}_1 + \quad + h(c_k)\,{}_k \ 2\ W
\end{aligned}
$$

for every polynomial h.

From the hypothesis, we can write $p = (x \quad c_i)q$, for some polynomial q.

$$
\begin{aligned}
\text{Also, } q \quad q(c_j) \quad & = \quad (x \quad c_j)h \\
) \quad q(T) \quad q(c_j) \quad & = \quad h(T)(T \quad c_j I) \\
) \quad q(T) \quad q(C_j) \quad & = \quad h(T)(T \quad c_j I) \quad = h(T)\,{}_1 \\
) \quad q(T) \quad Q(c_j) \quad & 2 \quad W
\end{aligned}
$$

Since

$$
\begin{aligned}
p \quad & = \quad (x \quad c_j)q \\
)\ \ p(T) \quad & = \quad (T \quad c_j I)q(T) \\
p(T) \quad & = \quad (T \quad c_j I)q(T) \\
o \quad & = \quad (T \quad c_j I)q(T) \qquad (\text{* } p \text{ is a minimal polynomial for } T) \\
) \quad 0 \quad & = \quad T[q(T)\ ] \quad c_j[q(T)\ ] \\
) \quad T[q(T)\ ] \quad & = \quad c_j[q(T)\ ]
\end{aligned}
$$

$)$ $q(T)$ is a characteristic vector of T.

$)$ $q(T)\ 2\ W$:

Here W is spanned by all characteristic vectors of T (or) any linear combinations of characteristic vectors of $T \in W$:

$\Rightarrow$ $q(c_j) \in W$ where $\notin W$:

$\Rightarrow$ The only possibility is that $q(c_j) = 0$:

Thus, we have $q \quad 0 = (x \quad c_j)h$

$\Rightarrow$ $q = (x \quad c_j)h$.

$\Rightarrow$ $p = (x \quad c_j)[(x \quad c_j)q]$

$\Rightarrow$ $p = (x \ c_j)^2 q$

Thus, the characteristic roots c_j is repeated twice, which is a contradiction to the hypothesis.

This completes the proof of the theorem.

7.2. Simultaneous Triangulation; Simultaneous Diagonalization

Definition 7.4. The subspace W is invariant under (the family of operators) $\mathcal{F}$ if W is invariant under each operator in $\mathcal{F}$.

Lemma 7.4. Let $\mathcal{F}$ be a commuting family of triangulable linear operators on V . Let W be a proper subspace of V which is invariant under $\mathcal{F}$. There exists a vector in V such that

(a) is not in W ;

(b) for each T in $\mathcal{F}$, the vector T is in the subspace spanned by and W .

Proof. Without loss of generality, assume that the family $\mathcal{F}$ contains only a finite number of operators.

$\Rightarrow$ we can find a vector $_1 \notin W$ and a scalar c_1 such that

$$(T_1 \quad c_1 I) \ _1 \in W \tag{7.7}$$

Let V_1 be the collection of all vectors in V such that $(T_1 \quad c_1 I) \in W$.

) V_1 is a subspace of V, which is properly larger than W.

Next we shall prove that V_1 is invariant under F .

Here F is commutting family of triangulable linear operators on V.

) If T commutes with T_1

i:e:; $TT_1 = T_1T = I$

$$\begin{aligned} \text{Consider } (T_1 \quad c_1I)(T\ ) &= (T_1T) \quad c_1I(T\ ) \\ &= (TT_1) \quad Tc_1I \\ &= T(T_1\ ) \quad T(c_1I\ ) \\ &= T(T_1 \quad c_1I) \end{aligned}$$

If $2\ V_1$, then $T(T_1 \quad c_1I)\ 2\ W$

) $T(T_1 \quad c_1I)\ 2\ W$) $(T_1 \quad c_1I)T\ 2\ W$.

) $T\ 2\ V_1\ 8\ 2\ V_1$and $T\ 2$ F .

) $T(V_1) \quad V_1;\ 8T\ 2$ F

Thus, V_1 is invariant under mathscrF .

Now, W is a proper subspace of V. Let $T_2\ 2$ F .

Let U_2 be the linear operator on V_1; which is obtained by restricting T_2 to the subspace V_1.

Thus, the minimal polynomial for U_2 divides the minimal polynomial for T_2.

i:e:; We can find a vector $_2$ in V_1 (not in W) and a scalar c_2 such that $(T_2\ c_2I)\ _2\ 2\ W$.

Thus, we have

(i) $_2$ is not in W.

(ii) $(T_1 \quad c_1I)\ _2\ 2\ W$.

(iii) $(T_2 \quad c_2I)\ _2\ 2\ W$.

Let V_2 be the collection of all vectors in V_1 such that $(T_2 \quad c_2I)\ 2\ W$:

) V_2 is invariant under F .

Let $T_3\ 2$ F .

Let U_3 be the restriction of T_3 to V_2.

) 9 a vector $_3$ in V_2 (not in W) and a scalar c_3 such that $(T_3 \quad c_3I)$ $_3$ 2 W:

If we continuing in this way, we get

) 9 a vector $_r$ in $V_{r\ 1}$ (not in W) and a scalar c_r such that $(T_r \quad c_rI)$ $_r$ 2 W:

In otherwords, 9 $=$ $_r$ (not in W) such that $(T_j \quad c_jI)$ alpha 2 W: $(j = 1; 2; \quad ; r)$

) $T_j \quad c_j$ 2 W (2 W)

) T_j 2 The subspace spanned by and W.

i:e:; T 2 The subspace spanned by and W 8T 2 F where 2 W. (T_j 2 F is arbitarary).

Theorem 7.3. Let V be a nite-dimensional vector space over the eld F. Let F be the commutting family of triangulable linear operator on V. Then there exists an ordered basis for V such that every operator in F is represted by a triangular matrix in that basis.

Proof. Prove the above lemma and prove theorem 7.1 (by replacing T by F)

Theorem 7.4. Let F be a commuting family of diagonalizable linear operators on the nite-dimensional vector space V. There exists an ordered basis for V such that every operator in F is represented in that basis by a diagonal matrix.

Proof. The proof is by induction on $\dim V = n$:

If $n = 1$, the result is quite obvious.

As part of induction, assume that the theorem is true for all vector spaces of dimension less than n.

Now let $\dim V = n$:

Choose any T 2 F , which is not a scalar multiple of the identity operator.

Let $c_1; c_2; \quad ; c_k$ be the distinct characteristic values of T. For every i, let W_i be the null space of $T \quad c_iI$.

If we x any i, then W_i is invariant under every operator that commutes with T.

Let $\mathcal{F}_i$ denote the family of linear operators on W_i wkhich are obtained by restricting the operators $\mathcal{F}$ to the invariant subspace W_i.

Then the minimal polynomial for any operator in $\mathcal{F}_i$ divides the minimal polynomial of the corresponding operator in $\mathcal{F}$.

i:e; Each operator in $\mathcal{F}_i$ is diagonalizable.

Here $\dim W_i < \dim V$.

$\Rightarrow$ The operators in $\mathcal{F}_i$ can be simultaneously diagonalised.

i:e:; W_i has a basis $\mathcal{B}_i$, which consists of vectors and are simulatenously characteristic vectors for every operator in $\mathcal{F}_i$.

Here $T \in \mathcal{F}$, which is a commuting family of diagonalisable linear opoerators on V.

$\Rightarrow$ T is diagonalizable.

$\Rightarrow$ $\mathcal{B} = \{\mathcal{B}_1; \mathcal{B}_2; \quad ; \mathcal{B}_k\}$ is a basis for V.

This basis is the requirement of the theorem.

Let us Sum Up:

In this unit, the students acquired knowledge to

explain the concept of annhiltor.

understand the concept of simulaneous triangulation and diagonalization.

Check Your Progress:

1. Let T be the linear operator on R^2, the matrix of which in the standard basis IS

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

Prove that the only subspaces of R^2 invariant under T are R^2 and the zero subspace.

2. Prove that every matrix A such that $A^2 = A$ is similar to a diagonal matrix.

3. Find an invertible real matrix P such that $P\ ^{1}AP$ and $P\ ^{1}BP$ are both diagonal, where A and B are the real matrices

(a)

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}; \quad B = \begin{bmatrix} 3 & 8 \\ 0 & 1 \end{bmatrix};$$

(b)

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}; \quad B = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix};$$

Suggested Readings:

1. M. Artin, Algebra , Prentice Hall of India Pvt. Ltd., 2005.

2. S.H. Friedberg, A.J. Insel and L.E Spence, Linear Algebra , 4th Edition, Prentice-Hall of India Pvt. Ltd., 2009.

3. I.N. Herstein, Topics in Algebra , 2nd Edition, Wiley Eastern Ltd, New Delhi, 2013.

4. J.J. Rotman, Advanced Modern Algebra , 2nd Edition, Graduate Studies in Mathematics, Vol. 114, AMS, Providence, Rhode Island, 2010.

5. G. Strang, Introduction to Linear Algebra , 2nd Edition, Prentice Hall of India Pvt. Ltd, 2013.

Block-IV

UNIT-8

DECOMPOSITIONS

Overview

In this unit, we shall describe how to decompose the underlying space V into a sum of invariant subspaces for T such that the restriction operators on those subspaces are simple.

Objectives

After successful completion of this lesson, students will be able to

- understand the concept of direct sum and interior direct sum.
- understand the concept of invariant direct sum.

8.1. Direct-Sum Decomposition

Definition 8.1. Let $W_1; W_2; \quad ; W_k$ be subspaces of the vector space V. We say that $W_1; W_2; \quad ; W_k$ are independent if

$$\begin{aligned} {}_1 + {}_2 + \quad + {}_k &= 0 \\ \Rightarrow \text{ each } {}_i &= 0 \end{aligned}$$

where ${}_i \in W_i$ $(i = 1; 2; \quad ; k)$

Note 8.1.

Let $k = 2$: i:e:; W_1 and W_2 are subspaces where ${}_1 \in W_1$ and ${}_2 \in W_2$.

Now, W_1 and W_2 are independent, if

$$\begin{aligned} {}_1 + {}_2 &= 0 \\ \Rightarrow \quad {}_1 &= 0; \ {}_2 = 0 \qquad ({}_1 \in W_1 \text{ and } {}_2 \in W_2) \\ \Rightarrow \quad W_1 \cap W_2 &= \{0\} \end{aligned}$$

Let $k > 2$: Then the independence of $W_1; W_2; \quad ; W_k$.

$\Rightarrow \quad W_1 \cap W_2 \cap \quad \cap W + k = \{0\}$.

In fact, it says that each subspace W_j intersects the sum of all other subspaces W_i only in the zero vector.

The following is the significane of the independence of subspaces:

Let $W = W_1 + W_2 + \quad + W_k$.

(i:e:; W is the subspace spanned by $W_1; W_2; \quad ; W_k$.)

$\Rightarrow$ Each vector in W can be expressed as a sum

$$= {}_1 + {}_2 + \quad + {}_k \quad ({}_i \in W_i) \tag{8.1}$$

If $W_1; W_2;$ $; W_k$ are independent, then the representation of in (8.1) is unique.

If possible, let

$$= \ _1 + \ _2 + \quad + \ _k \text{ where } \ _i \, 2 \, W_i \qquad (8.2)$$

From (8.1) and (8.2), we have

$$_1 + \ _2 + \quad + \ _k = \ _1 + \ _2 + \quad + \ _k$$

$$) \quad (\ _1 \quad _1) + (\ _2 \quad _2) + \quad + (\ _k \quad _k) = 0$$

where $_1 \ _1 \, 2 \, W_1;$ $;$ $_k \ _k \, 2 \, W_k$ and $W_1; W_2;$ $; W_k$ are independent.

$) \quad _1 \ _1 = 0; \ _2 \ _2 = 0; \quad ; \ _k \ _k = 0 .$

$) \quad _i = \ _i \quad 8i$

Hence the representation of in (8.1) is unique.

Thus, when $W_1;$ $W_2;$ $;$ W_k are independent, we can operate with the vectors in W as k-tuples $(\ _1; \ _2; \ ; \ _k);$ $_i$ in W_i, in the same way as we operate with vectors in R^k as k-tupples of numbers.

Theorem 8.1. Let V be a nite-dimensional vector space. Let $W_1; W_2 m$ $; W_k$ be subspaces of V and let $W = W_1 + W_2 + \quad + W_k$. The following are equivalent.

(a) $W_1;$ $; W_k$ are independent.

(b) For each j, $2 \quad j \quad k$, we have

$$W_j \setminus (W_1 + \quad + W_j \ _1) = \text{f}0\text{g}$$

(c) If B_i is an ordered basis for W_i, $q \quad i \quad k$, then the sequence $\mathsf{B} = (\mathsf{B}_1); \mathsf{B}_2;$ $; \mathsf{B}_k)$ is an ordered basis for W.

Proof. (a) $)$ (b) :

Assume that $W_1; W_2;$ $; W_k$ are independent.

Now, our aim is to prove that $W_j \setminus (W_1 + \quad + W_j \ _1) = \text{f}0\text{g}$.

i:e:; to prove that $2 \ W_j \setminus (W_1 + \quad + W_j \ _1)$ then $= 0$.

$$
\begin{aligned}
&\text{let} && 2\,W_j && \backslash && 2\,(W_1+W_2+\quad+W_{j\ 1})\\
&) && 2\,W_j && \text{and} && 2\,(W_1+W_2+\quad+W_{j\ 1})\\
&) && 2\,W_j && \text{and} && = {}_1+{}_2+\quad+{}_{j\ 1} \quad \text{where each } {}_i\,2\,W_i
\end{aligned}
$$

$$
\begin{aligned}
) \quad & {}_1+{}_2+\quad+{}_{j\ 1}+(\ 1) && = 0\\
) \quad & {}_1+{}_2+\quad+{}_{j\ 1}+(\ 1) && = 0\\
) \quad & {}_1={}_2=\quad={}_{j\ 1}= && = 0
\end{aligned}
$$

Hence (a)) (b).

(b)) (c):

Assume that for each j, $2 \quad j \quad k$, we have

$$W_j \backslash (W_1+W_2+\quad+W_{j\ 1}) = \text{f0g}$$

Claim: $W_1; W_2; \quad ; W_k$ are independent.

i:e:; to prove that ${}_1+{}_2+{}_k=0$) each ${}_i=0$ $({}_i\,2\,W_i)$.

$$\text{Let } 0 = {}_1+{}_2+\quad+{}_k \quad (\,{}_i\,2\,W_i \tag{8.3}$$

If possible, let some of the ${}^0_i$s are non-zero.

Let j be the largest interger i such that ${}_i\,6\!= 0$

i:e:; ${}_{j+1} = {}_{j+2} = \quad = {}_k = 0$:

) (8.3)) ${}_1 + \quad + {}_j = \quad (\,{}_j\,6\!= 0)$

) ${}_j = (\ 1)\,{}_1 + (\ 1)\,{}_2 + \quad + (\ 1)\,{}_{j\ 1}$

Here ${}_j\,2\,W_j$ and ${}_j \quad 0$.

Also, $(\ 1)\,{}_1 + (\ 1)\,{}_2 + \quad + (\ 1)\,{}_{j\ 1} \quad 0$.

) Both W_j and $W_1 + W_2 + \quad + W_{j\ 1}$ contains a non-zero element.

) $W_j \backslash (W_1 + \quad + W_{j\ 1}) \quad 0$.

This contradicts the hypothesis (b).

Thus, each ${}_i = 0$:

Hence (b)) (a):

Thus (a)) (b) and (b)) (a):

Now. we shall prove that (a)) (c) and (c)) (a).

(a)) (c) :

Assume that $W_1; W_2;\quad ; W_k$ are independent.

Given that B_i $(i = 1; 2;\quad ; k)$ is an ordered basis for W_i and $B = (B_1; B_2;\quad ; B_k)$:

We know that any linear relation between the vectors in B will be of the form

$$ {}_1 + {}_2 + \quad + {}_k = 0 \qquad (8.4)$$

where $_i$ is some linear combination of elements of B_i .

Since $W_1; W_2;\quad ; W_k$ are independent and hence each $_i = 0$.

Since each B_i is independent, which implies that the linear combination of elements of $B_i = 0$ $(i = 1; 2;\quad ; k)$

Thus, the associated scalars are all zero.

Hence, the relation (8.4) is a trivial relation.

This proves (a)) (c):

Similarly, we can prove that (c)) (a) .

This completes the proof of the theorem.

Definition 8.2. If any (and hence all) of the conditions of the previous lemma hold, we say that W is the direct sum of $W_1; W_2;\ :W_k$ and denote it by $W = W_1\ W_2\ W_k$:

Note 8.2. W is also called the independent sum of $W_1; W_2;\quad ; W_k$ (or) the interior direct sum of $W_1; W_2;\ ; W_k$.

Example 8.1. Let V be a finite-dimensional vector space over the field F and let f $_1$; $_2$; ; $_n$g be any basis for V . If W_i is the one-dimensional subspace spanned by $_i$; then $V = W_1\quad W_2\qquad W_n$:

Example 8.2. Let n be a positive integer and F a subfield of the complex numbers, and let V be the space of all $n\quad n$ matrices over F . Let W_i be the subspace of all symmetric matrices, i:e:; matrices A such that $A^t = A$. Let W_2 be the subspace of allskew-symmetric matrices, i:e:; matrices A such that $A^t = \quad A$. Then $V = W_1\quad W_2$. If A is any matrix in V , the unique expression for A as a sum of matrices, one in W_1 and the other in W_2 .

$$\begin{aligned} \text{i:e:;}\quad A &= A_1 + A_2 \\ \text{where } A_1 &= \frac{1}{2}(A + A^t) \\ A_2 &= \frac{1}{2}(A \quad A^t) \end{aligned}$$

Definition 8.3. Let V be a vector space. A linear operator E on V is called a projection of V if $E^2 = E$:

Remark 8.1. Suppose that E is a porojection. Let R be the range of E and let N be the null space of E .

1. The vector is in the range of R if and only if $E \quad = \quad$. If $\quad = E \quad$, then$E \quad = E^2 \quad = E \quad = \quad$. Conversely, if $\quad = E \quad$, then is in the range of E .

2. $V = R \quad N$:

3. The unique expression for as a sum of vectors in R and N is $\quad = E \quad + (\quad E \quad)$

From (1), (2) and (3) it is easy to see the following:

Definition 8.4. If R and N are sub-spaces of V such that $V = R \quad N$, there is one and only one projection operator E which has range R and null space N . That operator is called the projection on R along N.

Theorem 8.2. If $V = W_1 \quad W_2 \qquad W_k$: , then there exists k linear operators $E_1; E_2; \;; Ek$ on V such that

(i) each E_i is a projection $(E_i^2 = E_i)$;

(ii) $E_iE_j = 0$, if $i \neq j$;

(iii) $I = E_1 + \quad + E_k$;

(iv) the range of E_i is W_i .

Conversely, if $E_1; E_2; \;; E_k$ are k linear operators on V which satisfy conditions (i); (ii) and (iii) , and if we let W_i be the range of E_i , then $V = W_1 \; W_k$:

Proof. Assume that $V = W_1 \qquad W_k$.

For each j define an operator E_j on V as follows: E_j is well defined:

$$
\begin{aligned}
\text{Let } E_j &= E_j \; ; \quad 8\, j \\
) \quad {}_j &= {}_j \quad 8\, j = 1; 2; \quad ; k \\
) \quad {}_1 &= {}_1; \; {}_2 = {}_2; \quad ; \; {}_k = {}_k \\
) \quad {}_1 + {}_2 + \quad + {}_k &= {}_1 + {}_2 + \quad + {}_k \\
) &=
\end{aligned}
$$

E_j is linear:

Let $= {}_1 + {}_2 + \quad + {}_k; \quad = {}_1 + {}_2 + \quad + {}_k \; ({}_i; {}_i \, 2 \, W_i)$ and $c \, 2 \, F$.

$$
\begin{aligned}
\text{Consider } c \; + &= c({}_1 + \quad + {}_k) + ({}_1 + \quad + {}_k) \\
&= (c \, {}_1 + c \, {}_2 + \quad + c \, {}_k) + ({}_1 + \quad + {}_k) \\
&= (c \, {}_1 + {}_1; c \, {}_2 + {}_2; \quad ; c \, {}_k + {}_k) \\
) \; E_j(c \; + \;) &= c \, {}_j + {}_j \\
&= c \, E_j \; + E_j
\end{aligned}
$$

Thus, E_j is linear.

Range of E_j :

Let ${}_j \, 2$ Range of E_j.

i:e:; there exists an element 2 V such that $E_j = {}_j$ where ${}_j \, 2 \; W_j$.

i:e:; For all ${}_j \, 2 \; w_j$, there exists an element in V such that $E_j = {}_j$.

Thus, the range of E_j is W_j.

E_j is a projection:

$$
\begin{aligned}
\text{Consider } E_j^2 &= E_j(E_j \;) \\
&= E_j \\
) \quad E_j^2 &= E_j
\end{aligned}
$$

Null space of E_j :

We know that if 2 null space of E_j then $E_j = 0$ which implies ${}_j = 0$.

Thus, $= {}_1 + \quad + {}_{j\;1} + {}_{j+1} + \quad + {}_k$.

Hence, the null space of E_j is the subspace

$$W_1 + \quad + W_{j\;1} + W_{j+1} + \quad + W_k \tag{8.5}$$

$$\begin{aligned} \text{Now } E_j &= {}_j\, 8 j\\ ) \quad E_1 &= {}_1\\ E_2 &= {}_2\\ &:\\ E_k &= {}_k\\ ) &= {}_1 + {}_2 + \quad + {}_k\\ ) &= E_1 \quad + \quad + E_k 0\\ ) \quad I &= (E_1 \quad + \quad + E_k 0\ )\\ I &= E_1 \quad + \quad + E_k 0 \end{aligned}$$

Thus, the null space of E_i is $W_1 + \quad + W_{i\ 1} + W_{i+1} + \quad + W_k$.

(Note that when (i 6= j), the subspace W_j is part of the sum in the right hand side)

$$\begin{aligned} \text{If } i\ 6= j;\ \ E_iE_j(\ ) &= E_i(E_j(\ ))\\ &= E_i(\ _j)\\ &= 0 \end{aligned}$$

This proves the necessary part.

Su cient Part: Assume that $E_1; E_2; \quad ; E_k$ are linear operators on V which satisfy the conditions $(i); (ii)$ and (iii) and range of E_i is W_i:

$$\begin{aligned} \text{Now,} \quad &= {}_1 + {}_2 + \quad + {}_k\\ &= E_1 \ + E_2 \ + \quad + E_k \end{aligned}$$

where 2 V and E_j 2 $W_j(j = 1; 2; \quad ; k)$

) $V = W_1 + W_2 + \quad + W_k$.

It remains to prove that the expression for is unique.

Now $= {}_1 + {}_2 + \quad + {}_k$ where ${}_i$ 2 W_i , say ${}_i = E_i\ {}_i$

$$
\begin{aligned}
\text{consider } E_j \quad &= E_j\left(\sum_{i=1} {}_i\right)\\
&= \sum_{i=1}^{n} E_j \quad {}_i\\
&= \sum_{i=1}^{n} E_j(E_i \quad {}_i)\\
&= E_j^2 \quad {}_j\\
&= E_j \quad {}_j\\
&= \quad {}_j
\end{aligned}
$$

Thus, the expression for is unique.

$)$ $V = W_1 \qquad W_k$

Hence the theorem.

8.2. Invariant Direct Sums

Our aim is to study the direct sum decompositions $V = W_1 \quad W_k$, where each of the subspaces W_i is invariant under some given linear operator T.

Given such a decomposition, the linear operator T induces a linear operator T_i on each W_i, by means of restricting T to the subspace W_i.

In this context, we have the following:

$$
\begin{aligned}
\text{If} \quad 2\, V; \ 9 \quad {}_i\, 2\, W_i \text{ such that}&\\
= \quad {}_1 + \quad + \ {}_k&\\
\text{Then } T(\ ) = T_1 \ {}_1 + T_2 \ {}_2 + \quad + T_k \ {}_k&
\end{aligned}
$$

We describe this situation by saying that T is the Direct sum of the operators $T_1; T_2; \quad ; T_k$.

However, here

1. T_i are not linear operators on V but on the respective subspace W_i only.

2. $V = W_1 \qquad W_k$ enables us to associate with each in V, a unique

k-tuple (of vectgors $_i\ 2\ W_i$) say, ($_1$; $_2$; ; $_k$).

$i{:}e{:};$ by $= {}_1 + {}_2 + \ + {}_k$; in such a way that we can carry linear operation in V by working in the individual subspaces W_i.

3. The fact that, each W_i is invariant under T, enables us to view the action of T as the independent action of $T^0_i s$ on $W^0_i s$.

4. Our purpose is to study T by nding :invariant direct sum decomposition, in which T_i are opoerators of elementary nature.

5. Let us note the matrix analogue of this situation. Let B_i denote an ordered basis for each W_i and let B be the ordered basis for V, consisiting of the union of B_i arranged in the order $B_1; B_2;\ ; B_k$.

   If $A = [T]_B$ and $A_i = [T]_{B_i}$ then A has the block form

$$A = \begin{bmatrix} A_1 & 0 & & 0 \\ 0 & A_2 & 0 & \\ \cdot & & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & & A_k \end{bmatrix}$$

   where each A_i is a $(d_i\ d_i)$ matrix where $d_i = \dim W_i$ and $0^0\ s$ are rectangular matrices of zeros of various order. In this case, we say that A is the direct sum of the matrices $A_1; A_2;\ ; A_k$.

6. More often, we shall describe the subspace W_i by means of the associated projections E_i.

7. Hence, we need to phrase the invariance of the subspace W_i in terms of the E_i.

Theorem 8.3. Let T be a linear operator on the space V, and let $W_1; W_2;\ ; W_k$ and $E_1; E_2;\ :E_k$ be as in Theorem 8.2. Then a necessary and su cient condition that each subspace W_i be invariant under T is that T commute with each of the projections E_i $i{:}e{:};$

$$TE_i = E_iT; \quad i = 1; 2;\ ; k$$

Proof. Assume that T coomutes with each E_i. Let be in W_j. Then

$$i{:}e{:}; \quad TE_i = E_iT \quad (i = 1; 2;\ ; k)$$

Now, our claim is $W\ \ i$ is invariant under T.

i:e:; to prove that $T(W_j \quad W_j)$:

But, we know that range of $E_j \quad W_j$.

Hence, it is enough to prove that $T(W_j) \in$ Range of E_j .

For if, let $\in W_j \Rightarrow E_j =$:

$$\begin{aligned}
\text{Consider } T &= T(E_j \quad) \\
&= E_j(T \quad) \\
&\in \text{Range of } E_j \\
\text{i:e:; } T &\in \text{Range of } E_j; \text{ whenever } \in W_j \\
\Rightarrow T(W_j) & \quad \text{Range of } E_j:
\end{aligned}$$

This proves necessary part.

Sufficient Part:

Assume that each W_i is invariant under T .

i:e:; to prove that T commutes with each of the projection E_j .

i:e:; to prove that $TE_j = E_jT$:

Let be any vector in V . Then we know that

$$\begin{aligned}
&= E_1 \quad + \quad + E_k \\
T &= TE_1 \quad + \quad + TE_k \\
\text{Here } E_i &\in W_i \quad (i = 1; 2; \quad ; k) \\
\Rightarrow T(E_i \ ) &\in W_i \\
\Rightarrow T(E_i \ ) &= E_i \ _i \quad \text{for some vector } _i \\
\text{Now, consider } E_jTE_i &= E_jE_i \ _i \\
&= E_j \ _j \\
&= \begin{cases} 0 & \text{if } i \quad j \\ E_j \ _j & \text{if } i = j \end{cases} \\
) \ E_jT &= E_jT \\
&= E_j(TE_1 \quad + \quad + TE_k \ ) \\
&= E_jTE_1 \quad + \quad + E_jTE_j \quad + \quad + E_jTE_k \\
&= 0 + 0 + \quad + E_jTE_j \quad + 0 + 0 + \quad + 0 \\
&= E_j \ _j = T(E_j \ ) \\
\text{i:e:; } \forall; E_jT &= TE_j \\
\Rightarrow E_jT &= TE_j
\end{aligned}$$

This proves the sufficient part.

Theorem 8.4. Let T be a linear operator on a finite-dimensional space V. If T is diagonalizable and if $c_1; c_2; \quad ; c_k$ are the distinct characateristic values of T, then there exist linear operator $E_1; \ ; E_k$ on V such that

(i) $T = c_1E_1 + \quad + c_kE_k$;

(ii) $I = E_1 + \quad + E_k$;

(iii) $E_iE_j = 0; \quad i \neq j$;

(iv) $E_i^2 = E_i$ (E_i is a projection);

(v) the range of E_i, is the characteristic space for T associated with c_i.

Conversely, if there exists k distinct scalars $c_1; c_2; \ ; c_k$ and k non-zero linear operators $E_1; E_2; \ ; E_k$ which satisfy conditions (i); (ii) and (iii) , then T is diagonalizable, $c_1; c_2; \ ; c_k$ are the distinct characteristic values of T and conditions (iv) and (v) are satisifed also.

Proof. Suppose that T is diagonalizable, with distinct characteristic values $c_1; \ ; c_k$.

Let W_i be the space of characteristic vectors associated with the characteristic value c_i. In this case, we know that,

$$V = W_1 \quad W_2 \qquad W_k$$

Let $E_1; E_2; \ ; E_k$ be the projections associated with this decomposition as in Theorem 8.2.

$\Rightarrow$ conditons (ii) to (iv) are satisifed.

To verify (i) we proceed as follows:

$$\begin{aligned}
\text{now } I &= E_1 + \quad + E_k \\
\Rightarrow \quad I &= E_1 \quad + \quad + E_k \\
\Rightarrow \quad &= E_1 \quad + \quad + E_k \\
\Rightarrow \quad T &= T(E_1 \quad + \quad + E_k) \\
&= E_1T \quad + \quad + E_kT \\
&= E_1(c_1 \quad) + \quad + E_k(c_k \quad) \\
&= (c_1E_1) \quad + c_kE_k) \\
T &= c_1E_1) \quad + c_kE_k
\end{aligned}$$

This proves necessary part.

Sufficient Part:Assume that we are given a linear operator T along with distinct scalars $c_1; c_2; \cdots; c_k$ and non-zero operators $E_1; E_2; \cdots; E_k$ which satisfy (i); (ii) and (iii).

Now (iii) $\Rightarrow$ $E_iE_j = 0 \ \forall i \neq j:$

(ii) $\Rightarrow$ $I = E_1 + \cdots + E_i + \cdots + E_k$.

$\Rightarrow$ $E_i = E_1E_i + \cdots + E_{i-1}E_i + E_i^2 + E_{i+1}E_i + \cdots + E_kE_i$

$\Rightarrow$ $E_i = E_i^2$.

$$\begin{aligned}
\text{(i)} \Rightarrow \quad T &= c_1E_1 + \cdots + C_kE_k \\
\Rightarrow \quad TE_i &= c_1E_1E_i + \cdots + c_iE_iE_i + \cdots + c_kE_kE_i \\
\Rightarrow \quad TE_i &= c_iE_i^2 \\
\Rightarrow \quad TE_i &= c_iE_i \\
\Rightarrow \quad (T - c_iI)E_i &= 0 \qquad (8.6)
\end{aligned}$$

Therefore, any vector in the range of E_i is in the null space of $T - c_iI$.

Also given that operator $E_i \neq 0$:

$\Rightarrow$ (8.4), there exists $\alpha \neq 0$ in E_i such that $(T - c_iI)\alpha = 0$:

$\Rightarrow$ $T\alpha = c_i\alpha$.

ie:; c_i is a characteristic value of T $(i = 1; 2; \cdots; k)$.

Claim: These c_i are the only characteristic value of T.

i:e:; to prove that, if c is any other characteristic value of T, then $c_i = c$:

If possible, let c be any other characteristic value of T.

Then, by definition,

$$\begin{aligned}
T\alpha &= c\alpha \\
\Rightarrow \ (T - c_i)\alpha &= 0 \\
\text{Now, by part(i); } T &= c_1E_1 + \cdots + c_kE_k \\
\text{part (ii)} \Rightarrow \ I &= E_1 + \cdots + E_k \\
\Rightarrow \ cI &= cE_1 + \cdots + cE_k
\end{aligned}$$

$$
\begin{aligned}
) \; T \quad cI &= (c_1 \quad c)E_1 + \quad + (c_k \quad c)E_k \\
) \; (T \quad cI) &= (c_1 \quad c)E_1 \; + \quad + (c_k \quad c)E_k \\
0 &= (c_1 \quad c)E_1 \; + \quad + (c_k \quad c)E_k \\
) \; (c_i \;\; c)E_i &= 0 \quad 8 i = 1; 2; \quad ; k \\
) \; (c_i \;\; c) &= 0 \\
) \; C_i &= c
\end{aligned}
$$

This proves our claim.

We have shown that, every non-zero vector in the range of E_i is a characteristic vector of T .

Also, $I = E_1 + E_2 + \quad + E_k$ show that these characteristic vectors span V .

Thus, T is diagonalizable.

In order to complete the proof, it remains to show that the null space of $(T \;\; c_iI)$ is the range of E_i .

Let $\;\; 2$ the null space of $T \quad c_iI \;\;) \;\; t \;\; = c_i \;\;$.

We have $T \;\; = c_1E_1 \;\; + \quad + c_jE_j \;\; + \quad + c_kE_k$:

Also, $\;\; = E_1 \;\; + \quad + E_k$.

$) \quad c_i \;\; = c_iE_1 \;\; + \quad + c_iE_j \;\; + \quad + c_iE_k$

$); \quad T \quad c_i \;\; = (c_1 \quad c_i)E_1 \;\; + \quad + (c_j \quad c_i)E_j \;\; + \quad + (c_k \quad c_i)E_k$

$0 = \sum_{j=1}^{k} (c_j \quad c_i)E_j \;\;$.

$) \quad (c_j \quad c_i)E_j \;\; = 0 \;\; 8 j = 1; 2; \quad ; k$:

$) \quad E_j \;\; = 0; \;\; 8 j \; 6= \; i \;\; (* \; c_j \;\; ç \;\; 0 \;\; 8 i \quad j)$:

We know that $\;\; = E_1 \;\; + \quad + e_{i \;\; 1} \;\; + \quad + E_i \;\; + E_{i+1} \;\; + \quad + E_k$:

$= 0 + \quad + 0 + E_i \;\; + 0 + \quad + 0$:

i:e:; $\;\; = E_i$

alpha .

Thus $\;\; 2$ The range of E_i .

Hence the null space of $T \quad c_iI$ is the range of E_i .

This completes the proof.

8.3. The Primary Decomposition Theorem

In this section, we are trying to study a linear operator T on the finite-dimensional space V by decomposing T into a direct sum of operators whhich are in some sense elementary.

Theorem 8.5. (Primary Decomposition Theorem) Let T be a linear operator on the finite-dimensional vector space V over the field F . Let p be the minimal polynomial for T ,

$$p = p_1^{r_1} \quad p_k^{r_k}$$

where the p_i are distinct irreducible monic polynomials over F and the r_i are positive integers. Let W_i be the null space of $p_i(T)^{r_i}$; $i = 1; 2; \quad ; k$. Then

(i) $V = W_1 \qquad W_k$;

(ii) each W_i is invariant under T ;

(iii) if T_i is the operator induced on W_i by T , the minimal polynomial for ${}_i$ is $p_i^{r_i}$

Proof. The idea of the proof is as follows:

If we assume that the direct sum decomposition in part (i) is valid, what would we think of the projections $E_1; E_2; \quad ; E_k$ associated with this decomposition.

The fact is that such a projection E_i will be the identity on W_i and zero on the other W_j .

We have to find a polynomial say h_i such that $h_i(T)$ is the identity on W_i and on the other W_j , which will imply that

$$\begin{aligned} h_1(T) + \quad + h_{i\ 1}(T) + h_i(T) + h_{i+1}(T) + \quad + h_k(T) &= 0 + 0 + \quad + 0 + \quad + 0 \\ &= I \end{aligned}$$

Given that $p = p_1^{r_1} \quad p_k^{r_k}$ where p_i are irreducible, monic polynomial over F and r_i are integers.

For every $i = 1; 2; \quad ; k$, define

$$
\begin{aligned}
f_i &= \frac{p}{p_i^{r_i}} \\
&= \frac{p_1^{r_1} \quad p_{i\ 1}^{r_i\ 1} p_i^{r_i} p_{i+1}^{r_{i+1}}\ p_k^{r_k}}{p_i^{r_i}} \\
&= p_1^{r_1} \quad p_{i\ 1}^{r_i\ 1} p_{i+1}^{r_{i+1}} \quad p_k^{r_k} \\
&= \prod_{j=i} p_j^{r_j}
\end{aligned}
$$

Since $p_1; p_2; \quad ; p_k$ are distinct prime polynomials, the polynomials $f_1; f_2; \quad ; f_k$ are relatively prime, so we can nd polynomials $g_1; g_2; \ ; g_k$ such that

$$
\begin{aligned}
f_1g_1 + f_2g_2 + \quad + f_kg_k &= 1 \\
\text{i:e:;} \sum_{i=1}^{n} f_ig_i &= 1
\end{aligned}
$$

Note also that, if $i \ 6= j$, then $f_i f_j$ is divisible by the polynomial p, because $f_i f_j$ contains each $p_m^{r_m}$ as a factor.

Now, we shall prove that the polynomials $h_i = f_ig_i$ satisfy in the rst paragraph of this theorem.

For this purpose,

$$
\begin{aligned}
\text{Let } E_i &= h_i(T) \\
&= f_i(T)g_i(T)
\end{aligned}
$$

Then $h_1 + \quad + h_k = f_1g_1 + \quad + f_kg_k = 1$.

and $p = f_i f_j \ \ 8i \ 6= j;$ implies that $E_iE_j = 0$ if $i \ 6= j:$

Thus, E_i are the projections which correspond to some direct-sum decomposition of V.

We wish to show that the range of E_i is exactly the subspace of W_i.

Let 2 the range of E_i which implies that $E_i \ = \ :$

Given that W_i is the null space of $(p_i(T))^{r_i}$ $(i = 1; 2; \quad ; k)$

$$
\begin{aligned}
(p_i(T)^{r}\ {}_i &= (p_i(T)^{r}\ {}_i E_i \\
&= (p_i(T)^{r}\ f_i(T)g_i(T) \ = 0 \\
\text{i:e:;} \quad (p_i(T)^{r_i} &= 0; \quad \text{if} \quad 2 \text{ range of } E_i \\
) \quad & 2 \quad \text{The null space of } (p_i(T))^{r_i} \\
) \quad & 2 \quad W_i \\
) \quad \text{Range of } E_i & \quad W_i
\end{aligned}
$$

Now, let 2 Null space of $(p_i(T)^{r}$

We know that, if i 6= j , then f_iG_j is divisible by $_iP^{r_i}$

) $f_j(T)g_j(T)0$ (* p_i^r is prime)

(or) $f_j(T)g_j(T) = 0$

) $E_j = 0$ for j 6=

i Since

$$E_1 + E_2 + \quad + E_k = 1$$

$$) \; E_1 + \quad + E_{i\ 1} + E_i + E_{i+1} + \quad + E_k =$$

$$) \; E_i =$$

$$) \; 2 \text{ The Range of } E_i$$

$$\text{Thus } W_i \quad \text{Range of} E_i$$

) Range of $W_i = W_i$.

Thus, we have $V = W_1 \qquad W_k$.

This completes the proof of statement (i) .

Obviously by their construction, the subspaces W_i are invariant under T .

If T_i isthe operator induced by T on the subspace W_i .

W_i is the null space of $p_i(T)^{r_i}$.

) $p_i(T)^r= 0$ on W_i .

) $p_i(T)^r = 0$

) T_i satisfy the polynomial p_i^r.

Thus, the minimal polynomial for T_i divides $p_i^{r_i}$.

Conversely, let g be the minimal polynomial for T_i .

i:e:; let g be any polynomial such that $g(T_i) = 0$.

i:e:; $g(T) = 0$ (* T_i is induced by T on W_i).

i:e:; $g(T) f_i(T) = 0$

i:e:; T satisfy the polynomial $g f_i$.

i:e:; $f g_i$ is divisible by p .

i:e:; p divides $g f_i$

i:e:; $p_i^{r_i} f_i$ divides $g f_i$

i:e:; $p_i^{r_i}$ divides g.

$\Rightarrow$ $p_i^{r_i}$ divides the minimal polynomial for T_i.

Thus, the minimal polynomial for T_i is $p_i^{r_i}$.

This completes the proof of the theorem.

Definition 8.5. Let N be a linear operator on the vector space V. We say that N is nilpotent if there is some positive integer r such that $N^r = 0$:

Theorem 8.6. Let T be a linear operator on the finite-dimensional vector space over the field F. Suppose that the minimal polynomial for T decomposes over F into a product of linear polynomials. Then there is a diagonalizable operator D on V and a nilpotent operator N on V such that

(i) $T = D + N$,

(ii) $DN = ND$.

The diagonalizable operator D and the nilpotent operator N are uniquely determined by (i) and (ii) and each of them is a polynomial in T.

Proof. Recall the proof of Primary Decomposition Theorem. Using this notation, we may assume the special case that the minimal polynomial for T is a prdocut of first degree polynomials.

i:e:; p_i is of the form $p_i = x - c_i$:

We know that the range of $E_i = W_i =$ The null space of $(T - c_i I)^{r_i}$.

$$\text{Now, let } D = c_1E_1 + c_2E_2 + \cdots + c_kE_k \tag{8.7}$$

i:e:; D is a diagonalizable operator.

Let us now define $N = T - D$, where

$$\begin{aligned} T &= TE_1 + \cdots + TE_k \\ D &= c_1E_1 + \cdots + c_kE_k \\ \Rightarrow T - D &= (T - c_1I)E_1 + \cdots + (T - c_kI)E_k \\ \text{i:e:; } N &= (T - c_1I)E_1 + \cdots + (T - c_kI)E_k \end{aligned}$$

$$N^2 = (T \quad c_1I)^2E_1^2 + \quad + (T \quad c_kI)^2E_k^2$$

$$2\,((T \quad c_1I)(T \quad c_2I)E_1E_2 +$$

$$+ (T \quad c_{k\ 1}I)(T \quad c_kI)E_{k\ 1}E_k)$$

) $N^2 = (T \quad c_1I)^2E_1 + \quad + (T \quad c_kI)^2E_k \quad (*\ E_iE_j = 0$ 8i j:)

.

$N^r = (T \quad c_1I)^rE_1 + \quad + (T \quad c_kI)^rE_k \quad (*\ E_iE_j = 0 \qquad$ j:)

8i6=

) $N^r = 0$

Thus, N is nilpotent.

Thus, we have $T = D + N$, where D is diagonalizable and N is nilpotent, implies that D and N compute with each other and that they are polynomials in T.

) $DN = ND$:

) It remains to show that the representation $T = D + N$ is unique.

If possible, let $T = D^0 + N^0$ where D^0 is diagonalizable and N^0 is nilpotent, satisfying

$$D^0N^0 = N^0D^0:$$

Then to prove that $D = D^0$ and $N = N^0$.

Now, we shall prove that D^0 commutes with $T = D^0 + N^0$.

$$TD^0 = (D^0 + N^0)D^0$$

$$= D^{0\ 2} + N^0D^0$$

$$D^0T = D^0(D^0 + N^0)$$

$$= {}^{0\ 2} + D^0N^0$$

$$)\ TD^0 = D^0T$$

Thus, D^0 commutes with T.

Similarly, N^0 commutes with T.

Thus, both D^0 and N^0 commute with T.

) both D^0 and N^0 commute with any polynomial in T.

) both D^0 and N^0 commute with D and N.

Thus, we have $D + N = D^0 + N^0$

(or) $D \quad D^{P()} = N \quad N^0$

The above discussion implies that all the four operators $D; D^0; N; N^0$ commute with one another.

Now, $D; D^0$ are both diagonalizable and $DD^0 = D^0D$.

Thus D and D^0 are simultaneously diagonalizable and hence $D \; D^0$ is diagonalizable3.

Now, N and N^0 are nilpotent and $NN^0 = N^0N$ and hence $N \quad N^0$ is nilpotent.

Note that, using the fact that D and D^0 commute with each other, we see that $D \; D^0$ is nilpotent.

Thus, $D \quad D^0$ is a diagonalizable and nilpotent operator.

But we note that the only operator which is both diagonalizable and nilpotent is zero operator.

) $D \quad D^0 = 0$ and $N^0 \quad N = 0$.

i:e:; $D = D^0$ and $N^0 = N$.

Hence the representation of $T = D + N$ is unique.

This completes the proof of the theorem.

Corollary 8.1. Let V be a finite-dimensional vector space over an algebrically closed field F , e:g:; the field of complex numbers. Then every linear operator T on V can be written as the sum of a diagonalizable operator D and a nilpotent operator N which commute. These operators D and N are unique and each is a polynomial in T .

Proof. The field F is said to be algebrically closed if every prime polynomial over F has degree 1.

Also write the proof of the above theorem.

Let us Sum Up:

In this unit, the students acquired knowledge to

explain the concept of invariant subspaces.

understand the concept of primary decomposition theorem.

Check Your Progress:

1. Let V be a finite-dimensional vector space and let W_1 be any subspace of V. Prove that there is a subspace W_2 of V such that $V = W_1\ W_2$.

2. Let V be a finite dimensional vector space and let $W_1;\ W_2;\ ;\ W_k$ be subspaces of V such that

$$V = W_1 + W_2 + \quad + W_k \quad \text{and} \ \dim V = \dim W_1 + \quad + \dim W_k$$

3. Let T be the linear operator on R^2 , the matrix of which in the standard basis is

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

Let W_1 be the subspace of R^2 spanned by the vector $\ _1 = (1; 0)$.

(a) Prove that W_1 is invariant under T.

(b) Prove that there is no subspace W_2 which is invariant under T and which is complimentary to W_1

$$R^2 = W_1\ W_2$$

Suggested Readings:

1. M. Artin, *Algebra* , Prentice Hall of India Pvt. Ltd., 2005.

2. S.H. Friedberg, A.J. Insel and L.E Spence, *Linear Algebra* , 4th Edition, Prentice-Hall of India Pvt. Ltd., 2009.

3. I.N. Herstein, *Topics in Algebra* , 2nd Edition, Wiley Eastern Ltd, New Delhi, 2013.

4. J.J. Rotman, *Advanced Modern Algebra* , 2nd Edition, Graduate Studies in Mathematics, Vol. 114, AMS, Providence, Rhode Island, 2010.

5. G. Strang, *Introduction to Linear Algebra* , 2nd Edition, Prentice Hall of India Pvt. Ltd, 2013.

BLOCK - V

Unit – 9: The Rational Forms

Unit – 10: The Jordan Forms

Block-V

UNIT-9

THE RATIONAL FORMS

Overview

In this unit, we shall describe how to generate cyclic subspaces.

Objectives

After successful completion of this lesson, students will be able to

understand the concept of cyclic subspaces.

understand the concept of Rational Form.

9.1. Cyclic subspaces and Annihilators

Let V be a finite dimensional vector space over the field F and let T be a fixed (but arbitrary) linear operator on V .

Note that, if is any vector in V , then there exist a smallest subspace of V , which is invariant under T and contains . This subspace can be defined as The intersection of all T -invariant subspaces which contain .

Remark 9.1. If W is any subspace of V , which is invariant under T and contains , then W must also contain the vector T . Hence W must contain

$$T(T\ ) = T^2\ ; \quad T(T^2\ ) = T^3\ ;$$

[ ) It contains $T \ + T^2 \ +$

(i:e:;) if $g(T) = T + T^2 + T^3 +$

then $g(T)$ is contained in W , where the polynomial $g(x)$ 2 $F[x]$.]

) W must contain $g(T)$, for every polynomial g over F .

Note 9.1. The set of all vectors of the form $g(T)$, with g in $F[x]$, is clearly invariant under T , and is this the smallest T -invariant subspace which contains

Definition 9.1. If is any vector in V , the T -cyclic subspace generated by , is the subspace $Z(\ ; T) = V;$ then is called a cyclic vector for T .

Note 9.2. Another way of describing the subspace $Z(\ ; T)$ is the subspace spanned by the vectors $T^0\ ;\ T^1\ ;\ \ ; T^k\ (k\ 0)$.

) is a cyclic vector for T if and only if these vectors $T^k(\ )$ span V .

Important Cautions: The general operator T has no cyclic vectors.

1. For any T and , we are interested in linear relations of the form

$$c_0T \ + c_1T \ + c_2T^2 \ + \ + c_kT^k \ = 0 \tag{9.1}$$

   between the vectors T^j .

2. In otherwords, we are interested in the polynomials

$$g = c_0 + c_1 x + \quad + c_k x^k$$

which satisfy (9.1), (or) $g(T) = 0$.

3. The set of all g 2 $F[x]$ such that $g(T) = 0$, is an ideal in $F[x]$.

4. The minimal polynomial for T, say, $p(T)$, satis es $p(T) = 0$) $p(T) = 0$ i:e:; The minimal polynomial $p(x)$ is in this ideal.

) This ideal is non-zero.

De nition 9.2. Let be any vector in V . The T -annihilator of is the ideal $M(; T)$ in $F[x]$, consisting of all polynomials g over F such that $g(T) = 0$:

The unique monic polynomial p which generates this ideal will also be called the T -annihilator of .

Note 9.3. The T -annihilator p divides the minimal polynomial of the operator T.

Theorem 9.1. Let be any non-zero vector in V and let p be the T -annihilator of .

(i) The degree of p is equal to the dimension of the cyclic subspace $Z(; T)$.

(ii) If the degree of p is k, then the vectors $; T ; T^2 ; \quad ; T^{k \ 1}$ form a basis for $Z(; T)$.

(iii) If U is the linear operator on $Z(; T)$ induced by T, then the minimal polynomial for U is p :

Proof. Let g be any polynomial over the eld F .

Given that p is the T -annihilator of .

) The unique monic polynomial p generates the ideal $M(; T)$ in $F[x]$, which consists of all polynomials g over F such that $g(T) = 0$:

Now, given g and p, using the division algorithm, we get

$$g \quad = 7p \ q + r \tag{9.2}$$

where either $r = 0$ (or) $\deg (r) < \deg p = k$:

Since p is the generator of $M(\ ; T)$

i:e:; Any multiple of p 2 $M(\ ; T)$

) $p \ q$ 2 $M(\ ; T)$

) $p \ q(T) \ = 0$.

$$\begin{aligned} \text{Now } g(T) &= p \ q(T) \ + r(T) \\ &= 0 + r(T) \\ \text{i:e:; } g(T) &= r(T) \end{aligned}$$

Here $r = 0$ (or) $\deg r < k$:

Therefore, the vectors $r(T)$ is a linear combination of the vectors $; T \ ; \quad ; T^{k \, 1}$.

Thus, $g(T)$ s a linear combination of the vectors $; T \ ; \quad ; T^{k \, 1}$.

i:e; The k vectors $; T \ ; \quad ; T^{k \, 1}$ span $Z(\ ; T)$.

Claim: The vectors $; T \ ; \quad ; T^{k \ 1}$ are linearly independent.If

possible assume that these vectors are linearly dependent.

) there exists a scalars $c_0; c_1; \quad ; c_k$ in F not all zero such that

$$c_0 \quad + c_1 T \quad + \quad + c_{k \ 1} T^{k \ 1} \quad = \quad 0$$

) $$c_0 \quad + c_1 T \quad + \quad + c_{k \ 1} T^{k \ 1} \quad = \quad 0$$

) $$g(T) = 0$$

$$\text{where } g(T) = c_0 \quad + c_1 T \quad + \quad + c_{k \ 1} T^{k \ 1}$$

) $$\deg g(T) = k \ \ 1 < k = \deg (p \ )$$

) $$\deg (g) < \deg (p \ )$$

which contradicts the fact that p is the minimal polynomial for T .

Hence the vectors $; T \ ; \quad ; T^{k \ 1}$ are linearly independent.

Therefore, the vectors f $; T \ ; \quad ; T^{k \, 1}$ g form a basis for $Z(\ ; T)$.

This also implies that the dimension of $Z(\ ; T) = k = \deg \ p$:

Hence, we have proved parts (i) and (ii) .

Let U be the linear operator on $Z(\ ; T)$ induced by T . () $p \ (U) = p \ (T)$).

Let g be any polynomial over F .

$$
\begin{aligned}
\text{Then; } p\ (U)g(T) &= p\ (T)g(T)\\
&= g(T)\ p\ (T)\\
&= g(T)+0\\
&= 0
\end{aligned}
$$

i:e:; $p\ (U)g(T) = 0$ where $g(T) \in Z(\ ;T)$.

$\Rightarrow$ $p\ (U)$ sends every vector in $Z(\ ;T)$ to zero.

$\Rightarrow$ $p\ (U)$ is the zero operator on $Z(\ ;T)$.

Claim: p is the minimal polynomial for U.

i:e:; to prove that $p\ (U) = 0$ and no polynomial of degree less than deg p satisfies U.

If possible, let h be any other polynomial of degree less than $k = \deg\ p$:

$$
\begin{aligned}
\text{Let } h(U) &= 0\\
\Rightarrow\ h(U) &= 0\\
\Rightarrow\ h(T) &= 0 \quad (* \ U \text{ is induced by} T)\\
\Rightarrow\ h(T) &\in M(\ ;T)
\end{aligned}
$$

where $\deg h < \deg p$ where p which is a contradiction.

$\Rightarrow$ p is the minimal polynomial for U.

Hence part (iii).

This completes the proof of the theorem.

Companion matrix of the monic polynomial p :

Consider a linear operator U on a space W of dimension K which as a cyclic vector :

[If $Z(\ ;T) = V$, then is called a cyclic vector for T where T is a linear operator on V.

$\Rightarrow$ is a cyclic vector of U, where U is a linear operator on $W \Rightarrow Z(\ ;U) = W$:]

Then by above theorem, (i) the vectors $\ ;U\ ;U2\ ;\quad ;Uk\quad 1$ form a basis for the space W and (ii) the annihilator p of is the minimal polynomial for U (and hence p is also the characteristic polynomial for

U).

For $i = 1; 2; \quad ; k$, let $_i = U^{i\ 1}$; then the action of U on the ordered basis $B = f\ _1; \ _2; \quad ; \ _k g$ is

$$U\ _i = \ _{i+1}; \quad i = 1; 2; \quad ; k \quad 1$$

$$U\ _k = \quad c_0\ _1 \quad c_1\ _2 \qquad c_{k\ 1}\ _k$$

where $p = c_0 + c_1 x + \quad + c_{k\ 1} x^{k\ 1} + x^k$: is a minimal polynomial for U.

$$) \quad p(U) = 0$$

$$c_0 + c_1 U + c_2 U^2 + \quad + c_{k\ 1} U^{k\ 1} + U^k = 0$$

$$) \quad c_o \ + c_1 U \ + c_2 U^2 \ + \quad + c_{k\ 1} U^{k\ 1} + U^{k} = 0$$

$$) \quad U^k = \quad c_0 \quad c_1 U \quad c_2 U^2 \qquad c_{k\ 1} U^{k\ 1}$$

where $f \ ; U \ ; U^2 \ ; \quad ; U^{k\ 1} \ g$ is a basis for W.

) The matrix of U in the ordered basis B is

$$\begin{bmatrix} 0 & 0 & 0 & & 0 & c_0 \\ 1 & 0 & 0 & & 0 & c_1 \\ & 1 & 0 & & 0 & c_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & & 1 & c_{k\ 1} \end{bmatrix} \tag{9.3}$$

This is called the companion matrix of the monic polynomial p :

Theorem 9.2. If U is a linear operator on the nite-dimensional space W , then U has a cyclic vector if and only if there is some ordered basis for W in which U is represented by the companion matrix of the minimal polynomial for U .

Proof. Rewrite the above companion matrix.

Corollary 9.1. If A is the companion matrix of a monic polynomial p , then p is both the minimal and the characteristic polynomial of A .

Proof. Rewrite the proof of Theorem 9.1.

9.2. Cyclic Decompositions and the Rational Form

The primary purpose of this section is to prove that if T is any linear operator on a finite-dimensional space V, then there exists vectors $_1;\ _2;\quad ;\ _r$ in V such that V is a direct sum of T-cyclic subspaces.

$$V = Z(\ _1;T)\quad Z(\ _2;T)\qquad Z(\ _r;T)$$

This, in turn will show that the linear operator T is the direct sum of a finite number of linear operators, each of which has a cyclic vector.

Definition 9.3. If W is any subspace of a finite dimensional space V, then there exists a subspace W^{0} of V such that $V = W \quad W^0$. In fact, there will be many such subspaces W^0 satisfying $V = W\ W^0$. Each of this subspace is said to be complementary to W.

Now, the question is: When a T-invariant subspace has a complementary subspace, which is also invariant under T.

Remark 9.2. Assume that $V = W\ W^0$ where both W and W^0 are invariant under T. We can now see what is special about W?

If $\ 2\ V$, then

$$V = W\ W^0$$
$$)\qquad = \quad + \ ^0$$

where $\ 2\ W;\ \ ^0\ 2\ W^0$.

Let f be any polynomial over the scalar field. Then

$$f(T) \quad = \quad f(T) \ + f(T)\ ^0$$

Since W is invariant under T, which implies that $TW \quad W$.

$)\quad\ 2\ W\ )\quad f(T)\ 2\ W$.

similarly, W^0 is invariant under T, which implies that $TW^0 \quad W$.

$)\quad\ ^0\ 2\ W^0\ )\quad f(T)r^0\ 2\ W^0$.

$)$ $f(T)$ = a sum of an element in W + an element in W^0

$)$ $f(T)$ = an element of W if and only if this element of $W^0 = 0$

$)$ $f(T)$ 2 W if and only if $f(T)\ {}^0 = 0$

When $f(T)$ 2 w then $f(T) = f(T)$.

Definition 9.4. Let T be a linear operator on a vector space V and let W be a subspace of V. We say that W is T-admissible if

(i) W is invariant under T;

(ii) if $f(T)$ is in W, there exists a vector in W such that $f(T) = f(T)$:

Remark 9.3. If W is invariant and if W has a complementary invariant subspace, then W is admissible.

the Admissibility characterizes those invariant subspaces which have complementary invariant subspaces.

Discussion: Let W be a proper T-invariant subspace. Let us try to find a non-zero vector such that $W \setminus Z(; T) =$ f0g.

Choose some vector which is not in V.

Consider the T-conductor $S(; W) =$ fpolynomials g=$g(T)$ 2 Wg.

[Recall: The monic polynomial $f = S(; W)$ which generates the ideal $S(; W)$ is also called the T-generator of into W.]

Now, $f = S(; W)$ $)$ $f(T)$ 2 W .

$)$ If W is T-admissible, then by definition, there exists a inW such that

$$f(T) = f(T)$$

Let = and let g be any polynomial.

$)$ = where 2 W.

$)$ 2 W:

$)$ $g(T)$ will be in W if and only if $g(T)$ is in W.

$)$ $S(; W) = S(; W)$.

$)$ The polynomial f is also the T-conductor of into W.

But $f(T) = 0$ whichn implies that $g(T)$ 2 W if and only if $g(T) = 0$ $)$ $W = f0g$.

But, the subspace W is proper, i:e:; W $6= f0g$.

Thus, the only possibility left out is $Z(;T)$ and W are independent and f is the T-annihilator of :

Theorem 9.3 (Cyclic Decomposition Theorem). let T be a liner ooperator on a nite-dimensional vector space V and let W_0 be a proper T -admissible subspace of V . There exist non-zero vectors $_1;\ _2;\quad ;\ _r$ in V with respective T -annihilators $p_1;\ ;p_r$ such that

(i) $V = W_0 \quad Z(_1;T) \qquad Z(_r;T)$;

(ii) p_k divides $p_{k\ 1}$; $k = 2;3;\quad ;r$:

Furthermore, the inter r and the annihilators $p_1;\ p_2;\ ;\ p_r$ are uniquely determined by (i); (ii) and the fact that no $_k$ is 0 .

Proof. The proof is rather long; hence we shall divide into four steps.

Step I: Given that W_0 is a proper T-admissible subspace of V .

) By de nition, W_0 is invariant under T and if f 2 w_0 , then there exists 2 W_0 such that $f = f$:

Note that $S(;W)$ is the monic polynomial which generates the ideal $S(;W)$:

(0r) $S(;w)$ is a T-conductor of into W and it is the monic polynomial of least degree, which sends into W .

) Even the maximum of degree of such T-conductors cannot exceed the dimension of V .

$$\text{i:e:;}\quad 0 < \max\ \deg S(;W) \qquad \dim V$$

) We can choose a vector such that $\deg S(;W)$ attains the maximum.

Thus the subspace $W + Z(;T)$ is then T-invaraint and has a larger dimension than that of W . i:e:; so far.

Apply this process to $W = W_0$ and end up with a subspace $W_1 = W_0 + Z(;T)$ where $_1$ is a vector such that $\deg S(_1;W_1)$ is maximum.

If W_1 is an improper subspace, there is nothing to move.

) If W_1 is still proper: Apply the same process to W_1 and end up with $W_1 = W_1 + Z(\ _1; T)$ where $_2$ is such that deg $S(\ _2; W_2)$ is maximum.

Continuing in this process, we end up with $W_r = V$:

) There exists non-zero vectors $_1;\ _2;\ \ ;\ _r$ in V such that

(a) $V = W_0 + Z(\ _1; T) + Z(\ _2; T) + \ \ + Z(\ _r; T)$ and

(b) if $1\ \ k\ \ r$, and $W_k = W_0 + Z(\ _1; T) + Z(\ _2; T) + \ + Z(\ _k; T)$ then the conductor $p_k = S(\ _k; W_k)$ has maximum degree (among all T-conductros into the subspace $W_{k\ 1}$.)

$$\text{i:e:; } 8k;\ \ \deg p_k = \max_{2V} \deg S(\ ; W_{k\ 1})$$

Hence the step I.

Step II: Given that $_1;\ _2;\ \ ;\ _r$ are the non-zero vectors which satisfy step I.

Fix k, $1\ \ k\ \ r$. Let be any vector in V and let $f = S(\ ; W_{k\ 1})$.

Let $f\ \ =\ _0 + \sum_{1\ i\ k} g_i\ _i\ \ (\ _i\ 2\ W_i)$.

Claim: f =each polynomial g_i and $_0 = f\ _0$ where $_0\ 2\ W_0$.

Let $k = 1$: Then $2\ V$ and $f = S(\ : W_0)$ and $f\ \ =\ _0\ \ (\ _0\ 2\ W_0)$.

$$f\ \ =\ \ S(\ ; W_0)$$

$$)\ \ f \text{ is a } T\text{-conductor of into } W_0$$

$$)\ \ f\ \ 2\ \ W_0$$

Now, we prove Step-II for $k > 1$:

Using the division algorithm, we get

$$g_i\ \ =\ \ f h_i + r_i$$

where either $r_i = 0$ (or) $\deg r_i < \deg f$.

If we want to claim that f =g_i, then the remainder $r_i = 0$; $8i$.

$$\text{Let } r\ \ =\ \ \sum_{i=1}^{k\ 1} h_i\ _i$$

$$)\ \ r\ \ =\ \ \sum_{i=1}^{k\ 1} h_i\ _i$$

$$)\ \ r\ \ 2\ \ W_{k\ 1}$$

$$)\ \ S(r; W_{k\ 1})\ =\ S(\ ; W_{k\ 1}) = f$$

Furthermore,

$$f \quad = \quad {}_0 + \sum_{i=1}^{k\ 1} r_i\ {}_i$$

Now, we shall show that $r_i = 0 \ \forall i$.

If possible, let $r_i \neq 0$:

Let j be the largest value of i such that $r_i \neq 0$: Then

$$f \quad = \quad {}_0 + \sum_{i=1}^{j} r_i\ {}_i; \qquad r_j \quad 0 \tag{9.4}$$

and $\deg r_j < \deg f$:

Let $p = S(r : W_{j\ 1})$

Since, $W_{j\ 1} \quad W_{k\ 1};$ where $f = S(\ : W_{k\ 1});$ $p = S(r : W_{j\ 1})$

$\Rightarrow$ f must divide p.

$\Rightarrow$ $p = fg$:

Multiplying both sides of (9.4) by $g(T)$, we get

$$\begin{aligned} gf &= g\ {}_0 + \sum_{i=1}^{j} gr_i\ {}_i \\ \Rightarrow \quad p &= gr_j\ {}_j + g\ {}_0 + \sum_{1\ i<j} gr_i\ {}_i \end{aligned}$$

Since $p \in W_{j\ 1};$ $g\ {}_0;$ $\sum_{1\ i<j} gr_i\ {}_i \in W_{j\ 1}$

$\Rightarrow$ $gr_j\ {}_j \in W_{j\ 1}$.

Since gr_j sends ${}_j$ into $W_{j\ 1}$.

$\Rightarrow$ $gr_j = S(\ {}_j : W_{j\ 1})$

$$\begin{aligned} \deg(gr_j \quad & \deg S(\ {}_j; W_{j\ 1})) \\ &= \deg p_j \\ & \quad \deg S(\ ; W_{j\ 1}) \\ &= \deg p \\ &= \deg(fg) \\ \Rightarrow \quad \deg r\,j \quad & \quad \deg\ f \end{aligned}$$

which contradicts the choice of j.

) $r_i = 0$ where $g_i = f h_i + r_i$.

) $g_i = f h_i$

) f divides each g_i and also $_0 = f$:

) $_0 = f\ _0$ where $_0$ 2 W_0 .

Hence Step-II.

Step-III: There exist non-zero vectors $_1; \quad ; _r$ in V which satisfy conditions (i) and (ii) .

Start with the vectors $_1;\ _2; \quad ; _r$ available in Step-I.

Fix k, $1\ k\ r$.

We apply Step-II to the vector $= \ _k$ and $f = p_k$, we obtain

$$p_k\ _k = p_k\ _0 + \underset{1\ i<k}{X} p_k h_i\ _i \qquad (9.5)$$

where $_0$ is in W_0 and $h_1; \quad ; h_{k\ 1}$ are polynomials. Let

$$_k = \ _k\ \ _0 \quad \underset{1\ i<k}{X} h_i\ _i \qquad (9.6)$$

Since $_k\ \ _k$ is in $W_{k\ 1}$.

$$\begin{aligned}) \quad S(\ _k; W_{k\ 1}) &= S(\ _k; W_{k\ 1}) = p_k && (9.7) \\ \text{Now } p_k\ _k &= p_{k\ k}\ p_{k\ 0} \quad X \quad p_k h_{i\ i} = 0 \\ ) \quad p_k\ _k &= 0 \\ ) \quad W_{k\ 1} \backslash Z(\ _k; T) &= 0 && (9.8) \end{aligned}$$

Note that each $_k$ satisfies (9.7) and (9.8).

) It follows that

$$W_k = W_0 \quad Z(\ _1; T) \qquad Z(\ _k; T) \qquad (9.9)$$

where p_k is the annihilator of $_k$.

Thus, the vectors $_1;\ _2; \quad ; _r$ define the same sequence of subspace $W_1; W_2;$ as do the vectors $_1;\ _2; \quad ; _r$. Also the T-conductors $p_k = S(\ _k; W_{k\ 1})$ have the same maximality properties (because of condition (b) of Step-I). The vectors $_1;\ _2; \quad ; _r$ have the additional property that the subspace $W_0; Z(\ _2; T);\ Z(\ _2; T); \ ; Z(\ _r; T)$ are independent.

Since $p_{k\ k} = 0$ 8k .

$) \quad p_k \;\; _k = 0 + p_1 \;\; _1 + \quad + p_{k\,1} \;\; _{k\,1}$.

$) \quad p_k{=}p_1;\, p_2; \quad ;\, p_{k\,1}$.

This proves conditions (ii) of the theorem.

Step-IV: The number r and the polynomials $p_1; p_2; \quad ; p_r$ are uniquely determined by the conditions of Theorem.

If possible, let there exists another set of non-zero vectors $r_1; r_2; \quad ; r_s$ with respective T -annihilators $g_1; g_2; \quad ; g_s$ such that

$$V \;=\; W_0 \quad Z(r_1; T) \qquad Z(r_s; T) \tag{9.10}$$

and g_k divides $g_{k\,1}$ for $k = 2; 3; \quad ; s$

Claim: $r = s$ and $p_i = g_i \;\; 8 i$.

To prove this, rst we shall prove that $p_1 = g_1$.

First, we observe that the polynomial g_1 is determined from (9.10) as the T -conductor of V into W_0 .

Let $S\,(V; W_0) = \text{f}\text{polynomials } f{=}f \;\; _0\, 2\; W_0;\;\; 8 \;\; 2\; V$.

i:e:; $S\,(V; W_0)$ contains polynomial f such that the range of $f\,(T)$ is contained in W_0 .

i:e:; $S\,(V; W_0)$ is a non-zero ideal in the polynomial algebra whose monic generator is the polynomial g_1 .

Since $\;\; 2\; V$

$$\begin{aligned} &= \;\; _0 + f_1 \;\; _1 + \quad + f_s \;\; _2 \\ g_1 \;\; &= \; g_1 \;\; _0 + g_1 f_1 \;\; _1 + \quad + g_1 f_s \;\; _s \\ &= \; g_1 \;\; _0 + \sum_{i=1}^{s} g_1 f_i \;\; _i \end{aligned}$$

Since each g_i divides g_1 , we have $g_{1]} \;\; _i = 0$ for all i and $g_1 \;\; = g_1 \;\; _0$ is in W_0 .

Thus g_1 is in S $(V; W_0)$. Since g_1 is the monic polynomial of least degree which sends $\;\; _1$ into W_0 . We see that g_1 is the monic polynomial of least degree in the ideal S $(V; W_0)$. By the same argument, p_1 is the generator of that ideal, so $p_1 = g_1$.

Let W be a subspace of V and let f be a polynomial.

De ne $fW = \text{f} f \;\; = \;\; 2\; W \text{g}$.

Then, we have

1. $f Z(\; ; T) = Z(f \; ; T)$
2. If $V = V_1$ V_k, where each V_i is invariant under T, then
$$fV = fV_1 \qquad fV_k \tag{9.11}$$
3. If and have the same T-annihilator, then f and f have the same T-annihilator and $\dim Z(f \; ; T) = \dim Z(f \; ; T)$.

Now, we proceed by induction to show that $r = s$ and $p_i = g_i$ for $i = 2; \; ; r$:

Since $p_1 = g_1$ is already proved, Hence, it is enough to prove that $r = s$ and $p_i = g_i$ for $i = 2; \; ; r$:

Now, our claim is that if $r \; 2$ then $p_2 = g_2$.

Let $r \; 2$:

$)$ $\dim W_0 + \dim Z(\;_1; T) < \dim V$

Since we know that $p_1 = g_1$ which implies that $Z(\;_1; T)$ and $Z(\;_1; T)$ have the same dimension.

$)$ $\dim W_0 + \dim Z(\;_1; T) < \dim V$

which shows that $S > 2$:

Now, we have two decompositions of V namely.

$$V = W_0 \quad Z(\;_1; T) \qquad Z(\;_r; T) \text{ and} \tag{9.12}$$
$$V = W_0 \quad Z(\;_1; T) \qquad Z(\;_s; T) \tag{9.13}$$

Inturn, the subspace p_2V will have two decompositions as follows,

$$p_2V = p_2W_0 \quad Z(p_2 \;_1; T) \quad \text{and}$$
$$p_2V = p_2W_0 \quad Z(p_2 \;_1; T) \qquad Z(p_2r_S; T)$$

Now $\dim Z(p_2 \;_1; T) = \dim Z(p_2 \;_1; T)$.

$)$ $\dim Z(p_2; T) = \dim Z(p_2 \;_3; T) = \quad = \dim Z(p_2 \;_S; T) = 0$.

i:e:; $\dim Z(p_2 \;_i; T) = 0$ 8i 2.

We conclude that $p_2 \;_2 = 0$ and g_2 divides p_2.

By interchanging the roles of p_2 and g_2, we can prove that $p_2 = g_2$.

) We have $p_2 = g_2$.

Proceeding like this, using the principle of induction, we get $r = s$ and that $p_i = g_i$ for $i = 1; 2; \quad ; r$.

Hence Step-IV.

This completes the proof of the theorem.

Corollary 9.2. If T is a linear operator on a finite-dimensional vector space, then every T-admissible subspace has a complementary subspace, which is also invariant under T.

Proof. Let W_0 be an admissible subspace of V.

Case (i) : Let W_0 be an improper subspace of V.

i:e:; let $W_0 = V$:

In this complement $W_0^0 = \{0\}$ which is invariant under T such that $V = W_0 \quad W_0^0$.

Case (ii) : Let $W_0 \quad V$. Then by above theorem, the complement of W_0 namely W_0^0 is given by

$$W_0^0 = Z(\ _1; T) \qquad Z(\ _r; T)$$

Then also W_0^0 is invariant under T such that $V = W_0 \quad W_0^0$.

Hence the corollary.

Corollary 9.3. Let T be a linear operator on a finite-dimensional vector space V.

(a) There exists a vector   in V such that the T-annihilator of   is the minimal polynomial for T.

(b) T has a cyclic vector if and only if the characteristic and minimal polynomials for T are identical.

Proof. If $V = \{0\}$, the results are trivially true.

If $V \neq \{0\}$, let

$$V = Z(\ _1; T) \qquad Z(\ _r; T) \tag{9.14}$$

where the T-annihilators $p_1; p_2; \quad ; p_r$ are such tht $p_{k+1}=p_k \quad (1 \quad k \quad r \quad 1)$:

As we noted in the proof of the above theorem, p_i is the minimal polynomial for T.

(or) p_1 is the T-conductor of V into $\{0\}$.

Hence (a) proved.

We know that if T is a cyclic vector, then the minimal polynomial for T coincides with the characteristic polynomial.

This proves (b).

Hence the corollary.

Theorem 9.4. (Generalized Cayley-Hamilton Theorem) Let T be a linear operator on a finite-dimensional vector space V. Let p and f be the minimal and characteristic polynomial for T respectively,

(i) p divides f.

(ii) p and f have the same prime factors, except for multiplicities.

(iii) If

$$p = f_1^{T_1} \quad f_k^{T_k}$$

is the prime factorization of p, then

$$f = f_1^{d_1} \quad f_k^{d_k}$$

where d_i is the nullity of $f_i(T)^{r_i}$ divided by the degree of f_i.

Proof. Case (i): If $V = \{0\}$, then the theorem is trivially true.

Case (ii): Let $V \neq \{0\}$.

As in the previous corollary, there exists a decomposition of V of the for,

$$V = Z(\ _1; T) \qquad Z(\ _r; T)$$

where the T-annihilator $p_1; p_2; \quad ; p_r$ are such that p_{k+1} divides p_k $(1 \quad k \quad r\ 1)$.

Also, we have the T-annihilator of is the minimal polynomial for T. (i:e:;)$p_1 = p$:

let U_i is the restriction of T to $Z(\ _i; T)$.

ie:; U_i has a cyclic vector.

i:e:; p_i is both the minimal polynomial and the characteristic polynomial for U_i $(i = 1; 2; \quad ; r)$.

Given that f is the characteristic polynomial for T.

$$f = p_1 p_2 \quad p_r \tag{9.15}$$

i:e:; p_1 divides f.

i:e:; p divides f.

Hence part (i).

Also, any prime divisor of p_1 is a prime divisor of f.

Since, $p_1 = p$, which implies that any prime divisior of p is a prime divisor of f.

Conversely: Any prime divisor of f is a prime divisor of one of the factors $p_1; p_2; \quad ; p_r$.

Thus, any prime divisor of f divides p_1.

Since $p = p_1$, any prime divisor of f divides p.

Hence p and f have the same prime factors, except for multipliciites.

Hence part (ii).

Given that $p = f_1^{r_1} \quad f_k^{r_k}$ is the prime factorization of p.

Let p be the minimal polynomial for T.

$$p = p_1^{r_1} \quad p_k^{r_k}$$

Let W_i is the null space of $p_i(T)^{r_i}$. Then

$$V = W_1 \qquad W_k$$

If V_i is the null space of $f_i(T)^{r_i}$, then

$$V = V_1 \qquad V_k$$

where $f_i^{r_i}$ is the minimal polynomial of the operator T_i (which is obtained by restricting T to V_i).

Consider the operator T_i and apply part (ii) of this theorem. (i:e:; if p and f are the minimal and characteristic polynomial of T then p and f have the same prime factors, except for multiplicities.)

Since, the minimal polynomial for T_i is some power of the prime f_i .

Thus, the characteristic polynomial for T_i is of the form $f_i^{d_i}$, where $d_i \quad r_i$:

Obviously,

$$d_i \ = \ \frac{\dim V_i}{\deg\ f_i}$$

Since V_i is the null space of $f_i(T)^{r_i}$.

) dim V_i = dim [Null space of $f_i(T)^{r_i}$]

) dim V_i is the nullity of $f_i(T)\ ^{r}.$

) $d_i = \dfrac{\text{Nullity of } f_i(T)^{r_i}}{\text{degree of } f_i}$.

Also, $T = T_1 \qquad T_k$.

i:e:; Characteristic polynomial for T is the product of characteristic polynomial of $T_1; T_2; \quad ; T_k$.

i:e:; $f = f_1^{d_1} f_2^{d_2} \qquad f_k^{d_k}$.

Hence part (iii) .

This completes the proof of the theorem.

Rational Form:

Let us look at the matrix analogue of the cyclic decomposition theorem.

i:e:; Assume that we have an operator T such that

$$V \ = \ W_0 \quad Z(\ \ _1; T) \qquad Z(\ \ _r; T)$$

Where $\ _1; \ \ _2; \quad ; \ _r$ are non-zero vectors in V .

Let $\mathcal{B}_i \ = \ \{ \ _i; T \ \ _i; T^2 \ \ _i; \quad ; T^{k_i\ \ 1} \ \ _i\}$ be the cyclic ordered basis for $Z(\ \ _i; T)$:

where k_i = dimension of $Z(\ \ _i; T)$ = The degree of the annihilator p_i

The matrix of the induced operator T_i in the ordered basis $\mathcal{B}_i$ is the companion matrix of the polynomial p_i .

i:e:; If $\mathcal{B}$ is the ordered basis for V .

Then $\mathcal{B}$ is the union of $\mathcal{B}_i$, namely $\mathcal{B}_1 \quad \mathcal{B}_2 \qquad \mathcal{B}_r$.

If A_i denote the $k_i \quad k_i$ companion matrix of p_i and if A is the matrix of T in the ordered basis B. Then

$$A = \begin{bmatrix} A_1 & 0 & & 0 \\ 0 & A_2 & 0 & \\ \cdot & & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & & A_r \end{bmatrix}$$

$$A = a_1 \quad A_2 \qquad A_r$$

An $n \quad n$ matrix A, which is the direct sum of companion matrices $A_1; A_2; \quad ; A_r$ of non-scalar monic polynomials $p_1; p_2; \quad ; p_r$ such that p_{i+1} divides p_i $(i = 1; 2; \ ; r \ 1)$, is said to be in Rational Form.

Theorem 9.5. let F be a field and let B be an $n \ n$ matrix over F. Then B is similar over the field F to one and only one matrix which is in rational form.

Proof. Let T be the linear opertor on F^n and let T be represented by the matrix B, in the standard ordered basis.

i:e:; There is some ordered basis for F^n, in which the linear operator T is represented by a matrix (say) A, which is in rational form.

Then B is similar to A.

Claim: B is similar to only one and ony matrix, which is in rational form.

If possible, let B be similar to another matrix C which is in the rational form over F.

i:e:; There is some ordered basis for F^n, in which the linear operator T is represented by the matrix C.

Thus, C is the direct sum of companion matrices c_i of monic polynomials $g_1; g_2; \quad ; g_s$ such that $g_{i+1}{=}g_i$ for $i = 1; 2; \quad ; s \quad 1$.

By using cyclic decomposition theorem, there exists non-zero vectors $_1; \quad ; \ _s$ in V with respective T-annihilators $g_1; \quad ; g_s$ such that

$$V = Z(\ _1; T) \qquad Z(\ _s; T)$$

Since $g_1; g_2; \quad ; g_s$ are the T-annihilator with respect to the matrix C.

Similarly, $p_1; p_2; \quad ; p_r$ are the T-annihilators with respect to the matrix A.

Thus, the uniqueness of the cyclic decomposition theorem implies that the polynomial g_i are identical with the polynomials p_i.

Hence $C = A$:

This completes the proof of the theorem.

Let us Sum Up:

In this unit, the students acquired knowledge to

explain the concept of cyclic decomposition theorem.

understand the concept of Rational Forms.

Suggested Readings:

1. M. Artin, Algebra , Prentice Hall of India Pvt. Ltd., 2005.
2. S.H. Friedberg, A.J. Insel and L.E Spence, Linear Algebra , 4th Edition, Prentice-Hall of India Pvt. Ltd., 2009.
3. I.N. Herstein, Topics in Algebra , 2nd Edition, Wiley Eastern Ltd, New Delhi, 2013.
4. J.J. Rotman, Advanced Modern Algebra , 2nd Edition, Graduate Studies in Mathematics, Vol. 114, AMS, Providence, Rhode Island, 2010.
5. G. Strang, Introduction to Linear Algebra , 2nd Edition, Prentice Hall of India Pvt. Ltd, 2013.

Block-V

UNIT-10

THE JORDAN FORM

Overview

In this unit, we shall discuss the Jordan form.

Objectives

After successful completion of this lesson, students will be able to

understand the concept of JordanForm.

10.1. The Jordan Form

Let N be a linear ooperator on a vector space V. We say that N is nilpotent, if there is some positve integer r such that $N^r = 0$:

Thus, there exists non-zero vectors $_1;\ _2;\ \ ;\ _r$ in V with N-annihilators $p_1; p_2;\ \ ; p_r$ such that

(i) $V = Z(\ _1; N)\qquad Z(\ _r; N)$ and

(ii) $p_{i+1}{=}p_i$ for $i = 1{:}2;\ \ ; r\ \ 1$

Since N is nilpotent and thus the minimal polynomial for N is x^k for some $k\ \ n$.

) Each N-annihilator p_i is of the form $p_i x^{k_i}$ where $k_1\ \ k_2\qquad k_r$:

The companion matrix of X^{k_i} is the $k_i\ \ k_i$ given as follows:

$$A_i = \begin{bmatrix} 0 & 0 & & 0 & 0 \\ 1 & 0 & & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ . & . & & . & . \\ 0 & 0 & & 1 & 0 \end{bmatrix}$$

These matrices A_i are nilpotent and theri size decreases as i increases.

One sees from this that associated with a nilpotent $n\ \ n$ matrix is a postive integer r and r positive integers $k_1;\ \ ; k_r$ such that $k_1 +\ \ + k_r = n$ and $k_i\ \ k_{i+1}$; and these positive integers determine the rational form of the matrix, $i{:}e{:};$ determine the matrix upto similarity.

Moreover, the positive integer r is the nullity of N.

Claim: Infact, the null space has a basis of r vectors $N^{k_1\ 1}\ _1; N^{k_2\ 1}\ _2;\ \ ; N^{k_r\ 1}\ _r$.

Let is in the null space of N.

) $N\ = 0$, then

$$= f_1\ _1 +\ \ + f_r\ _r$$

where f_i is a polynomial with $\deg f_i < k_i$:

$$\begin{aligned} i:e:;\ N(_1\ _1 + \quad + f_r\ _r) &= 0 \\ N(f_1\ _1) + \quad + N(f_r\ _r) &= 0 \\ \Big) \quad N(f_i\ _i) &= ;\quad 8_i \\ N(0) &= N\,(N(f_i\ _i)) \\ 0 &= N\,f_i(N)\ _i \\ &= (x\,f_i)\ _i \end{aligned}$$

Thus, $x\,f_i$ is divisible by x^{k_i} and since $\deg\,(f_i > k_i)$, this means that

$$f_i = c_i x^{k_i\ 1}$$

Where c_i is some scalar. But

$$= c_1(x^{k_1\ 1}\ _1) + \quad + c_r(x^{k_r\ 1}\ _r)$$

Which shows that the vectors $N^{k_1\ 1}\ _1; \quad ; N^{k_r\ 1}\ _r$ forms a basis for the null space of N.

This proves our claim.

Now, let T be a linear operator on V and assume that the characteristic polynomialf for T can be factorised over F as

$$f = (x \quad c_1)^{d_1} \quad (x \quad c_k)^{d_k}$$

where $c_1; c_2; \quad ; c_k$ are distinct elements of F and each $d_i \quad 1$:

Thus, the minimal polynomial for T will be of the form

$$p = (x \quad c_1)^{r_1} \quad (x \quad c_k)^{r_k}$$

where $1 \quad r_i \quad d_i$:

Let W_i be the null space of $(T \quad c_i I)^{r_i}$ and let T_i be the operator induced by T on W_i.

By using primary decomposition theorem, we have

$$V = W_1 \qquad W_k$$

and the minimal polynomial of each T_i is of the form $(x \quad c_i)^{r_i}$.

Let N_i be the linear operator on W_i definedby

$$N_i \;=\; T_i \quad c_iI$$

Then N_i is nilpotent and has minimal polynomial x^{r_i}.

Now, we choose a basis the subspace W_i corresponding to the cyclic decomposition for the nilpotent operator N_i.

Thus, the matrix of T_i in this ordered basis will be the direct sum of matrices of the form

$$\begin{bmatrix} c & 0 & \cdots & 0 & 0 \\ 1 & c & 0 & 0 & \cdots \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & \cdots & 1 & c \end{bmatrix}$$

eachnwith $c = c_i$. A matrix of this form is called an elementary Jordan matrix with characteristic value c.

Now, let us put all the bases for the W_i together and obtain are ordered basis for V . Now, let us describe the matrix of T; (i:e:;) A in this ordered basis, as follows:

The matrix A is the direct sum

$$A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_k \end{bmatrix}$$

of matrices $A_1; \quad ; A_k$. Where Each A_i is of the form

$$A_i = \begin{bmatrix} J_1^{(i)} & 0 & \cdots & 0 \\ 0 & J_2^{(i)} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & J_{n_i}^{(i)} \end{bmatrix}$$

where each $J_j^{(i)}$ is an elementary Jordanf matrix with characteristic value c_i .

An $n \quad n$ matrix A described as above, is said to be Jordan Form.

Let us Sum Up:

In this unit, the students acquired knowledge to explain the concept of Jordan Form.

Suggested Readings:

1. M. Artin, Algebra , Prentice Hall of India Pvt. Ltd., 2005.
2. S.H. Friedberg, A.J. Insel and L.E Spence, Linear Algebra , 4th Edition, Prentice-Hall of India Pvt. Ltd., 2009.
3. I.N. Herstein, Topics in Algebra , 2nd Edition, Wiley Eastern Ltd, New Delhi, 2013.
4. J.J. Rotman, Advanced Modern Algebra , 2nd Edition, Graduate Studies in Mathematics, Vol. 114, AMS, Providence, Rhode Island, 2010.
5. G. Strang, Introduction to Linear Algebra , 2nd Edition, Prentice Hall of India Pvt. Ltd, 2013.

zzzzz