

# MAS242 선형대수학 Notes

한승우

December 11, 2023

# CONTENTS

<b>CHAPTER</b>	<b>LINEAR EQUATIONS</b>	<b>PAGE 3</b>
<b>CHAPTER</b>	<b>VECTOR SPACES</b>	<b>PAGE 4</b>
	2.1 Bases and Dimension	4
<b>CHAPTER</b>	<b>LINEAR TRANSFORMATIONS</b>	<b>PAGE 7</b>
	3.1 Linear Transformations	7
	3.2 The Algebra of Linear Transformations	8
	3.3 Isomorphism	10
	3.4 Representation of Transformation by Matrices	10
	3.5 Linear Functionals	11
	3.6 The Double Dual	13
	3.7 The Transpose of a Linear Transformation	14
<b>CHAPTER</b>	<b>POLYNOMIALS</b>	<b>PAGE 15</b>
	4.1 Algebras	15
<b>CHAPTER</b>	<b>DETERMINANTS</b>	<b>PAGE 19</b>
	5.1 Determinant Functions	19
<b>CHAPTER</b>	<b>ELEMENTARY CANONICAL FORMS</b>	<b>PAGE 23</b>
	6.1 Eigenvalues	23
	6.2 Annihilating Polynomials	25
	6.3 Invariant Subspaces	26
	6.4 Simultaneous Triangulation and Diagonalization	28
	6.5 Direct-Sum Decompositions	30
	6.6 Invariant Direct Sums	31
	6.7 The Primary Decomposition Theorem	31

<b>CHAPTER</b>	<b>THE RATIONAL AND JORDAN FORMS</b>	<b>PAGE 34</b>
7.1	Cyclic Subspaces and Annihilators	34
7.2	Cyclic Decompositions and the Rational Form	35
7.3	The Jordan Form	40

<b>CHAPTER</b>	<b>INNER PRODUCT SPACES</b>	<b>PAGE 43</b>
8.1	Inner Products	43
8.2	Inner Product Space	44
8.3	Linear Functionals and Adjoint	48
8.4	Normal Operators	51

# **Chapter 1**

## **Linear Equations**

# Chapter 2

## Vector Spaces

### 2.1 Bases and Dimension

#### Theorem 2.1.1

Any subset that is linearly independent can be extended to a basis of  $V$ .

#### Lemma 2.1.1

If  $W$  is a subspace of  $V$  and  $W \subsetneq V$ , then  $\dim W < \dim V$  provided that  $V$  is finite-dimensional.

**Proof.** Let  $S_0$  be a basis of  $W$ .  $S_0$  is linearly independent, so we can enlarge it to get a basis of  $V$ .  $S' \triangleq S_0 \cup \{v_1, v_2, \dots, v_r\}$  is a basis of  $V$ .  $|S'| \geq |S_0| + 1$ ; otherwise  $\text{span } S_0 = V$ .  $\square$

#### Theorem 2.1.2 Inclusion/Exclusion Principle for Vector Spaces

If  $W_1$  and  $W_2$  are finite-dimensional subspaces of  $V$ , then  $W_1 + W_2$  is a finite-dimensional vector space and  $\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2)$ .

**Proof.** Let  $a \triangleq \dim W_1$ ,  $b \triangleq \dim W_2$ ,  $c \triangleq \dim(W_1 + W_2)$ , and  $d \triangleq \dim(W_1 \cap W_2)$ . Choose  $\{\alpha_1, \alpha_2, \dots, \alpha_d\}$  as a basis for  $W_1 \cap W_2$ . We may extend this into bases of  $W_1$  and  $W_2$ . Let  $\{\alpha_1, \dots, \alpha_d, \beta_{d+1}, \beta_{d+2}, \dots, \beta_a\}$  and  $\{\alpha_1, \dots, \alpha_d, \gamma_{d+1}, \gamma_{d+2}, \dots, \gamma_b\}$  be bases for  $W_1$  and  $W_2$  respectively.

We now claim that

$$B \triangleq \{\alpha_1, \dots, \alpha_d, \beta_{d+1}, \dots, \beta_a, \gamma_{d+1}, \dots, \gamma_b\}$$

is a basis of  $W_1 + W_2$ .

- Let  $x \in W_1 + W_2$ . Then,  $x = w_1 + w_2$  where  $w_i \in W_i$ . Since  $w_1 \in \text{span}\{\alpha_1, \dots, \alpha_d, \beta_{d+1}, \dots, \beta_a\}$  and  $w_2 \in \text{span}\{\alpha_1, \dots, \alpha_d, \gamma_{d+1}, \dots, \gamma_b\}$ , On the other hand,  $B \subseteq W_1 + W_2$ . Hence,  $\text{span } B = W_1 + W_2$ .
- Suppose we have  $\sum a_i \alpha_i + \sum b_j \beta_j + \sum c_k \gamma_k = 0$  for some  $a_i, b_j, c_k \in F$ . Rearranging the terms, we get  $\sum a_i \alpha_i + \sum b_j \beta_j = -\sum c_k \gamma_k$ , which implies that  $\sum c_k \gamma_k \in W_1 \cap W_2$ . The fact that  $\gamma_k$ 's are linearly independent from  $\{\alpha_i\}$  implies that  $c_k = 0$  for all  $k$ . Similarly,  $b_j = 0$  for all  $j$ . Hence, we are left with  $\sum a_i \alpha_i = 0$ , in which  $\alpha_i$ 's are linearly independent;  $a_i = 0$ . Hence,  $B$  is linearly independent.

Therefore,  $\dim(W_1 + W_2) = a + b - d$ .  $\square$

**Definition 2.1.1: Ordered Basis**

Let  $V$  be a finite-dimensional vector space over  $F$ . An *ordered basis* of  $V$  is a sequence of vectors that forms a basis.

**Note:-**

Usually, we emphasize the ordered basis with semicolons like  $\{\beta_1; \beta_2\}$ .

**Lemma 2.1.2**

Let  $V$  be a finite-dimensional vector space over  $F$ . Suppose  $B = \{v_1; v_2; \dots; v_n\}$  is an ordered basis of  $V$ . Then, for each  $x \in V$ , there uniquely exists an expression of the form

$$x = x_1 v_1 + x_2 v_2 + \dots + x_n v_n$$

for some  $x_i \in F$ .

**Proof.** The existence of the form is obvious since  $x \in V = \text{span } B$ .

(Uniqueness) Suppose we have two such expressions:

$$x = \sum x_i v_i = \sum y_i v_i$$

where  $x_i, y_i \in F$ . Then, we have  $\sum (x_i - y_i) v_i = 0$ . The linear independence of  $B$  gives that  $x_i - y_i = 0$  for all  $i$ . Hence,  $x_i = y_i$ .  $\square$

**Definition 2.1.2: Coordinate Matrix**

Let  $V$  be a finite-dimensional vector space over  $F$ . Let  $B$  be an ordered basis of  $V$ . Let  $x \in V$  and write it as  $x = \sum_{i=1}^n x_i v_i$  with  $x_i \in F$ ,  $v_i \in B$ . Define

$$[x]_B \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

be the *coordinate matrix* of  $x$  with respect to the basis  $B$

**Theorem 2.1.3**

Let  $V$  be a finite-dimensional vector space over  $F$ . Let  $B$  and  $B'$  be two ordered bases of  $V$ . Then, there uniquely exists an invertible matrix  $P$  such that  $\forall x \in V$ ,  $[x]_B = P[x]_{B'}$  and  $[x]_{B'} = P^{-1}[x]_B$ .

**Proof.** Let  $B \triangleq \{\alpha_1; \dots; \alpha_n\}$  and  $B' \triangleq \{\alpha'_1; \dots; \alpha'_n\}$ . For  $\alpha'_j \in B'$ , since  $B$  is a basis, there are unique  $P_{ij} \in F$  ( $i \in [n]$ ) such that  $\alpha'_j = \sum_{i=1}^n P_{ij} \alpha_i$ .

Let  $x \in V$ . Write  $[x]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  and  $[x]_{B'} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ . Then,  $x = \sum_{j=1}^n x'_j \alpha'_j = \sum_{j=1}^n (\sum_{i=1}^n x'_j P_{ij}) \alpha_i$ .

By the uniqueness, we have  $x_i = \sum_{j=1}^n x'_j P_{ij}$  for each  $i$ . In other words, we have

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} P_{11} & \dots & P_{1n} \\ \vdots & \dots & \vdots \\ P_{n1} & \dots & P_{nn} \end{bmatrix} \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix}$$

Since  $B$  and  $B'$  are linearly independent,  $x = 0 \iff [x]_B = 0 \iff [x]_{B'} = 0$ . Hence,  $P$  is invertible.  $\square$

# Chapter 3

## Linear Transformations

### 3.1 Linear Transformations

#### Definition 3.1.1: Linear Transformation

Let  $V_1$  and  $V_2$  be vector spaces over  $F$ .  $T: V_1 \rightarrow V_2$  is said to be a *linear transformation* if

- $\forall x_1, x_2 \in V_1, T(x_1 + x_2) = T(x_1) + T(x_2)$
- $\forall x \in V_1, \forall c \in F, T(cx) = cT(x)$ .

#### Theorem 3.1.1

Let  $V$  and  $W$  be finite-dimensional vector spaces over  $F$ . where  $\{\alpha_1, \dots, \alpha_n\}$  is a basis of  $V$ . Let  $\{\beta_1, \dots, \beta_n\}$  be any given set of vectors of  $W$ . Then, there exists a unique transformation  $T: V \rightarrow W$  such that  $T(\alpha_i) = \beta_i$ .

**Proof.** Let  $T_0: V \rightarrow W$  be defined by

$$T_0\left(\sum_{i=1}^n x_i \alpha_i\right) = \sum_{i=1}^n x_i \beta_i.$$

This is a linear transformation indeed.

(Uniqueness) If there is another such  $U: V \rightarrow W$ , Then,  $U\left(\sum_{i=1}^n x_i \alpha_i\right) = \sum_{i=1}^n x_i U(\alpha_i)$ . Hence,  $U = T_0$ .  $\square$

#### Definition 3.1.2: Null Space and Range Space

Let  $T: V \rightarrow W$  be a linear transformation between vector spaces over  $F$ .

- $\text{null } T \triangleq \ker T \triangleq \{v \in V \mid T(v) = 0\}$
- $\text{range } T \triangleq \text{Im } T \triangleq \{w \in W \mid \exists v \in V, w = T(v)\}$

#### Note:-

$\ker T$  and  $\text{Im } T$  are subspaces of  $V$  and  $W$  respectively.

#### Definition 3.1.3

Let  $T: V \rightarrow W$  be a linear transformation between vector spaces over  $F$ .

$$\text{nullity}(T) \triangleq \dim \ker(T) \quad \text{and} \quad \text{rank}(T) \triangleq \dim \text{Im}(T)$$



### Theorem 3.1.2 Rank-Nullity Theorem

Let  $T: V \rightarrow W$  be a linear transformation between vector spaces over  $F$ . Then,  $\text{rank}(T) + \text{nullity}(T) = \dim V$ .

**Proof.** Let  $\{v_1, \dots, v_k\}$  be a basis for  $\ker T$  where  $k = \text{nullity } T$ . Choose  $v_{k+1}, \dots, v_n \in V$  such that  $\{v_i\}_{i=1}^n$  is a basis of  $V$ . We claim that  $\{T(v_{k+1}), \dots, T(v_n)\}$  is a basis of  $\text{Im } T$ .

Suppose  $\sum_{i=k+1}^n c_i T(v_i) = 0$  for some  $c_i \in F$ . Then, we have  $T(\sum_{i=k+1}^n c_i v_i) = 0$ ; hence  $\sum_{i=k+1}^n c_i v_i \in \ker T$ . Since  $\{v_1, \dots, v_k\}$  is a basis of  $\ker T$ , we have  $\sum_{i=k+1}^n c_i v_i = \sum_{i=1}^k a_i v_i$  for some  $a_i$ 's. Therefore, since  $\{v_1, \dots, v_n\}$  is linearly independent, all  $c_i$ 's and  $a_i$ 's are zero. This implies that  $\{T(v_i)\}_{i=k+1}^n$  is linearly independent.

Take any  $T(v) \in \text{Im } T$ . Then,  $v = \sum_{i=1}^n c_i v_i$  for some  $c_i \in F$ . Then,  $T(v) = \sum_{i=k+1}^n c_i T(v_i)$ . Hence,  $\text{Im } T \subseteq \text{span}\{T(v_{k+1}), \dots, T(v_n)\}$ .

The two paragraphs imply that  $\text{rank } T = n - k$ . □

### Theorem 3.1.3

Let  $A$  be a  $m \times n$  matrix. Then  $\dim \text{span}(\text{rows}) = \dim \text{span}(\text{columns})$ .

**Proof.**  $V = F^n$ ,  $W = F^m$ . Then,  $\dim \text{span}(\text{columns}) = \dim \text{Im } T = \text{rank } T$ , so  $\text{nullity } T = n - \text{rank } T = n - \text{colrank } T$ .

The number of rows with leading one's in  $\text{rref } A$  equals the dimension of the row space of  $A$ , which is simply the number of columns with the leading ones. It is equal to the dimension of the column space. Hence,  $\text{nullity } T = n - \text{colrank } T$  □

## 3.2 The Algebra of Linear Transformations

### Definition 3.2.1

Let  $T: V \rightarrow W$  be a linear transformation between vector spaces over  $F$ .  $L(V, W) \triangleq \{T: V \rightarrow W \mid T \text{ is a linear transformation}\}$

### Theorem 3.2.1

Let  $T: V \rightarrow W$  be a linear transformation between vector spaces over  $F$ . Then,  $L(V, W)$  is a vector space over  $F$  under usual addition and multiplication.

### Theorem 3.2.2

Let  $V$  and  $W$  be  $n$ - and  $m$ -dimensional vector spaces over  $F$ , respectively. Then,  $\dim L(V, W) = mn$ .

**Proof.** Let  $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$  and  $\mathcal{B}' = \{\beta_1, \dots, \beta_m\}$  be bases for  $V$  and  $W$ , respectively. For each  $p \in [n]$  and  $q \in [m]$ , Define

$$E^{p,q}(\alpha_i) = \begin{cases} 0 & \text{if } i \neq p \\ \beta_q & \text{if } i = p \end{cases}.$$

Then,

- These  $E^{p,q}$  are linear transformations
- These are linearly independent.

- They span  $L(V, W)$ .

□

### Lemma 3.2.1

Let  $T: V \rightarrow W$  and  $U: W \rightarrow Z$  be linear transformations between vector spaces over  $F$ . Then,  $U \circ T \in L(V, Z)$ .

### Definition 3.2.2: Linear Operator (Endomorphism)

Let  $T: V \rightarrow V$  be a linear transformation from a vector space  $V$  to itself. Then,  $T$  is called a *linear operator*. (Or an *endomorphism*.)

#### Note:-

For each  $T, U \in L(V, V)$ ,  $T \circ U \in L(V, V)$ .  $(T_1 + T_2) \circ U = T_1 \circ U + T_2 \circ U$ . And many more...  $(L(V, V), +, \circ)$  is a non-commutative ring.

### Definition 3.2.3: Injectivity and Surjectivity

A linear transform  $T: V \rightarrow W$  is

- *injective* (or, nonsingular) if  $T(v) = 0 \implies v = 0$ .
- *surjective* if  $T(V) = W$ .
- *invertible* if  $\exists$  linear transform  $U: W \rightarrow V$ ,  $U \circ T = \text{id}_V \wedge T \circ U = \text{id}_W$ .

### Exercise 3.2.1

$T: V \rightarrow W$  is injective and surjective if and only if  $T$  is invertible.

### Exercise 3.2.2

If  $T: V \rightarrow W$  is a nonsingular linear transformation, then, for any linearly independent subset  $S \subseteq V$ ,  $T(S)$  is linearly independent.

### Exercise 3.2.3

Suppose  $V$  and  $W$  are finite-dimensional vector spaces. If  $T: V \rightarrow W$  is invertible, then  $\dim V = \dim W$ .

### Theorem 3.2.3

Let  $V$  and  $W$  be finite-dimensional vector spaces over  $F$  with  $\dim V = \dim W$ . Let  $T: V \rightarrow W$  be a linear transform. TFAE

- $T$  is invertible.
- $T$  is injective.
- $T$  is surjective.

**Proof.**  $T$  is injective  $\iff$  nullity  $T = 0 \iff$  rank  $T = n \iff \text{Im } T = W \iff T$  is onto □

### Definition 3.2.4: General Linear Group

Let  $\text{GL}(V) \triangleq \{ T \in L(V, V) \mid T \text{ is invertible} \}$ . Then,  $(\text{GL}(V), \circ)$  is called the *general linear group* of  $V$ .

**Note:-**

The general linear group is actually a group.

### 3.3 Isomorphism

#### Definition 3.3.1: Isomorphism

Let  $V$  and  $W$  be vector spaces over  $F$ . We say that a linear transformation  $T: V \rightarrow W$  is an *isomorphism* if  $T$  is an invertible linear transformation.

We say  $V$  and  $W$  are *isomorphic* if there exists an isomorphism  $T: V \rightarrow W$ , if  $V$  and  $W$  are isomorphic, then we write  $V \simeq W$ .

#### Theorem 3.3.1

Let  $V$  be a vector spaces over  $F$  of dimension  $n$ . Then,  $V \simeq F^n$ .

**Proof.** Let  $B = \{\alpha_1; \dots; \alpha_n\}$  be a basis of  $V$ . Define  $T: V \rightarrow F^n$  by  $v \mapsto [v]_B$ .

Suppose  $T(v) = 0$ . Then,  $v = 0 \cdot \alpha_1 + \dots + 0 \cdot \alpha_n = 0$ . Hence,  $T$  is injective. By Theorem 3.2.3,  $T$  is isomorphism.  $\square$

### 3.4 Representation of Transformation by Matrices

#### Theorem 3.4.1

Let  $V$  and  $W$  be vector spaces over  $F$  with  $\dim V = n$  and  $\dim W = m$ . Let  $B$  and  $B'$  be bases of  $V$  and  $W$ , respectively. If  $T: V \rightarrow W$  is a linear transformation, then there uniquely exists  $m \times n$  matrix  $A$  such that  $[T(v)]_{B'} = A[v]_B$ . We write  $[T]_{B,B'} \triangleq A$ .

**Proof.**  $A = \begin{bmatrix} [T(v_1)]_{B'} & [T(v_2)]_{B'} & \dots & [T(v_n)]_{B'} \end{bmatrix}$  where  $v_i$  is the  $i^{\text{th}}$  basis vector of  $B$ .  $\square$

#### Theorem 3.4.2

Let  $V \xrightarrow{T} W \xrightarrow{U} Z$  be linear transformations. Let  $A_1 = [T]_{B,B'}$  and  $A_2 = [U]_{B'',B'}$ . Then,  $[U \circ T]_{B,B''} = A_2 A_1$ .

#### Theorem 3.4.3

Let  $V$  be finite-dimensional vector space over  $F$  with two (possibly different) bases  $B_1$  and  $B_2$ . Let  $T \in L(V, V)$ . Let  $P$  be the matrix such that  $[v]_{B_1} = P[v]_{B_2}$ . Then,  $[T]_{B_1} \triangleq [T]_{B_1, B_1}$  are related by

$$[T]_{B_2} = P^{-1}[T]_{B_1}P.$$

#### Definition 3.4.1: Similar Matrices

Suppose  $M$  and  $N$  are  $n \times n$  matrices.  $M$  and  $N$  are *similar* if there exists an invertible  $P$  such that  $N = P^{-1}MP$ .

**Proof.**  $[T(v)]_{B_1} = [T]_{B_1}[v]_{B_1} = [T]_{B_1}P[v]_{B_2}$ .  $[T(v)]_{B_1} = P[T(v)]_{B_2} = P[T]_{B_2}[v]_{B_2}$ .

Since  $v$  was arbitrary,  $P[T]_{B_2} = [T]_{B_1}P$ .  $\square$

**Note:-**

- A linear transformation  $T: V \rightarrow V$  gives varying matrices  $[T]_B$  that are all similar when the basis  $B$  is changed.
- On linear operators, we will have various definitions.
- Characteristic (eigen) polynomial has  $(-1)^{\deg}(\text{constant term})$  as  $\det T$  and  $-(n - 1 \text{ deg term})$  as  $\text{tr } T$ .

### 3.5 Linear Functionals

**Definition 3.5.1: Linear Functional**

Let  $V$  be a vector space over  $F$ . A linear transformation  $T: V \rightarrow F$  is called a (*linear*) *functional*.

**Definition 3.5.2: Dual Vector Space**

Let  $V$  be a vector space over  $F$ . We normally write  $V^* \triangleq L(V, F)$  and call it the *dual vector space* of  $V$ .

**Note:-**

By Theorem 3.2.2, we know that  $\dim V^* = \dim V$  if  $V$  is a finite-dimensional vector space.

**Lemma 3.5.1**

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $n = \dim V$ . Let  $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$  be a basis for  $V$ . Define  $f_i \in V^*$  by declaring  $f_i(\alpha_j) = \delta_{ij}$ . Then,  $\{f_1, \dots, f_n\}$  is a basis for  $V^*$ .

**Proof.** Since  $\dim V^* = \dim V = n$ , we only need to show that the set is linearly independent.

Suppose  $\sum_{i=1}^n c_i f_i = 0$  for some  $c_i \in F$ . Then, for each  $j \in [n]$ , as  $f_i(\alpha_j) = \delta_{ij}$ ,  $0 = (\sum_{i=1}^n c_i f_i)(\alpha_j) = c_j f_j(\alpha_j) = c_j$ . Hence, they are linearly independent.  $\square$

**Definition 3.5.3: Dual Basis**

The set  $\{f_1, f_2, \dots, f_n\} \subseteq V^*$  in Lemma 3.5.1 is called the *dual basis* of the basis  $\{\alpha_1, \dots, \alpha_n\}$  for  $V$ .

**Lemma 3.5.2**

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $n = \dim V$ . Let  $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$  be a basis for  $V$ . Let  $\{f_1, \dots, f_n\} \subseteq V^*$  be the dual basis of it.

- For each  $f \in V^*$ ,  $f = \sum_{i=1}^n f(\alpha_i) f_i$ .
- For each  $v \in V$ ,  $v = \sum_{i=1}^n f_i(v) \alpha_i$ .

**Proof.**

- There exists  $x_i \in F$  such that  $f = \sum_{i=1}^n x_i f_i$ . Evaluating at  $\alpha_j$  for each  $j \in [n]$ , we get  $f(\alpha_j) = x_j$ .
- There exists  $y_i \in F$  such that  $v = \sum_{i=1}^n y_i \alpha_i$ . Applying  $f_j$  for each  $j \in [n]$ , we get  $f_j(v) = y_j$ .

$\square$

**Definition 3.5.4: Hyperspace**

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $n = \dim V$ . A subspace  $W$  of  $V$  which has the dimension  $n - 1$  is called a *hyperspace* in  $V$ .

**Example 3.5.1**

If  $f : V \rightarrow F$  is a nonzero functional, then  $\ker f$  is an example of a hyperspace in  $V$ .

**Definition 3.5.5: Annihilator**

Let  $V$  be a finite-dimensional vector space over  $F$  with dimension  $n$ . Let  $\emptyset \subsetneq S \subseteq V$ . The *annihilator* of  $S$ ,  $S^\circ = \text{Ann } S$  is defined to be

$$S^\circ = \{f \in V^* \mid \forall \alpha \in S, f(\alpha) = 0\}.$$

**Note:-**

- $S^\circ$  is a subspace of  $V^*$
- $\text{Ann } \{0\} = V^*$ .
- $\text{Ann } V = \{0\}$ .

**Theorem 3.5.1**

Let  $V$  be a finite-dimensional vector space over  $F$  with dimension  $n$ . Let  $W$  be a subspace of  $V$ . Then,

$$\dim W + \dim W^\circ = \dim V.$$

**Proof.** Let  $k \triangleq \dim W$  and  $\{\alpha_1, \dots, \alpha_k\} \subseteq W$  be a basis for  $W$ . We may extend it to the basis for  $V$  so that  $\{\alpha_1, \dots, \alpha_k, \alpha_{k+1}, \dots, \alpha_n\}$  is a basis for  $V$ . Let  $\{f_1, \dots, f_k, f_{k+1}, \dots, f_n\}$  be the dual basis of  $\{\alpha_1, \dots, \alpha_n\}$ .

For each  $i \in \{k+1, \dots, n\}$ , by the construction of the dual basis,  $f_i(\alpha_j) = 0$  for each  $j \in [k]$ . Hence,  $f_{k+1}, \dots, f_n \in W^\circ$ .

Take any  $f \in W^\circ$ . Then,  $f = \sum_{i=1}^n f(\alpha_i)f_i$ . For each  $i \in [k]$ ,  $f(\alpha_i) = 0$ . Hence,  $f = \sum_{i=k+1}^n f(\alpha_i)f_i$ ;  $\{f_{k+1}, \dots, f_n\}$  spans  $W^\circ$ . Therefore,  $\{f_{k+1}, \dots, f_n\}$  is a basis for  $W^\circ$ .  $\square$

**Corollary 3.5.1**

Let  $V$  be a finite-dimensional vector space over  $F$  with dimension  $n$ . Let  $W$  be a  $k$ -dimensional subspace of  $V$ . Then,  $W$  is the intersection of  $n - k$  hyperspaces in  $V$  of the form  $\ker f_i$  for some  $f_i \in V^* \setminus \{0\}$ .

**Proof.** Let  $\{\alpha_1, \dots, \alpha_k\}$  be a basis for  $W$  and extend it to  $\{\alpha_1, \dots, \alpha_n\}$  so that it becomes a basis for  $V$ . Let  $\{f_1, \dots, f_n\} \subseteq V^*$  be the dual basis of  $\{\alpha_1, \dots, \alpha_n\}$ . Then,  $W = \bigcap_{i=k+1}^n \ker f_i$ .  $\square$

**Corollary 3.5.2**

Let  $V$  be a finite-dimensional vector space over  $F$  with dimension  $n$ . Let  $W$  be a hyperspace in  $V$ . Then,  $W = \ker f$  for some  $f \in V^* \setminus \{0\}$ .

## 3.6 The Double Dual

**Note:-**

Take  $\alpha \in V$ . Let us define  $L_\alpha \in V^{**}$  as follows:

$$\begin{aligned} L_\alpha : V^* &\longrightarrow F \\ f &\longmapsto f(\alpha). \end{aligned}$$

Then, define  $\mathcal{L}$  by

$$\begin{aligned} \mathcal{L} : V &\longrightarrow V^{**} \\ \alpha &\longmapsto L_\alpha. \end{aligned}$$

Then,  $\mathcal{L}$  is an injective linear transformation.

### Theorem 3.6.1

Let  $V$  be a finite-dimensional vector space over  $F$  with dimension  $n$ . Then,  $\mathcal{L} : V \rightarrow V^{**}$  is an isomorphism of vector spaces.

**Proof.** We have  $\dim V = \dim V^* = \dim V^{**} = n$  by Theorem 3.2.2. The result follows from Theorem 3.2.3.  $\square$

### Definition 3.6.1: Proper Subspace

Let  $V$  be a vector space over  $F$ . Then, a subspace  $W$  of  $V$  is a *proper subspace* of  $V$  if  $W \subsetneq V$ .

### Definition 3.6.2: Maximal Subspace

A proper subspace  $W$  of  $V$  is said to be *maximal* if, there exists no subspace  $Z$  of  $V$  such that  $W \subsetneq Z \subsetneq V$ .

### Definition 3.6.3: Hyperspace

Let  $V$  be a vector space over  $F$ . A maximal proper subspace  $W$  of  $V$  is called a *hyperspace* in  $V$ .

**Note:-**

In case of  $\dim V = n$ , a proper maximal subspace of  $V$  is of dimension  $n - 1$ .

### Theorem 3.6.2

Let  $V$  be a vector space over  $F$ . Let  $f \in V^* \setminus \{0\}$ . Then,  $\ker f$  is a hyperspace in  $V$ .

**Proof.**  $\ker f$  is proper, since, otherwise,  $f = 0$ .

It is enough to show that, for each  $\alpha \in V \setminus \ker f$ ,  $\text{span}\{\ker f, \alpha\} = V$ . Take any  $\beta \in V$ . Let  $\alpha \in V \setminus \ker f$ . Define  $c \triangleq f(\alpha)^{-1}f(\beta)$  and  $\gamma \triangleq \beta - c\alpha$ . Then,  $f(\gamma) = f(\beta) - cf(\alpha) = 0$ ;  $\gamma \in \ker f$ . Hence,  $\beta = \gamma + c\alpha \in \text{span}[\ker f, \alpha]$ .  $\square$

### Theorem 3.6.3

Let  $V$  be a vector space over  $F$ . Let  $W$  be a hyperspace in  $V$ . Then, there exists  $f \in$

$V^* \setminus \{0\}$  such that  $W = \ker f$ .

**Proof.** There exists  $\alpha \in V \setminus W$  such that  $\text{span}\{W, \alpha\} = V$ . Hence, every  $\beta \in V$  can be written as  $\beta = \gamma + c\alpha$  where  $\gamma \in W$  and  $c \in F$ . Note that  $\gamma$  and  $c$  are uniquely determined by  $\beta$ .

Define  $g: V \rightarrow F$  by  $g(\beta) = c$ . Then,  $g$  is a linear functional, and  $\ker g = W$  by definition.  $\square$

**Note:-**

Theorem 3.6.2 and Theorem 3.6.3 together imply that the set of hyperspaces in  $V$  and the set of null spaces of functionals have a one-to-one correspondence.

## 3.7 The Transpose of a Linear Transformation

### Definition 3.7.1: Transpose

Let  $T: V \rightarrow W$  be a linear transformation. The map  $T^t: W^* \rightarrow V^*$  defined by  $g \mapsto g \circ T$  is called the *transpose* of  $T$ .

### Lemma 3.7.1

Let  $T: V \rightarrow W$  be a linear transformation. Then,  $T^t$  is a linear transformation.

### Theorem 3.7.1

Let  $T: V \rightarrow W$  be a linear transformation between finite-dimensional vector spaces over  $F$ . Fix ordered bases  $\mathcal{B}$  and  $\mathcal{B}'$  for  $V$  and  $W$ , respectively. Let  $\mathcal{B}^*$  and  $\mathcal{B}'^*$  be their dual bases. Let  $A \triangleq [T]_{\mathcal{B}, \mathcal{B}'}$  and  $A' \triangleq [T^t]_{\mathcal{B}'^*, \mathcal{B}^*}$ . Then,  $a_{ij} = a'_{ji}$ .

**Proof.** Let  $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ ,  $\mathcal{B}' = \{\beta_1, \dots, \beta_m\}$ ,  $\mathcal{B}^* = \{f_1, \dots, f_n\}$ , and  $\mathcal{B}'^* = \{g_1, \dots, g_m\}$ . Then, we have  $T\alpha_j = \sum_{i=1}^m a_{ij}\beta_i$  for each  $j \in [n]$  and  $T^tg_j = \sum_{i=1}^n b_{ij}f_i$  for each  $j \in [m]$ .

For each  $i \in [n]$  and  $j \in [m]$ ,  $(T^tg_j)(\alpha_i) = g_j(T\alpha_i) = g_j\left(\sum_{k=1}^m a_{ki}\beta_k\right) = \sum_{k=1}^m a_{ki}g_j(\beta_k) = a_{ji}$ . Hence, since  $T^tg_j$  is a linear functional on  $V$ ,  $T^tg_j = \sum_{i=1}^n (T^tg_j)(\alpha_i)f_i = \sum_{i=1}^n a_{ji}f_i$ . Therefore,  $a_{ij} = b_{ji}$  for each  $i \in [n]$  and  $j \in [m]$ .  $\square$

### Theorem 3.7.2

Let  $T: V \rightarrow W$  be a linear transformation.

- (i)  $\ker T^t = (\text{Im } T)^\circ$ .
- (ii) If  $V$  and  $W$  are finite-dimensional, then  $\text{rank } T^t = \text{rank } T$ .
- (iii) If  $V$  and  $W$  are finite-dimensional, then  $\text{Im } T^t = (\ker T)^\circ$ .

**Proof.**

- (i)  $g \in \ker T^t \iff T^t(g) = 0 \iff g \circ T = 0 \iff g \in (\text{Im } T)^\circ$
- (ii) Let  $n \triangleq \dim V$  and  $m \triangleq \dim W$ . Let  $r = \text{rank } T$ . Then, by Theorem 3.5.1,  $\dim(\text{Im } T)^\circ = m - r$ . By (i),  $(\text{Im } T)^\circ = \ker T^t$ ; hence  $\text{nullity } T^t = m - r$ . By the rank-nullity theorem,  $\text{rank } T^t = r = \text{rank } T$ .
- (iii) Take any  $f \in \text{Im } T^t$ . Then, there exists  $g \in W^*$  such that  $f = g \circ T$ . Then, for any  $\alpha \in \ker T$ ,  $f(\alpha) = g(T(\alpha)) = 0$ . Hence,  $f \in (\ker T)^\circ$ ;  $\text{Im } T^t \subseteq (\ker T)^\circ$ . But since the two spaces have the same dimension, it must be the equality to hold.  $\square$

# Chapter 4

## Polynomials

### 4.1 Algebras

#### Definition 4.1.1: $F$ -algebra

Let  $F$  be a field. A vector space  $\mathcal{A}$  with a map  $\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$  such that

- (i)  $\forall \alpha, \beta, \gamma \in \mathcal{A}, \alpha(\beta\gamma) = (\alpha\beta)\gamma$
- (ii)  $\forall \alpha, \beta, \gamma \in \mathcal{A}, \alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$  and  $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$
- (iii)  $\forall c \in F, \forall \alpha, \beta \in \mathcal{A}, c(\alpha\beta) = (c\alpha)\beta = \alpha(c\beta)$

is called a  $F$ -algebra or a linear algebra over  $F$ .

- If there is an element  $1$  in  $\mathcal{A}$  such that  $1\alpha = \alpha 1 = \alpha$  for each  $\alpha \in \mathcal{A}$ , then we call  $\mathcal{A}$  a  $F$ -algebra with identity.
- The algebra  $\mathcal{A}$  is called *commutative* if  $\alpha\beta = \beta\alpha$  for all  $\alpha, \beta \in \mathcal{A}$ .

#### Definition 4.1.2: Polynomial

Let  $F[x]$  be the subspace of  $F^\omega$  spanned by the vectors  $1, x, x^2, \dots$ . An element of  $F[x]$  is called a *polynomial over  $F$* .

#### Definition 4.1.3: Degree

For each  $f \in F[x] \setminus \{0\}$ ,  $\deg f \triangleq \max\{k \in \mathbb{N} \cup \{0\} \mid f_k \neq 0\}$ .

#### Theorem 4.1.1

Let  $f, g \in F[x] \setminus \{0\}$ .

- (i)  $fg \neq 0$
- (ii)  $\deg(fg) = \deg f + \deg g$
- (iii)  $fg$  is monic if  $f$  and  $g$  are monic.
- (iv)  $fg$  is scalar polynomial if  $f$  and  $g$  are scalar polynomials.
- (v) If  $f + g \neq 0$ , then  $\deg(f + g) \leq \max\{\deg f, \deg g\}$ .

#### Theorem 4.1.2 Euclidean Algorithm

Let  $f, g \in F[x]$  and  $g \neq 0$ . Then, there uniquely exists  $q, r \in F[x]$  such that

- $f = gq + r$  and
- either  $r = 0$  or  $\deg r < \deg g$ .



**Definition 4.1.4: Divisibility**

Let  $f, g \in F[x]$ . If  $f = gq$  for some  $q \in F[x]$ , then we write  $g \mid f$ .

**Lemma 4.1.1**

Let  $f \in F[x] \setminus \{0\}$  and  $c \in F$ . Then,  $(x - c) \mid f \iff f(c) = 0$ .

**Proof.** There exists  $q, r \in F[x]$  such that  $f = (x - c)q + r$  with either  $r = 0$  or  $\deg r = 0$ . Note that  $f(c) = r$ . Hence,  $f(c) = 0 \iff (x - c) \mid f$ .  $\square$

**Definition 4.1.5: Evaluation**

Let  $F$  be a field. Let  $\alpha \in F$  be fixed. Then, the function  $\text{ev}_\alpha: F[x] \rightarrow F$  defined by  $f \mapsto f(\alpha)$  is called the *evaluation of  $\alpha$  in  $f(x)$* .

**Definition 4.1.6: Ideal**

An ideal  $M \subseteq F[x]$  is an  $F$ -subspace if for every  $f \in F[x]$  and  $g \in M$ , we have  $fg \in M$ .

**Definition 4.1.7: Principal Ideal**

An ideal of the form

$$M = \{g_0 h \mid h \in F[x]\} = (g_0)$$

for a fixed  $g_0$  is called a *principal ideal*.

**Theorem 4.1.3**

Let  $F$  be a field. Let  $M \subseteq F[x]$  be a nonzero ideal. Then,  $M$  is a principal ideal given by a monic polynomial in  $F[x]$ .

**Proof.**  $M$  does contain nonzero polynomials. Hence, we may let  $g_0 \in \arg\min_{g \in M \setminus \{0\}} \deg g$  by the well-orderedness of natural numbers. WLOG,  $g_0$  is monic.

We shall claim that  $M = (g_0)$ . Take any  $f \in M$ . By the Euclidean algorithm,  $\exists q, r \in F[x]$ ,  $f = g_0 q + r$  with either  $r = 0$  or  $\deg r < \deg g_0$ . If  $r \neq 0$ , then  $r = f - g_0 q \in M$  but  $\deg r < \deg g_0$ , which contradicts the minimality of  $\deg g_0$ . Hence,  $r = 0$ , and thus  $f = g_0 q \in (g_0)$ .  $\square$

**Note:-**

By putting “monic” assumption, such  $g_0$  is unique as well.

**Corollary 4.1.1**

Let  $p_1, \dots, p_n \in F[x]$  be a finite number of polynomials where not all of them are zero. Then, there uniquely exists monic  $g_0 \in F[x]$  such that

- (i)  $p_1 F[x] + p_2 F[x] + \dots + p_n F[x] = (g_0)$
- (ii)  $\forall i \in [n], g_0 \mid p_i$
- (iii)  $(\forall i \in [n], f \mid p_i) \implies f \mid g_0$

Such  $g_0$  is called the *greatest common divisor* of  $p_1, \dots, p_n$ . Sometimes this is denoted by  $(p_1, \dots, p_n) = (g_0)$ .

**Proof.**  $p_1F[x] + p_2F[x] + \cdots + p_nF[x]$  is an ideal. By Theorem 4.1.3, there uniquely exists monic  $g_0$  that generates it. (ii) directly follows from (i).  $g_0 = \sum_{j=1}^n p_j g_j = f \sum_{j=1}^n h_j g_j$ .  $\square$

#### Definition 4.1.8: Relatively Prime

Let  $p_1, \dots, p_n$  be nonzero polynomials. We say that they are *relatively prime* if  $(p_1, \dots, p_n) = (1)$ .

#### Definition 4.1.9: Reducibility

Let  $F$  be a field. We say  $f \in F[x] \setminus \{0\}$  is *reducible* if  $f = gh$  for some  $g, h \in F[x]$  with  $\deg g, \deg h \geq 1$ . If  $f$  is not reducible, we say  $f$  is *irreducible*.

#### Definition 4.1.10: Prime Element

Let  $F$  be a field. We say that  $f \in F[x]$  is a *prime element* if, for every  $g, h \in F[x]$ ,  $f \mid gh \implies (f \mid g \vee f \mid h)$ .

#### Example 4.1.1

- Let  $F$  be a field. Then any polynomial over  $F$  with degree one is irreducible.
- $F = \mathbb{R}$ .  $f(x) = x^2 + ax + b$  is irreducible iff  $D < 0$ .
- $F = \mathbb{F}_p = \mathbb{Z}/p$ . There are quite many irreducible polynomial of degree  $d$ .

#### Theorem 4.1.4

Let  $p \in F[x] \setminus \{0\}$  be a polynomial. Then,  $p$  is irreducible if and only if  $p$  is prime.

**Proof.**

( $\implies$ ) Suppose  $p \mid gh$  for some  $g, h \in F[x]$ . If  $g$  or  $h$  is zero, then it is done. Hence, WMA that  $g, h \neq 0$ . Let  $(p, g) = (d)$ .  $d \mid p$  implies that  $d = 1$  or  $d = p$  since  $p$  is irreducible. If  $d = p$ , then  $d \mid g$ , i.e.,  $p \mid g$ . If  $d = 1$ , then there exists  $p_0, g_0$  such that  $pp_0 + gg_0 = 1$ . Hence,  $php_0 + ghg_0 = h$ . Hence,  $p \mid h$ .

( $\impliedby$ ) Suppose  $p$  is reducible. Then,  $p = gh$  for some  $g, h$  with nonzero degrees. Since  $p$  is prime,  $p \mid g$  or  $p \mid h$ . This implies  $\deg p \leq \deg g$  or  $\deg p \leq \deg h$ . This is a contradiction since  $\deg p = \deg g + \deg h \leq 2 \deg p$  arises.  $\square$

#### Theorem 4.1.5 Unique Factorization of Polynomials

Let  $F$  be a field. Every non-constant polynomial  $f \in F[x]$  factors into a product of irreducible polynomials  $f = p_1 p_2 \cdots p_r$ . Moreover, the representation is unique up to multiplying nonzero constants and relabeling.

**Proof.** WLOG,  $f$  is monic.

(existence) If  $\deg f = 1$ , then  $f(x) = x - a$  for some  $a \in F$ , which is itself irreducible.

Suppose  $\deg f > 1$ . Suppose the theorem holds for all  $g \in F[x]$  with  $\deg g < \deg f$ . If  $f$  is itself irreducible, then done. Otherwise, there are  $g_1, g_2 \in F[x]$  with  $\deg g_i \geq 1$  such that  $f = g_1 g_2$ . Then,  $\deg g_1$  and  $\deg g_2$  are less than  $f$ . Hence,  $g_1 = p_1 p_2 \cdots p_r$  and  $g_2 = q_1 q_2 \cdots q_s$  where  $p_j$  and  $q_j$  are irreducible, yielding  $f = p_1 \cdots p_r q_1 \cdots q_s$ .

(uniqueness) Suppose we have two factorization  $f = p_1 \cdots p_r = q_1 \cdots q_s$ .  $p_1 \mid q_1 \cdots q_s$ . Hence,  $p_1 \mid q_j$  for some  $j \in [s]$ . Since  $q_j$  is irreducible, this means  $p_1$  is a nonzero constant

multiple of  $q_j$ . Relabeling,  $p_1 = q_1$ , we have  $p_2 \cdots p_r = q_2 \cdots q_s$ . Proceeding in this way, we get  $r = s$  and  $p_j = q_j$  for each  $j$ .  $\square$

#### Definition 4.1.11: (Formal) Derivative

For  $f(x) = a_0 + a_1x + \cdots + a_nx^n \in F[x]$ , we define

$$f'(x) \triangleq a_1 + 2a_2x + \cdots + na_nx^{n-1}.$$

#### Note:-

- $(f + g)' = f' + g'$
- $(fg)' = f'g + fg'$

#### Theorem 4.1.6

$f$  is a product of distinct irreducible polynomials if and only if  $f$  and  $f'$  are relatively prime.

**Proof.**  $(\Leftarrow)$  Suppose  $f$  and  $f'$  are relatively prime but  $f = p^2h$  for some irreducible polynomial  $p$  for the sake of contradiction. Then,  $f' = p(2p'h + ph')$ , which contradicts  $(f, f') = (1)$ .  
 $(\Rightarrow)$   $\square$

#### Definition 4.1.12: Algebraically Closed

A field  $F$  is said to be *algebraically closed* if every irreducible polynomial in  $F[x]$  is of degree 1.

#### Note:-

$F$  is algebraically closed.

- $\Leftrightarrow$  Every  $f \in F[x]$  with  $\deg f \geq 1$  has precisely  $n$  roots counting multiplicity.
- $\Leftrightarrow$  Every non-constant  $f \in F[x]$  factors into linear polynomials.

#### Note:-

$\mathbb{C}$  is algebraically closed while  $\mathbb{R}$  is not.

# Chapter 5

## Determinants

### 5.1 Determinant Functions

#### Definition 5.1.1: $n$ -linear and Iterating

Let  $K$  be a ring. Let  $\mathcal{D} : K^{n \times n} \rightarrow K$  be a function. This is considered as a function on  $n$  row vectors.

- (i) We say  $\mathcal{D}$  is  $n$ -linear if  $\mathcal{D}$  is a linear function on the  $i^{\text{th}}$  row while fixing all other rows.

$$\mathcal{D} \begin{bmatrix} \cdots & a_1 + a'_1 & \cdots \\ \cdots & a_2 & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & a_n & \cdots \end{bmatrix} = \mathcal{D} \begin{bmatrix} \cdots & a_1 & \cdots \\ \cdots & a_2 & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & a_n & \cdots \end{bmatrix} + \mathcal{D} \begin{bmatrix} \cdots & a'_1 & \cdots \\ \cdots & a_2 & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & a_n & \cdots \end{bmatrix}$$

- (ii) An  $n$ -linear function  $\mathcal{D}$  is called *iterating* if  $\mathcal{D}(A) = 0$  when two rows are equal.

#### Note:-

If  $\mathcal{D}$  is iterating, and if  $A'$  is obtained by switching  $i^{\text{th}}$  and  $j^{\text{th}}$  rows of  $A$ , then  $\mathcal{D}(A') = -\mathcal{D}(A)$ .

#### Definition 5.1.2: Determinant

Let  $K$  be a commutative ring with unity. Let  $\mathcal{D} : K^{n \times n} \rightarrow K$  be a function. We say  $\mathcal{D}$  is a determinant function if

- (i)  $\mathcal{D}$  is  $n$ -linear,
- (ii)  $\mathcal{D}$  is alternating, and
- (iii)  $\mathcal{D}(I_n) = 1$ .

#### Definition 5.1.3: Minor Matrix

Let  $K$  be a commutative ring with unity. Let  $A \in K^{n \times n}$  where  $n > 1$ . For each  $i, j \in [n]$ , define  $A(i | j)$  be the  $(n-1) \times (n-1)$  matrix with the  $i^{\text{th}}$  row and the  $j^{\text{th}}$  column are removed.  $A(i | j)$  is called  $(i, j)$ -minor of  $A$ .

#### Theorem 5.1.1

There exists a determinant function  $\mathcal{D} : K^{n \times n} \rightarrow K$ .

**Proof.** We shall prove by exploiting mathematical induction. If  $n = 1$ , the identity function is a determinant function.

Suppose we have found a function  $\mathcal{D}: K^{(n-1) \times (n-1)}$  which is  $(n-1)$ -linear and alternating. We shall denote  $\mathcal{D}(A(i | j)) = D_{ij}(A)$ . Define  $E_i(A) \triangleq \sum_{j=1}^n (-1)^{i+j} A_{ij} D_{ij}(A)$  for each  $j \in [n]$ .

**Claim.**  $E_j$  is an  $n$ -linear function on  $K^{n \times n}$ .

$D_{ij}(A)$  is independent from the entries of the  $i$ -th row and the  $j$ -th column. Hence,  $D_{ij}$  is  $n$ -linear as  $\mathcal{D}$  is  $(n-1)$ -linear. Furthermore,  $A \mapsto A_{ij} D_{ij}(A)$  is also  $n$ -linear; thus  $E_j$  is linear combination of  $n$ -linear functions.

**Claim.**  $E_j$  is an alternating function on  $K^{n \times n}$ .

For the sake of simplicity, suppose  $A$  has two equal rows at  $\alpha_k$  and  $\alpha_{k+1}$ . Hence, when  $i \neq k$  and  $i \neq k+1$ ,  $A(i | j)$  has two identical rows; thus  $D_{ij}(A) = \mathcal{D}(A(i | j)) = 0$ . Thus,  $E_j(A) = (-1)^{k+j} A_{kj} D_{kj}(A) + (-1)^{k+j+1} A_{(k+1),j} D_{(k+1),j}(A)$ .

$$\begin{aligned} E_j(A) &= (-1)^{k+j} A_{kj} D_{kj}(A) + (-1)^{k+j+1} A_{(k+1),j} D_{(k+1),j}(A) \\ &= (-1)^{k+j} (A_{kj} D_{kj}(A) - A_{(k+1),j} D_{(k+1),j}(A)) = 0 \end{aligned}$$

**Claim.**  $E_j(I_n) = 1$ .

$$I_n(i | j) = I_{n-1}.$$

□

### Corollary 5.1.1

The function defined recursively in the proof of Theorem 5.1.1 is a determinant function.

### Definition 5.1.4: Permutation

Let  $S$  be a set. A permutation  $\sigma$  of  $S$  is a bijective function  $\sigma: S \rightarrow S$ .  $S_n$  is the set of bijective functions from  $[n]$  onto  $[n]$ .

### Definition 5.1.5: Transposition

$\tau \in S_n$  is called a *transposition* if it interchanges just the values of two members. A transposition that interchanges  $i$  and  $j$  is usually written as  $(i, j)$ .

### Definition 5.1.6: Cycle

A cycle is like:

$$i_1 \mapsto i_2 \mapsto i_3 \mapsto \cdots \mapsto i_n \mapsto i_1.$$

This is written as  $(i_1, i_2, \dots, i_n)$ .

### Note:-

- Every permutation can be written as a product of disjoint cycles.
- Every cycle can be written as a product of transpositions.
- Every permutation can be written as a product of transpositions.

### Theorem 5.1.2

For any permutation  $\sigma \in S_n$ , the number of transpositions needed to express  $\sigma$  modular 2 is an invariant of  $\sigma$ .

**Definition 5.1.7: Sign of Permutation**

$$\text{sign}(\sigma) \triangleq \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}$$

**Corollary 5.1.2**

For  $\sigma_1, \sigma_2 \in S_n$ ,  $\text{sign}(\sigma_1 \sigma_2) = \text{sign}(\sigma_1) \text{sign}(\sigma_2)$ .

**Theorem 5.1.3**

There exists a unique determinant function  $\mathcal{D}: K^{n \times n} \rightarrow K$ , which is equal to

$$\mathcal{D}(A) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \prod_{j \in [n]} A_{j, \sigma(j)}.$$

**Proof.** Let  $e_1, \dots, e_n$  be the rows of  $I_n$ . For  $A \in K^{n \times n}$ , let  $\alpha_i$  be the  $i$ -th rows of  $A$ . Then,  $\alpha_i = \sum_{j=1}^n A_{ij} e_j$ .

Note that, if  $j_i = j_{i'}$ , then  $\mathcal{D}(e_{j_1}, \dots, e_{j_n}) = 0$ . Also, if  $\sigma \in S_n$ ,  $\mathcal{D}(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \text{sign}(\sigma) \mathcal{D}(I_n) = \text{sign}(\sigma)$ .

$$\begin{aligned} \mathcal{D}(A) &= \mathcal{D}(\alpha_1, \alpha_2, \dots, \alpha_n) \\ &= \mathcal{D}\left(\sum_{j=1}^n A_{1j} e_j, \alpha_2, \dots, \alpha_n\right) \\ &= \sum_{j=1}^n A_{1j} \mathcal{D}(e_j, \alpha_2, \dots, \alpha_n) \\ &= \dots = \sum_{j_1=1}^n \sum_{j_2=1}^n \dots \sum_{j_n=1}^n A_{1j_1} A_{2j_2} \dots A_{nj_n} \mathcal{D}(e_{j_1}, \dots, e_{j_n}) \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \prod_{j \in [n]} A_{j, \sigma(j)} \end{aligned}$$

Note that, if  $\mathcal{D}$  is a  $n$ -linear and alternating, then  $\mathcal{D}(A) = \det A \cdot \mathcal{D}(I_n)$ . □

**Corollary 5.1.3**

$$\det(AB) = \det A \cdot \det B$$

**Corollary 5.1.4**

Any cofactor expansion gives the same value.

**Corollary 5.1.5**

$$\det A^t = \det A$$

**Proof.** Theorem 3.7.1 and Theorem 5.1.3. □

**Exercise 5.1.1**

Let  $A$  be  $r \times r$  matrix and  $C$  be an  $s \times s$  matrix. Then,

$$\det \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = \det A \cdot \det C.$$

Hint) Fixing  $A, B$ , define  $\mathcal{D}(A, B, C)$ .

**Definition 5.1.8: Adjoint Matrix**

Let  $A$  be an  $n \times n$  matrix.  $C_{ij} \triangleq (-1)^{i+j} \det(A(i \mid j))$  for each  $i, j \in [n]$  is called the  $(i, j)$ -cofactor. Then,  $\text{adj}A \triangleq C^t$  where  $(C)_{ij} = C_{ij}$  is called the *adjoint* of  $A$ .

**Corollary 5.1.6**

$A \cdot \text{adj}A = (\det A)I_n$ . If  $\det A \in K$  is invertible, then  $A^{-1} = (\det A)^{-1} \text{adj}A$ .

# Chapter 6

## Elementary Canonical Forms

### 6.1 Eigenvalues

#### Definition 6.1.1: Eigenvalue

Let  $V$  be a vector space over  $F$ . Let  $T: V \rightarrow V$  be a linear operator.

- $c \in F$  is said to be an *eigenvalue* (or a *characteristic value*) of  $T$  if there exists  $v \in V \setminus \{0\}$  such that  $T(v) = cv$ . Such  $v$  is called an *eigenvector* (or a *characteristic vector*) of  $T$  associated to  $c$ .
- For each  $c \in F$ ,  $E_c \triangleq \{v \in V \mid T(v) = cv\}$  is called an *eigenspace* (or a *characteristic space*) associated to  $c$ .

#### Theorem 6.1.1

Let  $V$  be a vector space over  $F$ . Let  $T: V \rightarrow V$  be a linear operator. Then, TFAE.

- (i)  $c \in F$  is an eigenvalue of  $T$ .
- (ii)  $T - cI$  is singular.
- (iii)  $\det(T - cI) = 0$ .

**Proof.** The equivalence of (i) and (ii) is trivial. The equivalence of (ii) and (iii) is evident from Corollary 5.1.6.  $\square$

#### Definition 6.1.2: Characteristic Polynomial

Let  $A$  be an  $n \times n$  matrix over  $F$ . Define  $f(x) \triangleq \det(xI - A) \in F[x]$ . Then,  $f$  is a monic polynomial in  $x$  of degree  $n = \dim V$ .

If there exists a basis  $\mathcal{B}$  for  $V$  and  $A = [T]_{\mathcal{B}}$ , then we call  $f(x) = \det(xI - A)$  the *characteristic polynomial* of  $T$ .

#### Note:-

The choice of basis does not affect the characteristic polynomial. See Theorem 3.4.3.

#### Note:-

If  $f$  is a characteristic polynomial of  $T$ , then  $f(c) = 0$  if and only if  $c$  is an eigenvalue of  $T$ .

#### Corollary 6.1.1

If  $T$  is a linear operator on  $V$ , then there are at most  $n$  eigenvalues of  $T$ .



**Proof.** Every polynomial of degree  $n$  has at most  $n$  solutions.  $\square$

### Definition 6.1.3: Diagonalizability

Let  $V$  be a finite-dimensional vector space over  $F$ . Let  $T \in L(V)$ . We say  $T$  is *diagonalizable* if there exists a basis  $\mathcal{B}$  such that it consists of eigenvectors of  $T$ .

#### Note:-

- If  $\mathcal{B} = \{v_1, \dots, v_n\}$  and  $Tv_i = c_i v_i$  for each  $i \in [n]$ , then  $[T]_{\mathcal{B}} = \text{diag}(c_1, c_2, \dots, c_n)$ .
- If  $T \in L(V)$  is diagonalizable, then the characteristic polynomial can be completely decomposed into a product of linear factors.

### Lemma 6.1.1

Let  $V$  be a finite-dimensional vector space over  $F$ . Let  $T \in L(V)$ . Suppose  $c_1, \dots, c_k \in F$  are all the possible distinct characteristic values of  $T$ . Let  $W_i$  be the eigenspace of  $c_i$ , i.e.,  $W_i = \ker(T - c_i I)$ . Then, if  $\mathcal{B}_i$  is a basis for  $W_i$  for each  $i \in [k]$ ,  $\bigcup_{i=1}^k \mathcal{B}_i$  is a basis for  $\sum_{i=1}^k W_i$ .

**Proof.** Suppose  $\sum \beta_i = 0$  where  $\beta_i \in W_i$ . Then, applying  $T, T^2, \dots, T^{k-1}$ , we get

$$\sum_{i=1}^k c_i^j \beta_i = 0$$

for each  $j \in \{0, \dots, k-1\}$ . As

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ c_1 & c_2 & \dots & c_k \\ \vdots & \vdots & \ddots & \vdots \\ c_1^{k-1} & c_2^{k-1} & \dots & c_k^{k-1} \end{bmatrix}$$

is invertible since  $c_i$ 's are distinct, we get  $\beta_i = 0$  for each  $i$ .  $\square$

#### Note:-

Lemma 6.1.1 also implies that  $\dim(\sum_{i=1}^k W_i) = \sum_{i=1}^k \dim W_i$ .

### Theorem 6.1.2

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Let  $T \in L(V)$ . Suppose  $c_1, \dots, c_k \in F$  are all the possible distinct characteristic values of  $T$ . Let  $W_i$  be the eigenspace of  $c_i$ , i.e.,  $W_i = \ker(T - c_i I)$ . TFAE.

- $T$  is diagonalizable.
- The characteristic polynomial is  $p(x) = \prod_{i=1}^k (x - c_i)^{d_i}$  where  $d_i = \dim W_i$ .
- $\sum_{i=1}^k d_i = n$ .

**Proof.** ((i)  $\Rightarrow$  (ii)) Let  $\mathcal{B}$  be the basis for  $V$  that consists of eigenvectors of  $T$ . If  $\mathcal{B}_i$  is the part of  $\mathcal{B}$  that only consists of eigenvectors corresponding to  $c_i$ , we have  $\text{span } \mathcal{B}_i = W_i$ . Hence, on

rearranging,  $[T]_{\mathcal{B}} = \text{diag}(\overbrace{c_1, \dots, c_1}^{d_1}, \overbrace{c_2, \dots, c_2}^{d_2}, \dots, \overbrace{c_k, \dots, c_k}^{d_k})$ .

((ii)  $\Rightarrow$  (iii)) A direct consequence of Lemma 6.1.1.

((iii)  $\Rightarrow$  (i))  $\dim \sum W_i = \sum \dim W_i = \sum d_i = n$ . Hence,  $\sum W_i = V$ , i.e.,  $V$  has a basis consisting of characteristic vectors.  $\square$

## 6.2 Annihilating Polynomials

### Note:-

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Let  $T \in L(V)$ .  $\{f \in F[x] \mid f(T) = 0\}$  is a nonzero ideal as  $\{I, T, T^2, \dots, T^{n^2}\}$  is linearly dependent.

### Definition 6.2.1: Minimal Polynomial

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Let  $T \in L(V)$ . The monic generator of the nonzero ideal  $\{f \in F[x] \mid f(T) = 0\}$  is called the *minimal polynomial* of  $T$ .

### Theorem 6.2.1

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Let  $T \in L(V)$ . If  $p(x)$  is the characteristic polynomial of  $T$  and  $m(x)$  is the minimal polynomial of  $T$ , then  $p(x)$  and  $m(x)$  has the same solutions in  $F$ .

**Proof.** ( $\Rightarrow$ ) Suppose  $m(c) = 0$ . Then,  $m(x) = (x - c)q(x)$  for some  $q \in F[x]$ . As  $m$  is minimal,  $q(T) \neq 0$ . This means that  $q(T)(\beta) \neq 0$  for some  $\beta \in V$ . However,  $m(T)(\beta) = ((T - cI)q(T))(\beta) = 0$ ; hence  $q(T)(\beta) \in \ker(T - cI)$ , i.e.,  $c$  is an eigenvalue. This means that  $p(c) = 0$ .

( $\Leftarrow$ ) Suppose  $p(c) = 0$ , i.e.,  $T(\alpha) = c\alpha$  for some nonzero  $\alpha \in V$ . As  $T^k(\alpha) = c^k\alpha$  for all  $k \in \mathbb{N} \cup \{0\}$ , for any polynomial  $f \in F[x]$ , we have  $f(T)(\alpha) = f(c)\alpha$ . In particular,  $0 = m(T)\alpha = m(c)\alpha$ , i.e.,  $m(c) = 0$ .  $\square$

### Theorem 6.2.2 Cayley-Hamilton

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Let  $T \in L(V)$ . If  $p(x)$  is the characteristic polynomial of  $T$ , then  $p(T) = 0$ .

**Proof.** Let  $K \triangleq \{h(T) \mid h \in F[x]\}$  be a commutative ring. Let  $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$  be a basis for  $V$ . Let  $A \triangleq [T]_{\mathcal{B}}$  so that  $T(\alpha_i) = \sum_{j=1}^n A_{ji}\alpha_j$ . This is equivalent to  $\sum_{j=1}^n (\delta_{ij}T - A_{ji}I)\alpha_j = 0$ .

Let  $B_{ij} \triangleq \delta_{ij}T - A_{ji}I \in K$  and  $B \triangleq [B_{ij}]$ . Then,  $(\text{adj } B)B = B(\text{adj } B) = (\det B)I$ . By construction,  $\sum_{j=1}^n (\text{adj } B)_{ki}B_{ij}\alpha_j = 0$  for all  $k, i \in [n]$ .

Taking sum over  $i$ , we have

$$\begin{aligned} 0 &= \sum_{i=1}^n \sum_{j=1}^n (\text{adj } B)_{ki}B_{ij}\alpha_j \\ &= \sum_{j=1}^n \left( \sum_{i=1}^n (\text{adj } B)_{ki}B_{ij} \right) \alpha_j \\ &= \sum_{j=1}^n \delta_{kj}(\det B)\alpha_j = (\det B)\alpha_k \end{aligned}$$

for each  $k \in [n]$ . As  $\{\alpha_1, \dots, \alpha_n\}$  is a basis of  $V$ , we have  $\det B = 0$ , i.e.,  $p(T) = 0$ .  $\square$

## 6.3 Invariant Subspaces

### Definition 6.3.1: $T$ -Invariant Subspace

Let  $V$  be a finite-dimensional vector space over  $F$  and  $W$  be a subspace of  $V$ . Let  $T \in L(V)$ . Then,  $W$  is said to be a  $T$ -invariant subspace if  $T(W) \subseteq W$ .

#### Note:-

If  $W$  is a  $T$ -invariant subspace of  $V$ , then  $T|_W$  is a naturally induced linear operator on  $W$ .

### Example 6.3.1

Let  $V$  be a finite-dimensional vector space over  $F$  and  $T \in L(V)$ .

- $W = \{0\}$  is a  $T$ -invariant subspace.
- For every eigenvalue  $c$  of  $T$ ,  $E_c = \ker(T - cI)$  is a  $T$ -invariant subspace.

### Lemma 6.3.1

Let  $V$  be a finite-dimensional vector space over  $F$  and  $T \in L(V)$ . Let  $W$  be a  $T$ -invariant subspace of  $V$ . Then,  $m_W \mid m$  and  $f_W \mid f$  where  $m_W$  and  $m$  are minimal polynomials of  $T|_W$  and  $T$ , and  $f_W$  and  $f$  are characteristic polynomials of  $T|_W$  and  $T$ .

**Proof.** Let  $\mathcal{B}' = \{\alpha_1, \dots, \alpha_k\}$  be a basis for  $W$ , and extend it to  $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$  so  $\mathcal{B}$  is a basis for  $V$ . As  $W$  is  $T$ -invariant, we have

$$M \triangleq [T]_{\mathcal{B}} = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

where  $A = [T|_W]_{\mathcal{B}'}$ . Hence,  $f(x) = \det(xI - M) = \det(xI - A) \det(xI - C) = f_W(x) \det(xI - C)$ .

Now, noting that  $M^r = \begin{bmatrix} A^r & * \\ 0 & C^r \end{bmatrix}$ , whenever  $p(x) \in F[x]$  satisfies  $p(M) = 0$ , we always have  $p(A) = 0$  as  $A$  is invertible;  $m(A) = 0$ . By the definition of  $m_W$ , we have  $m_W \mid m$ .  $\square$

### Definition 6.3.2: $T$ -conductor

Let  $V$  be a finite-dimensional vector space over  $F$  and  $T \in L(V)$ . Let  $W$  be a  $T$ -invariant subspace of  $V$ . Then, for each  $\alpha \in V$ , the set

$$S_T(\alpha; W) \triangleq \{g \in F[x] \mid g(T)\alpha \in W\}$$

is called the  $T$ -conductor of  $\alpha$  to  $W$ .

### Lemma 6.3.2

Let  $V$  be a finite-dimensional vector space over  $F$  and  $T \in L(V)$ . Let  $W$  be a  $T$ -invariant subspace of  $V$ . Then, for each  $\alpha \in V$ ,  $S_T(\alpha; W)$  is a nonzero ideal.

**Proof.**  $S_T(\alpha, W)$  is nonzero as the characteristic polynomial is contained in the set by Theorem 6.2.2.

It is evident that  $S_T(\alpha, W)$  is a subspace of  $F[x]$ . Now, take any  $h \in F[x]$  and  $g \in S_T(\alpha; W)$ . Then,  $(hg)(T)\alpha = h(T)g(T)\alpha \in W$  as  $W$  is  $T$ -invariant and  $g(T)\alpha \in W$ .  $\square$

**Definition 6.3.3:  $T$ -conductor**

Let  $V$  be a finite-dimensional vector space over  $F$  and  $T \in L(V)$ . Let  $W$  be a  $T$ -invariant subspace of  $V$ . Due to Lemma 6.3.2 and Theorem 4.1.3, there uniquely exists the monic generator  $g_{T,\alpha,W}$  of  $S_T(\alpha, W)$ .  $g_{T,\alpha,W}$  is also often called the  $T$ -conductor of  $\alpha$  to  $W$ .

**Note:-**

Since  $m(T) = f(T) = 0$  where  $m$  and  $f$  are minimal and characteristic polynomials of  $T$ , they are elements of  $S_T(\alpha, W)$  for any  $\alpha, W$ . Hence,

$$g_{T,\alpha,W} \mid m \mid f.$$

**Definition 6.3.4: Triangulable Matrix**

Let  $V$  be a finite-dimensional vector space over  $F$  and  $T \in L(V)$ .  $T$  is said to be *triangulable* if there exists basis  $\mathcal{B}$  for  $V$  such that  $[T]_{\mathcal{B}}$  is upper triangular matrix.

**Note:-**

If  $T$  is diagonalizable, then  $T$  is triangulable.

**Lemma 6.3.3**

Let  $V$  be a finite-dimensional vector space over  $F$ . Let  $T : V \rightarrow V$  be a linear operator on  $V$  such that the minimal polynomial  $m$  of  $T$  has the form of

$$m(x) = \prod_{i=1}^k (x - c_i)^{r_i}.$$

If  $W$  is a proper  $T$ -invariant subspace of  $V$ , then there exists  $\alpha \in V \setminus W$  and an eigenvalue  $c \in F$  such that  $(T - cI)\alpha \in W$ . In other words,  $x - c$  is the  $T$ -conductor of  $\alpha$  on  $W$ .

**Proof.** Take  $\beta \in V \setminus W$ . Then,  $g \triangleq g_{T,\beta,W} \mid m$ , i.e.,

$$g(x) = \prod_{i=1}^k (x - c_i)^{e_i}.$$

By the definition of  $g$ , and since  $\beta \notin W$ , there exists  $j \in [k]$  such that  $e_j \geq 1$ .  $g(x) = (x - c_j)h(x)$  for some  $h \in F[x]$ . By the minimality of  $g$ ,  $\alpha \triangleq h(T)\beta \notin V \setminus W$  but  $(T - c_jI)\alpha = (T - c_jI)h(T)\beta = g(T)\beta \in W$ .  $\square$

**Note:-**

For  $\alpha \notin W$  and  $T \in L(V)$ , TFAE.

- (i)  $(T - cI)\alpha \in W$  for some  $c \in F$ .
- (ii)  $x - c$  is the  $T$ -conductor of  $\alpha$  on  $W$  for some  $c \in F$ .
- (iii)  $T\alpha \in \text{span}\{W, \alpha\}$ .

**Theorem 6.3.1**

Let  $V$  be a finite-dimensional vector space over  $F$ . Let  $T : V \rightarrow V$  be a linear operator on  $V$ . Then,  $T$  is triangulable if and only if the minimal polynomial of  $T$  is a product of linear polynomials over  $F$ .

**Proof.** ( $\Rightarrow$ ) Since  $T$  is triangulable, there exists a basis  $\mathcal{B}$  such that  $A = [T]_{\mathcal{B}}$  is upper triangular. Hence, the characteristic polynomial is  $\det(xI - A) = \prod_{i=1}^n (x - (A)_{ii})$ . The result follows due to Theorem 6.2.1.

( $\Leftarrow$ ) Suppose  $m(x) = \prod_{i=1}^k (x - c_i)^{r_i}$ . We shall make use of Lemma 6.3.3 repeatedly over different choices of  $W$ . With  $W = \{0\}$ , we have  $\alpha \in V \setminus \{0\}$  such that  $(T - d_1 I)\alpha_1 = 0$  for some eigenvalue  $d_1$ . Inductively define  $\alpha_i$  by:

- $W_i = \text{span}\{\alpha_1, \dots, \alpha_i\}$ .
- Thanks to Lemma 6.3.3, take  $\alpha_{i+1} \in V \setminus W_i$  such that  $(T - d_{i+1} I)\alpha_{i+1} \in W_i$  where  $d_{i+1}$  is an eigenvalue.

Then, by construction,  $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$  is a basis for  $V$  and  $[T]_{\mathcal{B}}$  is an upper triangular matrix since  $T\alpha_{i+1} \in \text{span}\{\alpha_1, \dots, \alpha_i\} + d_{i+1}\alpha_{i+1}$ .  $\square$

### Corollary 6.3.1

Let  $V$  be a  $n$ -dimensional vector space over an algebraically closed field  $F$ . Then, every linear operator on  $V$  is triangulable.

### Theorem 6.3.2

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Let  $T \in L(V)$ . Then,  $T$  is diagonalizable if and only if the minimal polynomial is  $m(x) = \prod_{i=1}^k (x - c_i)$  where  $c_1, \dots, c_k$  are all the distinct eigenvalues of  $T$ .

**Proof.** ( $\Rightarrow$ ) By Theorem 6.1.2 and Theorem 6.2.1, we already have  $m(x) = \prod_{i=1}^k (x - c_i)^{e_i}$  where  $e_i \geq 1$ . Now, we claim that  $S \triangleq \prod_{i=1}^k (T - c_i I) = 0$ .

From assumption, there exists a basis  $\{\alpha_1, \dots, \alpha_n\}$  for  $V$  which consists of eigenvectors. Let  $\alpha_j$  corresponds to the eigenvalue  $c_{i(j)}$ . Then, for each  $j \in [n]$ ,  $(T - c_{i(j)} I)\alpha_j = 0$ , i.e.,  $S\alpha_j = 0$ . Therefore,  $S = 0$ .

( $\Leftarrow$ ) Let  $W$  be the subspace spanned by eigenvectors of  $T$ . For the sake of contradiction, suppose  $W \subsetneq V$ . As  $W$  is  $T$ -invariant, by Lemma 6.3.3, there exists  $\alpha \in V \setminus W$  and an eigenvalue  $c_j \in F$  such that  $\beta \triangleq (T - c_j I)\alpha \in W$ .

Write  $m(x) = (x - c_j)h(x)$  so  $h$  does not have  $x - c_j$  as a factor of it. As  $h(x) - h(c_j)$  has  $x = c_j$  as a root,  $h(x) - h(c_j) = (x - c_j)q(x)$  for some  $q$ . Then, we have

$$h(T)\alpha - h(c_j)\alpha = q(T)(T - c_j I)\alpha = q(T)\beta \in W$$

since  $W$  is  $T$ -invariant.

Moreover,  $0 = m(T)\alpha = (T - c_j I)h(T)\alpha$  and thus  $h(T)\alpha \in E_{c_j} \subseteq W$ . This implies that  $h(c_j)\alpha \in W$  but  $\alpha \notin W$ ; thus  $h(c_j) = 0$ , implying the multiplicity of  $x - c_j$  in the minimal polynomial.  $\square$

## 6.4 Simultaneous Triangulation and Diagonalization

### Definition 6.4.1: Commuting Family of Linear Operators

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . A set of linear operators  $\mathcal{F}$  is said to be a *commuting family* of linear operators if  $T_1 T_2 = T_2 T_1$  for each  $T_1, T_2 \in \mathcal{F}$ .

**Definition 6.4.2:  $\mathcal{F}$ -invariant**

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . A subspace  $W$  of  $V$  is said to be  $\mathcal{F}$ -invariant if it is  $T$ -invariant for all  $T \in \mathcal{F}$ .

**Lemma 6.4.1**

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Suppose  $\mathcal{F}$  is a commuting family of triangulable linear operators on  $V$ . Suppose a proper subspace  $W$  of  $V$  is  $\mathcal{F}$ -invariant. Then, there exists  $\alpha \in V \setminus W$  such that  $\forall T \in \mathcal{F}, T\alpha \in \text{span}\{W, \alpha\}$ .

**Proof.** Suppose  $\{T_1, \dots, T_r\}$  is a basis for the subspace spanned by  $\mathcal{F}$ . Note that  $\text{span } \mathcal{F}$  is still a commuting family of triangulable linear operators.

Let  $V_0 = V$ . Construct  $V_1, \dots, V_r$  and  $\beta_1, \dots, \beta_r$  as follows. For each  $i \in [r]$ ,

- (i) Let  $U_i = T_i|_{V_{i-1}}$ . Then,  $U_i \in L(V_{i-1})$  by (iii)-(c).
- (ii) Take  $\beta_i \in V_{i-1} \setminus W$  and  $c_i \in F$  such that  $(U_i - c_i I)\beta_i \in W$ . Their existence is guaranteed by Lemma 6.3.3 and (iii)-(b).
- (iii) Let  $V_i \triangleq \{\beta \in V_{i-1} \mid (T_i - c_i I)\beta \in W\}$ . Then, by construction, the following hold.
  - (a)  $\beta_i \in V_i \setminus W$
  - (b)  $W \subsetneq V_i \subseteq V_{i-1}$
  - (c)  $V_i$  is  $\mathcal{F}$ -invariant as, for each  $T \in \mathcal{F}$  and  $\beta \in V_i$ ,  $(T_i - c_i I)(T\beta) = T(T_i - c_i I)\beta \in W$ , i.e.,  $T\beta \in V_i$ .

Then,  $\beta_r$  satisfies  $T_i \beta_r \in \text{span}\{W, \beta_r\}$  for each  $i \in [r]$ . □

**Corollary 6.4.1**

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Let  $\mathcal{F}$  be a commuting family of triangulable linear operators on  $V$ . Then, there exists a basis  $\mathcal{B}$  for  $V$  such that  $[T]_{\mathcal{B}}$  is an upper triangular matrix for all  $T \in \mathcal{F}$ .

**Proof.** Take any  $\alpha_1 \in V$ . Now, construct  $\alpha_2, \dots, \alpha_n$  as following. For each  $i \in [n-1]$ ,

- Let  $W_i \triangleq \text{span}\{\alpha_1, \dots, \alpha_i\}$ .
- Take  $\alpha_{i+1} \in V \setminus W_i$  such that  $T\alpha_{i+1} \in \text{span}\{\alpha_1, \dots, \alpha_{i+1}\}$  for each  $T \in \mathcal{F}$ . The existence is guaranteed by Lemma 6.4.1.

Then,  $\mathcal{B} = \{\alpha_1; \dots; \alpha_n\}$  is the ordered basis we are looking for. □

**Theorem 6.4.1**

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Let  $\mathcal{F}$  be a commuting family of diagonalizable linear operators on  $V$ . Then, there exists a basis  $\mathcal{B}$  for  $V$  such that  $[T]_{\mathcal{B}}$  is a diagonal matrix for all  $T \in \mathcal{F}$ .

**Proof.** We will apply the mathematical induction over  $\dim V$ . If  $\dim V = 1$ , there is nothing to prove. Hence, suppose the theorem holds for any finite-dimensional vector space  $V$  over  $F$  with dimension less than  $n$ .

If  $\mathcal{F}$  only consists of multiples of identity, it is done. So we may assume the existence of  $T \in \mathcal{F}$  which is not a multiple of identity. Let  $c_1, \dots, c_k$  be its distinct characteristic values. For each  $i \in [k]$ , let  $\mathcal{F}_i \triangleq \{T|_{W_i} \in L(W_i, V) : T \in \mathcal{F}\}$  where  $W_i$  is the eigenspace associated to  $c_i$ . Then:

- (i) As  $T$  is not a multiple of identity,  $k > 1$  and  $\dim W_i < n$ .
- (ii) As  $W_i$  is  $\mathcal{F}$ -invariant,  $\mathcal{F}_i \subseteq L(W_i)$ .

- (iii) For all  $T' \in \mathcal{F}$ , if  $m_i$  and  $m$  are minimal polynomials of  $T'|_{W_i}$  and  $T'$ ,  $m_i \mid m$  thanks to Lemma 6.3.1.
- (iv) By (iii) and Theorem 6.3.2, every linear operator in  $\mathcal{F}_i$  is diagonalizable.
- (v) By (i), (iv), and the induction hypothesis, there exists a basis  $\mathcal{B}_i$  for  $W_i$  that simultaneously diagonalize all linear operators in  $\mathcal{F}_i$ .

Now,  $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_k)$  is an ordered basis for  $V$  due to Lemma 6.1.1, and  $\mathcal{B}$  is the basis we are looking for.  $\square$

### Corollary 6.4.2

Let  $V$  be a  $n$ -dimensional vector space over an algebraically closed field  $F$ . Let  $\mathcal{F}$  be a commuting family of linear operators on  $V$ . Then, there exists a basis  $\mathcal{B}$  for  $V$  such that  $[T]_{\mathcal{B}}$  is a diagonal matrix for all  $T \in \mathcal{F}$ .

## 6.5 Direct-Sum Decompositions

### Definition 6.5.1: Independent Subspaces

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . We say subspaces  $W_1, \dots, W_k$  of  $V$  are *independent* if, whenever  $a_1 + \dots + a_k = 0$  where  $a_i \in W_i$ ,  $a_i = 0$  for all  $i \in [k]$ .

### Definition 6.5.2: Direct Sum

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Let  $W_1, \dots, W_k$  be the finite number of subspaces of  $V$ . Then, we say that the sum  $W = \sum_{i=1}^k W_i$  is *direct* if  $W_1, \dots, W_k$  are independent. We write  $W = W_1 \oplus \dots \oplus W_k = \bigoplus_{i=1}^k W_i$  if the sum is direct.

### Definition 6.5.3: Projection

Let  $V$  be a vector space over  $F$ . A linear operator  $E \in L(V)$  such that  $E^2 = E$  is called a *projection*.

### Example 6.5.1

Suppose  $V = V_1 \oplus V_2$ . Then,  $P_1 \in L(V)$  defined by  $v_1 + v_2 \mapsto v_1$  where  $v_1 \in V_1$  and  $v_2 \in V_2$  is a projection.

### Lemma 6.5.1

Let  $V$  be a vector space over  $F$ . Let  $E \in L(V)$  be a projection. Then,  $V = V_1 \oplus V_2$  for some subspaces  $V_1$  and  $V_2$  of  $V$  such that  $E$  can be represented by  $E(v_1 + v_2) = v_1$  where  $v_1 \in V_1$  and  $v_2 \in V_2$ .

**Proof.** Take  $V_1 = \text{Im } E$  and  $V_2 = \ker E$ . Take any  $v \in V$ . Then,  $v = Ev + (v - Ev)$  while  $Ev \in V_1$  and  $v - Ev \in \ker E$ . Hence,  $V = V_1 + V_2$ .

Take any  $v_1 \in \text{Im } E$  and  $v_2 \in \ker E$  and suppose  $v_1 + v_2 = 0$ . Then, there exists  $v'_1 \in V$  such that  $v_1 = E(v'_1)$ . Then,  $0 = E(v_1 + v_2) = E(v_1) = E^2(v'_1) = E(v'_1) = v_1$ . Hence, the sum is direct. It is also shown that  $E(v_1 + v_2) = E(v_1) = v_1$ .  $\square$

### Theorem 6.5.1

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Suppose  $V = \bigoplus_{i=1}^k W_i$  for some subspaces  $W_i$  of  $V$ . Then, for each  $i \in [k]$ , there exists  $E_i \in L(V)$  such that

- (i)  $E_i$  is a projection for each  $i \in [k]$ ,
- (ii)  $E_i E_j = 0$  if  $i \neq j$ .
- (iii)  $I = \sum_{i=1}^k E_i$ .
- (iv)  $\text{Im } E_i = W_i$  for each  $i \in [k]$ .

**Proof.** All  $v \in V$  can be uniquely written as  $v = \sum_{i=1}^k v_i$  where  $v_i \in W_i$  for each  $i \in [k]$ . Hence, define  $E_i: V \rightarrow V$  by  $v \mapsto v_i$ . Then,  $E_i$ 's satisfy the four constraints.  $\square$

## 6.6 Invariant Direct Sums

### Theorem 6.6.1

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Let  $T \in L(V)$ . Suppose  $V = \bigoplus_{i=1}^k W_i$  for some subspaces  $W_i$  of  $V$ . Let  $E_1, \dots, E_k$  be the projections in Theorem 6.5.1. Then,  $W_i$  is  $T$ -invariant for all  $i \in [k]$  if and only if  $T$  commutes with all  $E_i$ 's.

**Proof.** ( $\Rightarrow$ ) Suppose  $W_i$  is  $T$ -invariant for each  $i \in [k]$ . Take any  $\alpha \in V$  and write  $\alpha = \sum_{i=1}^k \alpha_i$  where  $\alpha_i \in W_i$  for each  $i \in [k]$ . Then,  $E_i T \alpha = \sum_{j=1}^k E_i T \alpha_j = T \alpha_i = T E_i \alpha$ .

( $\Leftarrow$ ) Suppose  $T E_i = E_i T$ . Take any  $\alpha_i \in W_i$ . Then,  $T \alpha_i = T E_i \alpha_i = E_i (T \alpha_i) \in W_i$  by the definition of  $E_i$ . Hence,  $W_i$  is  $T$ -invariant.  $\square$

### Theorem 6.6.2

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Let  $T \in L(V)$ . If  $T$  is diagonalizable and  $c_1, \dots, c_k$  are all the distinct eigenvalues, we have projections  $E_i \in L(V)$  for each  $i \in [k]$  on  $W_i = E_{c_i}$  such that  $T = \sum_{i=1}^k c_i E_i$  and  $V = \bigoplus_{i=1}^k W_i$  with  $I = \sum_{i=1}^k E_i$  and  $E_i E_j = \delta_{ij} E_i$ .

#### Note:-

The converse of Theorem 6.6.2 also holds.

## 6.7 The Primary Decomposition Theorem

### Theorem 6.7.1 Primary Decomposition Theorem

Let  $V$  be a  $n$ -dimensional vector space over  $F$ . Let  $T \in L(V)$  and  $m \in F[x]$  be its minimal polynomial. Write  $m(x) = \prod_{i=1}^k p_i^{r_i}$  where  $p_i$ 's are irreducible polynomials in  $F[x]$  and  $r_i \geq 1$ . Let  $W_i \triangleq \ker(p_i(T)^{r_i})$ . Then, the following hold.

- (i)  $V = \bigoplus_{i=1}^k W_i$ .
- (ii) Each  $W_i$  is  $T$ -invariant.
- (iii) The minimal polynomial of  $T_i = T|_{W_i}$  is  $p_i^{r_i}$  for each  $i \in [k]$ .

**Proof.** If  $k = 1$ , there is nothing to prove. Hence, we may assume  $k \geq 2$ .



Define for each  $i \in [k]$ ,  $f_i \triangleq \prod_{j \in [n] \setminus \{i\}} p_j^{r_j}$  so that  $(f_i, p_i^{r_i}) = 1$ . Since  $f_1, \dots, f_k$  are also relatively prime, there exists  $g_1, \dots, g_k \in F[x]$  such that  $f_1 g_1 + \dots + f_k g_k = 1$ . Define  $h_i \triangleq f_i g_i$  so  $\sum_{i=1}^k h_i(T) = I$ . When  $i \neq j$ , we have  $m \mid f_i f_j$  and  $f_i(T) f_j(T) = 0$ .

Define  $E_i \triangleq h_i(T) \in L(V)$ . Then, we have  $\sum_{i=1}^k E_i = I$  and  $E_i E_j = f_i(T) f_j(T) g_i(T) g_j(T) = 0$  for each  $i \neq j$ . Moreover,  $E_j = E_j \sum_{i=1}^k E_i = E_j^2$ , i.e.,  $E_j$  is a projection for each  $j \in [k]$ . Then,  $V = \bigoplus_{i=1}^k \text{Im } E_i$  and each  $\text{Im } E_i$  is  $T$ -invariant.

Now, we claim that  $\text{Im } E_i = W_i = \ker(p_i(T)^{r_i})$ .

- Take any  $\alpha \in \text{Im } E_i$ . Then,  $\alpha = E_i \alpha$ . This implies  $p_i(T)^{r_i}(\alpha) = p_i(T)^{r_i} f_i(T) g_i(T) \alpha = 0$  as  $p_i^{r_i} f_i = m$ . Hence,  $\text{Im } E_i \subseteq W_i$ .
- Take any  $\alpha \in \ker(p_i(T)^{r_i})$ . If  $j \neq i$ , then  $p_i^{r_i} \mid f_j \mid f_j g_j$ . This implies that  $f_j(T) g_j(T) \alpha = h_j(T) \alpha = 0$ . In other words,  $E_j \alpha = 0$  for each  $j \neq i$ , this restricts to the only left option:  $\alpha \in \text{Im } E_i$ . Hence,  $W_i \subseteq \text{Im } E_i$ .

It remains to show that  $T_i = T|_{W_i}$  has the minimal polynomial  $p_i^{r_i}$ . Let  $m_i$  be the minimal polynomial of  $T_i$ . By the definition of  $W_i$ , we have  $p_i(T)^{r_i}|_{W_i} = 0$ . Hence,  $m_i \mid p_i^{r_i}$ . Let  $g$  be any polynomial in  $F[x]$  such that  $g(T_i) = 0$ . We now claim that  $p_i^{r_i} \mid g$ . For each  $\alpha_j \in W_j = \ker(p_j(T)^{r_j})$  with  $i \neq j$ , we have  $f_i(T) \alpha_j = 0$ . Hence,  $g(T) f_i(T) = 0$ .  $m \mid g f_i$ . However, as  $(p_i^{r_i}, f_i) = (1)$ ,  $m = \prod_{j=1}^k p_j^{r_j} \mid g \prod_{i \neq j} p_j^{r_j}$  directly implies that  $p_i^{r_i} \mid g$ .  $\square$

### Corollary 6.7.1

If  $E_1, \dots, E_k$  are projections associated to the primary decomposition of  $V$  with respect to  $T$ , then each  $E_i$  is a polynomial in  $T$ .

In particular, if  $U \in L(V)$  commutes with  $T$ , then  $U$  commutes with all  $E_i$  so each  $W_i$  is  $U$ -invariant.

### Definition 6.7.1: Nilpotent Linear Operator

Let  $V$  be a finite-dimensional vector space over  $F$ .  $T \in L(V)$  is called a *nilpotent* operator if  $T^N = 0$  for some  $N \in \mathbb{N}$ .

### Theorem 6.7.2

Let  $V$  be a finite-dimensional vector space over  $F$ . Let  $T \in L(V)$  be a triangulable linear operator. Then, there *uniquely* exists a diagonalizable  $D \in L(V)$  and a nilpotent  $N \in L(V)$  such that

- $T = D + N$  and
- $DN = ND$ .

**Proof.** Let  $m(x) = \prod_{i=1}^k (x - c_i)^{r_i}$  be the minimal polynomial of  $T$ . As in Theorem 6.7.1, take  $W_i \triangleq \ker(T - c_i I)^{r_i} = \text{Im } E_i$  where  $E_i$  is the projection to  $W_i$ .

Take  $D = \sum_{i=1}^k c_i E_i$  and  $N = T - D$ . Then,  $D$  is diagonalizable. Now, we claim that  $N$  is nilpotent. As  $I = \sum_{i=1}^k E_i$ ,  $D = \sum_{i=1}^k (T - c_i I) E_i$ . Hence,  $N^r = \sum_{i=1}^k (T - c_i I)^r E_i$  as  $T$  and  $E_i$  commute, and as  $E_i$ 's are projections onto independent subspaces. Hence,  $N^{\max_{i=1}^k r_i} = 0$ ;  $N$  is nilpotent. Furthermore,  $D$  and  $N$  are polynomials in  $T$ ; hence they commute.

Now, we are left with the proof for uniqueness. Suppose we have another  $D'$  and  $N'$  that satisfy (i) and (ii).  $D + N = T = D' + N'$  implies that  $A = D - D' = N' - N$  is both diagonalizable and nilpotent. In other words,  $A = 0$ , i.e.,  $D = D'$  and  $N = N'$ .  $\square$

**Note:-**

$D$  and  $N$  in Theorem 6.7.2 are called *diagonalizable part* and *nilpotent part* of  $T$ , respectively.

# Chapter 7

## The Rational and Jordan Forms

### 7.1 Cyclic Subspaces and Annihilators

#### Definition 7.1.1: $T$ -cyclically Generated Subspace

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $T \in L(V)$ . For  $\alpha \in V$ , the subspace

$$Z(\alpha; T) = \{g(T)\alpha \mid g \in F[x]\}$$

of  $V$  is called the  $T$ -cyclic subspace generated by  $\alpha$ . If  $Z(\alpha; T) = V$ , then we say  $V$  is cyclically generated by  $\alpha$ , and  $\alpha$  is a cyclic vector for  $T$ .

#### Note:-

Some immediate facts:

- $Z(\alpha; T)$  is  $T$ -invariant.
- $Z(0; T) = \{0\}$ .
- If  $\alpha \neq 0$  is an eigenvector, then  $Z(\alpha; T) = \text{span}\{\alpha\}$ .
- If  $\dim Z(\alpha; T) = 1$ , then  $\alpha \neq 0$  and  $Z(\alpha; T) = \text{span}\{\alpha\}$ ; thus  $\alpha$  is an eigenvector.

So, we need  $\alpha$  be neither too bad nor too good to utilize  $Z(\alpha; T)$ .

#### Definition 7.1.2: $T$ -annihilator

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $T \in L(V)$ . For  $\alpha \in V$ , the  $T$ -annihilator of  $\alpha$  is the subspace

$$M(\alpha; T) \triangleq \{g \in F[x] \mid g(T)\alpha = 0\}.$$

In other words,  $M(\alpha; T) = S_T(\alpha; \{0\})$ .

#### Note:-

$T$ -annihilator of  $\alpha$  is the  $T$ -conductor of  $\alpha$  to  $\{0\}$ ,  $M(\alpha; T)$  is a nonzero ideal and thus has a unique monic generator  $p_\alpha$ .  $p_\alpha$  is also called the  $T$ -annihilator of  $\alpha$ . Hence, as the minimal polynomial  $m$  of  $T$  resides in  $M(\alpha; T)$ , we have  $p_\alpha \mid m$ .

#### Theorem 7.1.1

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $T \in L(V)$ . Let  $\alpha \in V \setminus \{0\}$  be fixed. Let  $p_\alpha$  be the  $T$ -annihilator of  $\alpha$ .

- (i) If  $k = \deg p_\alpha$ ,  $\{\alpha, T\alpha, \dots, T^{k-1}\alpha\}$  is a basis for  $Z(\alpha; T)$ , hence  $\deg p_\alpha = \dim Z(\alpha; T)$ .

(ii) Let  $U \triangleq T|_{Z(\alpha; T)} \in L(Z(\alpha; T))$ . Then, the minimal polynomial of  $U$  is  $p_\alpha$ .

**Proof.**

- (i) Let  $g \in F[x]$  be arbitrary. By Theorem 4.1.2, we have  $g = p_\alpha q + r$  where  $q, r \in F[x]$  in which either  $r = 0$  or  $r \neq 0$  and  $\deg r < \deg p_\alpha$ . As  $(p_\alpha) = M(\alpha; T)$ , we also have  $p_\alpha q \in M(\alpha; T)$ , and thus

$$g(T)\alpha = q(T)p_\alpha(T)\alpha + r(T)\alpha = r(T)\alpha.$$

Hence,  $Z(\alpha; T) = \text{span}\{\alpha, T\alpha, \dots, T^{k-1}\alpha\}$ . We are left with proving that they are linearly independent.

Suppose they are not linearly independent for the sake of contradiction. Then there exist  $c_0, \dots, c_{k-1} \in F$  not all zero such that  $(\sum_{i=0}^{k-1} c_i T^i)\alpha = 0$ , which means  $g_0(x) = \sum_{i=0}^{k-1} c_i x^i \in M(\alpha; T)$  with  $\deg g_0 < \deg p_\alpha$ , violating the minimality of  $p_\alpha$ . Hence, they are linearly independent.

- (ii) Let  $m$  be the minimal polynomial of  $U$ . Then,  $m(U)\alpha = 0$  for all  $\alpha \in V$ , and thus  $p_\alpha \mid m$ .

Take any  $v \in Z(\alpha; T)$ . Then, there exists  $g \in F[x]$  so  $v = g(T)\alpha$ . Then,  $p_\alpha(U)v = g(T)p_\alpha(T)\alpha = 0$ . Hence,  $p_\alpha(U) = 0$ ; thus  $m \mid p_\alpha$ .  $\square$

**Note:-**

With respect to the ordered basis  $\mathcal{B} = \{\alpha; T\alpha; \dots; T^{k-1}\alpha\}$  for  $Z(\alpha; T)$ . Then,

$$[U]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{k-1} \end{bmatrix}$$

where  $p_\alpha(x) = \sum_{i=0}^{k-1} c_i x^i + x^k$ .

### Definition 7.1.3: Companion Matrix

The matrix  $[U]_{\mathcal{B}}$  above is called the *companion matrix* of  $p_\alpha$ .

## 7.2 Cyclic Decompositions and the Rational Form

### Definition 7.2.1: Complementary $T$ -invariant Subspace

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $T \in L(V)$ . Let  $W$  be a  $T$ -invariant subspace of  $V$ . If  $W'$  is a  $T$ -invariant subspace of  $V$  such that  $V = W \oplus W'$ , we call it a *complementary  $T$ -invariant subspace* of  $W$ .

### Definition 7.2.2: $T$ -admissible Subspace

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $T \in L(V)$ . We say a subspace  $W$  of  $V$  is  *$T$ -admissible* if

- (i)  $W$  is  $T$ -invariant and
- (ii)  $\forall f \in F[x], \forall \beta \in V, (f(T)\beta \in W \implies \exists \gamma \in W, f(T)\beta = f(T)\gamma)$ .

### Lemma 7.2.1

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $T \in L(V)$ . Suppose  $W$  and  $W'$  are  $T$ -invariant subspaces such that  $V = W \oplus W'$ . Then,  $W$  and  $W'$  are  $T$ -admissible.

**Proof.** The condition (i) is already true. Suppose  $f(T)\beta \in W$  where  $f \in F[x]$  and  $\beta \in V$ . We can write  $\beta = \gamma + \gamma'$  where  $\gamma \in W$  and  $\gamma' \in W'$ . Then,  $f(T)\beta = f(T)\gamma + f(T)\gamma'$ . As  $W$  and  $W'$  are  $T$ -invariant, we have  $f(T)\beta - f(T)\gamma = f(T)\gamma' \in W \cap W'$ . Hence,  $f(T)\beta = f(T)\gamma$ .  $\square$

### Notation 7.1

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $T \in L(V)$ . If  $T$  is the only subjective linear transform, we may write

- $f\alpha$  instead of  $f(T)(\alpha)$  for each  $f \in F[x]$  and  $\alpha \in V$ .
- $fW$  instead of  $f(T)(W)$  for each  $f \in F[x]$  and  $W \subseteq V$ .

### Lemma 7.2.2

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $T \in L(V)$ . Then, the following hold.

- $fZ(\alpha; T) = Z(f\alpha; T)$  for each  $\alpha \in V$  and  $f \in F[x]$ .
- If  $V = \bigoplus_{i=1}^k W_i$  where each  $W_i$  is  $T$ -invariant, then  $fV = \bigoplus_{i=1}^k fW_i$ .
- For  $\alpha, \gamma \in V$ , if  $\alpha$  and  $\gamma$  have the same  $T$ -annihilator, then  $f\alpha = f\gamma$  has the same  $T$ -annihilator. Therefore,  $\dim Z(f\alpha; T) = \dim Z(f\gamma; T)$ .

**Proof.**

- $fZ(\alpha; T) = \{fg\alpha \mid g \in F[x]\} = \{gf\alpha \mid g \in F[x]\} = Z(f\alpha; T)$
- It is evident that  $fV = \sum_{i=1}^k fW_i$ . Suppose  $\sum_{i=1}^k f\alpha_i = 0$  for some  $\alpha_i \in W_i$ . As  $f\alpha_i \in V_i$ , we have  $f\alpha_i = 0$  for all  $i \in [k]$ . Hence,  $fW_1, \dots, fW_k$  are independent.
- We have  $M(\alpha; T) = M(\gamma; T)$ , i.e.,  $\forall g \in F[x], (g\alpha = 0 \iff g\gamma = 0)$ . Hence,  $M(f\alpha; T) = \{g \in F[x] \mid gf\alpha = 0\} = \{g \in F[x] \mid gf\gamma = 0\} = M(f\gamma; T)$ .  $\square$

### Theorem 7.2.1 Cyclic Decomposition Theorem

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $T \in L(V)$ . Let  $W_0$  be a proper  $T$ -admissible subspace of  $V$ . Then, there exist  $\alpha_1, \dots, \alpha_r \in V \setminus \{0\}$  such that

- $V = W_0 \oplus \left( \bigoplus_{i=1}^r Z(\alpha_i; T) \right)$  and
  - $p_{i+1} \mid p_i$  for each  $i \in [r-1]$
- where  $p_i$  is the  $T$ -annihilator of  $\alpha_i$ . Furthermore,  $r$  and  $p_1, \dots, p_r$  are uniquely decided.

**Proof.** In this proof we denote the monic generator of  $S_T(\alpha; W)$  as  $s_T(\alpha; W)$  for conciseness.

**Claim 0.** For  $\alpha, \beta \in V$  and a  $T$ -invariant subspace  $W$  of  $V$ , if  $\alpha - \beta \in W$ , then  $S_T(\alpha; W) = S_T(\beta; W)$ . Moreover, if  $W$  is  $T$ -invariant, then  $W + Z(\alpha; T) = W + Z(\beta; T)$ .

Let  $\gamma \triangleq \alpha - \beta \in W$ . Then,  $g \in S_T(\alpha; W) \iff g\alpha \in W \iff g(\beta + \gamma) \in W \iff g\beta \in W \iff g \in S_T(\beta; W)$ .

Assuming  $W$  is  $T$ -invariant, we have, for each  $g\alpha \in Z(\alpha; T)$ ,  $g\alpha = g(\beta + \gamma) \in Z(\beta; T) + W$ ; hence  $Z(\alpha; T) + W \subseteq Z(\beta; T) + W$ .  $\checkmark$

**Claim 1.** For a proper  $T$ -admissible subspace  $W$  of  $V$ , there exists  $\alpha \in V \setminus W$  such that  $s_T(\alpha; W)\alpha = 0$ .

Take any  $\beta \in V \setminus W$ . Let  $f \triangleq s_T(\beta; W)$  so  $f\beta \in W$ . By  $T$ -admissibility,  $\exists \gamma \in W, f\beta = f\gamma$ . Let  $\alpha \triangleq \beta - \gamma$  so that  $f\alpha = 0$ . Moreover,  $S_T(\alpha; W) = S_T(\beta; W) = (f)$  as  $W$  is  $T$ -invariant.

Hence,  $f = s_T(\beta; W) = s_T(\alpha; W)$ . and  $f \in M(\alpha; T)$ , which implies  $(f) = S_T(\alpha; W) \subseteq M(\alpha; T)$ . Conversely, if  $g \in M(\alpha; T)$ , then  $g\alpha = 0 \in W$  and thus  $M(\alpha; T) \subseteq S_T(\alpha; W)$ ;  $f$  is the  $T$ -annihilator of  $\alpha$  as well.

**Claim 2.** Let  $W$  be a subspace of  $V$ . If  $s_T(\alpha; W)\alpha = 0$ , then  $S_T(\alpha; W) \subseteq M(\alpha; T)$  and  $W \cap Z(\alpha; T) = \{0\}$ .

It is easily shown that  $S_T(\alpha; W) \subseteq M(\alpha; T)$ . Suppose  $g\alpha \in W \cap Z(\alpha; T)$ . Then,  $g \in S_T(\alpha; W) \subseteq M(\alpha; T)$ , and thus  $g\alpha = 0$ .  $\checkmark$

**Claim 3.** For a proper  $T$ -invariant subspace  $W$  of  $V$ ,  $\beta \in \operatorname{argmax}_{\alpha \in V} \deg s_T(\alpha; W)$  exists. Moreover,  $W \cap \operatorname{argmax}_{\alpha \in V} \deg s_T(\alpha; W) = \emptyset$ . As a corollary,  $W + Z(\beta; T)$  is a  $T$ -invariant subspace of  $V$  which has  $W$  as its proper subspace.

If  $p$  is the characteristic polynomial of  $T$ , then  $p\alpha = 0 \in W$  for all  $\alpha \in V$  by Theorem 6.2.2, i.e.,  $p \in S_T(\alpha; T)$ . Therefore,  $\deg s_T(\alpha; W)$  is bounded above by  $\deg p = \dim V$ . Hence,  $A = \operatorname{argmax}_{\alpha \in V} \deg s_T(\alpha; W) \neq \emptyset$ , thus we may take  $\beta \in A$ .

If  $\beta \in W$ , we will have  $s_T(\alpha; W) = 1$  for all  $\alpha \in V$  and thus  $\alpha = s_T(\alpha; W)\alpha \in W$ , contradicting  $W \subsetneq V$ .  $\checkmark$

---

**Algorithm:** Construct  $\beta_1, \dots, \beta_r$  and  $W_1, \dots, W_r$

$i \leftarrow 0$ ;

**while**  $W_i \neq V$  **do**

    Take any  $\beta_{i+1} \in \operatorname{argmax}_{\alpha \in V} \deg s_T(\alpha; W_i)$ ;

$\triangleright$  well-defined by **Claim 3**

$W_{i+1} \leftarrow W_i + Z(\beta_{i+1}, W_i)$ ;

$i \leftarrow i + 1$ ;

---

This algorithm above eventually ends in at most  $\dim V$  loops until we have  $V = W_0 + \sum_{i=1}^r Z(\beta_i, W_{i-1})$  by **Claim 3**. Also, by the construction,  $W_k = W_{k-1} + Z(\beta_k, W_{k-1})$  for each  $k \in [r]$ , and each  $W_k$  is  $T$ -invariant.

$$W_k = W_0 + \sum_{i=1}^{k-1} Z(\beta_i; W_{i-1})$$

**Claim 4.** For each  $k \in [r]$  and  $\beta \in V$ , write  $f\beta = \beta_0 + \sum_{i=1}^{k-1} g_i\beta_i$  where  $f = s_T(\beta; W_{k-1})$ ,  $g_i \in F[x]$ ,  $\beta_0 \in W_0$ , and  $\beta_i \in W_i$  for each  $i \in [k-1]$ . Then,  $f \mid g_i$  for each  $i \in [k-1]$ , and  $\beta_0 = f\gamma_0$  for some  $\gamma_0 \in W_0$ .

Fix  $k \in [r]$  for now. By Theorem 4.1.2,  $g_i = f q_i + r_i$  for some  $q_i, r_i \in F[x]$  such that it is either  $r_i = 0$  or  $\deg r_i < \deg f$ . Let  $\gamma \triangleq \beta - \sum_{i=1}^{k-1} q_i\beta_i$ . Then, we have:

$$\begin{aligned} f\gamma &= f\beta - \sum_{i=1}^{k-1} f q_i\beta_i \\ &= (\beta_0 + \sum_{i=1}^{k-1} g_i\beta_i) - \sum_{i=1}^{k-1} (g_i - r_i)\beta_i \\ &= \beta_0 + \sum_{i=1}^{k-1} r_i\beta_i. \end{aligned}$$

Note that, by **Claim 0**,  $S_T(\gamma; W_{k-1}) = S_T(\beta; W_{k-1}) = (f)$ .

For the sake of contradiction, suppose  $r_i \neq 0$  for some  $i \in [k-1]$  and let  $j$  be the maximum among such  $i$  so  $f\gamma = \beta_0 + \sum_{i=1}^j r_i\beta_i$ . Let  $p \triangleq s_T(\gamma; W_{j-1})$ . As  $W_j \subseteq W_{k-1}$ , we have  $p \in S_T(\gamma; W_{k-1}) = (f)$ , i.e.,  $p = fg$  for some  $g \in F[x]$ . Then,

$$p\gamma = gf\gamma = g\beta_0 + \sum_{i=1}^{j-1} g r_i\beta_i + g r_j\beta_j.$$

Then,  $p\gamma \in W_{j-1}$  by the definition of  $p$  and  $g(\beta_0 + \sum_{i=1}^{j-1} r_i\beta_i) \in W_{j-1}$  as  $W_{j-1}$  is  $T$ -invariant. Hence, we have  $g r_j\beta_j \in W_{j-1}$ , i.e.,  $g r_j \in S_T(\beta_j; W_{j-1})$ . Hence, by the construction of  $\beta_j$ ,

$$\deg(g r_j) \underbrace{\geq}_{\text{by definition}} \deg s_T(\beta_j; W_{j-1}) \underbrace{\geq}_{\text{by construction of } \beta_j} \deg s_T(\gamma; W_{j-1}) = \deg p = \deg(fg).$$

Therefore,  $\deg r_j \geq \deg f$ , which is a contradiction. Hence,  $r_i = 0$  for all  $i \in [k-1]$ ;  $f \mid g_i$ .

Now, we are left with  $\beta_0 = f\gamma$ . By  $T$ -admissibility of  $W_0$ , there exists  $\gamma_0 \in W_0$  such that  $f\gamma_0 = f\gamma = \beta_0$ . ✓

Fix any  $k \in [r]$ . Let  $p_k \triangleq s_T(\beta_k; W_{k-1})$ . Then, by **Claim 4**,  $p_k\beta_k = p_k\gamma_0 + \sum_{i=1}^{k-1} p_k h_i \beta_i$  for some  $\gamma_0 \in W_0$  and  $h_i \in F[x]$ . Let  $\alpha_k \triangleq \beta_k - \gamma_0 - \sum_{i=1}^{k-1} h_i \beta_i$  so that  $p_k\alpha_k = 0$  and  $\alpha_k - \beta_k \in W_{k-1}$ . Then, by **Claim 0** and **Claim 2**, we have:

- $(p_k) = S_T(\beta; W_{k-1}) = S_T(\alpha_k; W_{k-1}) \subseteq M(\alpha_k; T)$
- $W_{k-1} \cap Z(\alpha_k; T) = \{0\}$
- $W_{k-1} + Z(\alpha_k; T) = W_{k-1} + Z(\beta_k; T)$

As  $k$  is arbitrary, we have

$$W_k = W_0 \oplus \left( \bigoplus_{i=1}^k Z(\alpha_i; T) \right).$$

Moreover, note that  $\alpha_1, \dots, \alpha_r$  retains the defining property of  $\beta_1, \dots, \beta_r$ , i.e.,  $\alpha_k \in \operatorname{argmax}_{\alpha \in V} \deg s_T(\alpha; W_{k-1})$ . **Claim 4** holds when  $\beta_1, \dots, \beta_r$  are replaced with  $\alpha_1, \dots, \alpha_r$ . Hence, applying the alternative version of **Claim 4** to the trivial equation

$$p_k\alpha_k = 0 \cdot 0 + \sum_{i=1}^{k-1} p_i\alpha_i,$$

we have  $p_k \mid p_i$  for each  $i \in [k-1]$ . The existence part of the theorem is now proven. ✓

Now, we shall show the uniqueness of such decomposition. Suppose  $V = W_0 \oplus \left( \bigoplus_{i=1}^s Z(\gamma_i; T) \right)$  is another cyclic decomposition where  $q_{k+1} \mid q_k$  for each  $k \in [s-1]$ . ( $q_i$  is the  $T$ -annihilator of  $\gamma_i$ .) Let  $S_T(V; W_0) \triangleq \{f \in F[x] \mid fV \subseteq W_0\}$  so  $S_T(V; W_0)$  is an ideal as  $W_0$  is  $T$ -invariant.

**Claim 5.**  $p_1$  and  $q_1$  are the monic generator of  $S_T(V; W_0)$ . Thus,  $p_1 = q_1$ .

As it is already  $S_T(V; W_0) \subseteq S_T(\alpha_1; W_0)$ , it is enough to show that  $p_1 \in S_T(V; W_0)$ . Take any  $\beta \in V$  and write  $\beta = \beta_0 + \sum_{i=1}^r f_i \alpha_i$  where  $\beta_0 \in W_0$  and  $f_i \in F[x]$ . Then,  $p_1\beta = p_1\beta_0 + \sum_{i=1}^r p_1 f_i \alpha_i$ . As  $p_i \mid p_1$  for each  $i \in [r]$ , we have  $p_1\alpha_i = 0$ . Hence,  $p_1\beta \in W_0$ ;  $p_1 \in S_T(V; W_0)$ . Similarly,  $q_1 \in S_T(V; W_0)$ . ✓

Now, we will prove the main question by induction.

**Claim 6.** Fix any  $1 \leq k < r$ . If  $s \geq k$  and  $p_i = q_i$  for each  $i \in [k]$ , then  $s > k$  and  $p_{k+1} = q_{k+1}$ .

As  $k < r$ , we have  $\dim W_0 + \sum_{i=1}^k \dim Z(\alpha_i; T) < \dim V$ . By (i) of Theorem 7.1.1,  $\dim W_0 + \sum_{i=1}^k \dim Z(\gamma_i; T) < \dim V$  as  $p_i = q_i$ . Hence,  $s > k$ . Now, we may discuss facts about  $q_{k+1}$ .

The two decompositions give

$$\begin{aligned} p_{k+1}V &= p_{k+1}W_0 \oplus \left( \bigoplus_{i=1}^k Z(p_{k+1}\alpha_i; T) \right) \quad \text{and} \\ p_{k+1}V &= p_{k+1}W_0 \oplus \left( \bigoplus_{i=1}^s Z(p_{k+1}\gamma_i; T) \right) \end{aligned}$$

together with (i) and (ii) of Lemma 7.2.2. (For the first representation, we only add up to  $k$  since  $p_{k+1}\alpha_i$  for all  $i \in \{k+1, \dots, r\}$ .) Now, we have

$$\begin{aligned} \dim(p_{k+1}V) &= \dim(p_{k+1}W_0) + \sum_{i=1}^k \dim Z(p_{k+1}\alpha_i; T) \\ &= \dim(p_{k+1}W_0) + \sum_{i=1}^s \dim Z(p_{k+1}\gamma_i; T) \\ &= \dim(p_{k+1}W_0) + \sum_{i=1}^k \dim Z(p_{k+1}\alpha_i; T) + \sum_{i=k+1}^s \dim Z(p_{k+1}\gamma_i; T), \end{aligned}$$

by (iii) of Lemma 7.2.2, and thus  $\dim Z(p_{k+1}\gamma_{k+1}; T) = 0$ , i.e.,  $p_{k+1}\gamma_{k+1} = 0$ . Thus,  $q_{k+1} \mid p_{k+1}$ . Similarly, we also have  $p_{k+1} \mid q_{k+1}$ , and thus  $p_{k+1} = q_{k+1}$ . ✓

Using **Claim 6**, we reach  $r \leq s$  and  $p_i = q_i$  for each  $i \in [r]$  by mathematical induction. By symmetry, we also have  $s \leq r$  and thus the theorem is proven. □

### Corollary 7.2.1

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $T \in L(V)$ . Let  $W$  be a  $T$ -invariant subspace of  $V$ . Then,  $W$  is  $T$ -admissible if and only if there exists another  $T$ -invariant subspace  $W'$  of  $V$  such that  $V = W \oplus W'$ .

**Proof.**  $(\Rightarrow)$  Theorem 7.2.1  $(\Leftarrow)$  Lemma 7.2.1 □

#### Note:-

- Every  $T \in L(V)$  has  $\alpha \in V$  such that  $T$ -annihilator of  $\alpha$  equals the minimal polynomial of  $T$ . ( $\alpha_1$  when  $W_0 = \{0\}$ .  $T$ -conductor of  $\alpha_1$  to  $W_0$  is  $T$ -annihilator of  $\alpha_1$  and it is the minimal polynomial.)
- If  $T \in L(V)$  has a cyclic vector, then characteristic polynomial of  $T$  equals the minimal polynomial of  $T$ .

### Corollary 7.2.2

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $T \in L(V)$ . Let  $V = \bigoplus_{i=1}^r Z(\alpha_i; T)$  be the cyclic decomposition obtained by applying Theorem 7.2.1 to  $W_0 = \{0\}$ .

- $p_1$ , the  $T$ -annihilator of  $\alpha_1$ , is the minimal polynomial of  $T$ .
- The characteristic polynomial of  $T$  equals  $\prod_{i=1}^r p_i$ .
- $T$  has a cyclic vector if and only if the characteristic polynomial and the minimal polynomial are identical.

**Proof.**

- $p_1$  is the monic generator of  $S_T(V; \{0\}) = \{g \in F[x] \mid \forall \alpha \in V, g\alpha = 0\}$  by **Claim 6** in the proof of Theorem 7.2.1. In other words,  $p_1$  is the minimal polynomial of  $T$ .
- Let  $T_i \triangleq T|_{Z(\alpha_i; T)}$  be a linear operator on  $Z(\alpha_i; T)$ . By Theorem 7.1.1,  $p_i$  is the characteristic polynomial of  $T_i$ . Hence,  $\prod_{i=1}^r p_i$  is the characteristic polynomial of  $T$ .
- $\exists \alpha \in V, T = Z(\alpha; T) \iff r = 1 \iff p_1 = (\text{the characteristic polynomial of } T)$  □  

$\underbrace{\iff}_{\text{uniqueness}}$

$\underbrace{\iff}_{(ii)}$

### Theorem 7.2.2 Generalized Cayley-Hamilton Theorem

Let  $V$  be a finite-dimensional vector space over  $F$  and let  $T \in L(V)$ . Let  $m$  and  $f$  be the minimal and the characteristic polynomial, respectively. Then,

- $m \mid f$ .
- $m$  and  $f$  have the same prime factors except possibly for multiplicities.
- Suppose  $m = \prod_{i=1}^k f_i^{r_i}$  and  $f = \prod_{i=1}^k f_i^{d_i}$  are the prime factorizations. Then,  $d_i = (\text{nullity } f_i(T)^{r_i}) / (\deg f_i)$ .

**Proof.**

- Theorem 6.2.2
- Apply Theorem 7.2.1 to  $W_0 = \{0\}$ . Then, we have  $V = \bigoplus_{i=1}^r Z(\alpha_i; T)$ . and  $p_{i+1} \mid p_i$  for each  $i \in [r-1]$  where  $p_i$  is the  $T$ -annihilator of  $\alpha_i$ . (ii) of Corollary 7.2.2 says  $f = \prod_{i=1}^r p_i$ . If a prime factor divides  $f$ , then it divides one of  $p_i$ , which divides  $p_1 = m$ . Hence,  $f$  and  $m$  has the same prime factors.
- Apply Theorem 6.7.1 with respect to  $V$  and  $T$ . Then, we have  $V = \bigoplus_{i=1}^k W_i$  where  $W_i = \ker f_i(T)^{r_i}$ . Take  $T_i = T|_{W_i}$  so  $f_i^{r_i}$  is the minimal polynomial of  $T_i$ . By (ii) and the supposition, the characteristic polynomial  $g_i$  of  $T_i$  is  $f_i^{d_i}$  where  $r_i \leq d_i$ . Hence,  $\dim W_i = \deg g_i = d_i \cdot \deg f_i$ . Therefore,  $d_i = (\dim W_i) / (\deg f_i) = (\text{nullity } f_i(T)^{r_i}) / (\deg f_i)$ .



□

### Corollary 7.2.3

Let  $V$  be an  $n$ -dimensional vector space over  $F$  and let  $N \in L(V)$  be nilpotent. Then, the characteristic polynomial of  $N$  is  $x^n$ .

**Proof.** As  $N^k = 0$  for some  $k \in \mathbb{N}$ , the minimal polynomial is  $x^r$  for some  $r$ . Hence, by (ii) of Theorem 7.2.2, we have  $x^n$  as the characteristic polynomial of  $N$ . □

### Definition 7.2.3: Rational Form

Let  $F$  be a field. A matrix  $A \in F^{n \times n}$  is said to be in a *rational form* if

$$A = \begin{bmatrix} A_1 & \mathbf{0}_{k_1 \times k_2} & \cdots & \mathbf{0}_{k_1 \times k_r} \\ \mathbf{0}_{k_2 \times k_1} & A_2 & \cdots & \mathbf{0}_{k_2 \times k_r} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{k_r \times k_1} & \mathbf{0}_{k_r \times k_2} & \cdots & A_r \end{bmatrix}$$

where

- $A_i$  is a companion matrix of  $p_i \in F[x]$  with  $\deg p_i = k_i$  for each  $i \in [r]$  and
- $p_{i+1} \mid p_i$  for each  $i \in [r-1]$ .

### Theorem 7.2.3

Let  $F$  be a field. For any  $B \in F^{n \times n}$ , there uniquely exists a matrix in a rational form which is similar to  $B$ .

**Proof.** Let  $F^n = \oplus_{i=1}^r Z(\alpha_i; B)$  be a cyclic decomposition obtained by applying Theorem 7.2.1 to  $W_0 = \{0\}$ . By (i) of Theorem 7.1.1,  $\mathcal{B}_i = \{\alpha_1, \dots, T^{k_i-1}\alpha_i\}$  is a basis for  $B_i = Z(\alpha_i; T)$  where  $k_i = \dim Z(\alpha_i; B)$ . Let  $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_r)$  be a basis for  $F^n$ . Then,

$$[B]_{\mathcal{B}} = \begin{bmatrix} [B_1]_{\mathcal{B}_1} & 0 & \cdots & 0 \\ 0 & [B_2]_{\mathcal{B}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & [B_r]_{\mathcal{B}_r} \end{bmatrix}$$

is in a rational form since  $[B_i]_{\mathcal{B}_i}$  is the companion matrix of  $p_i$ , the  $T$ -annihilator of  $\alpha_i$ .

Now, suppose that  $A$  is a  $n \times n$  matrix over  $F$  which is similar to  $B$  and is in a rational form. Let  $A$  be composed of  $r$  companion matrices of  $p_1, \dots, p_r$  where  $p_{i+1} \mid p_i$  for each  $i \in [r-1]$ . Then, we have  $\beta_1, \dots, \beta_r \in F^n$  where  $p_i$  is the  $A$ -annihilator of  $\beta_i$ , which is the  $B$ -annihilator of  $\beta_i$  as well. Hence,  $F^n = \oplus_{i=1}^r Z(\beta_i; B)$ . By Theorem 7.2.1,  $r$  and  $p_1, \dots, p_r$  are uniquely determined. □

## 7.3 The Jordan Form

### Example 7.3.1

Let  $V$  be an  $n$ -dimensional vector space over  $F$  and let  $N \in L(V)$  be nilpotent. Applying Theorem 7.2.1 to  $N$ , we have  $V = \oplus_{i=1}^r Z(\alpha_i; N)$  and  $p_{i+1} \mid p_i$  for each  $i \in [r-1]$  where  $p_i$  is the  $T$ -annihilator of  $\alpha_i$ .  $p_1$ , the minimal polynomial of  $N$ , is of the form  $x^k$

where  $1 \leq k \leq n$ . Therefore, the divisibility relation simply asserts  $k = k_1 \geq k_2 \geq \dots \geq k_r$  where  $p_i(x) = x^{k_i}$ . So the rational form of  $N$  is

$$\begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{bmatrix}$$

where  $A_i$  is the companion matrix of  $x^{k_i}$ , which is the  $k_i \times k_i$  matrix

$$A_i = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

Hence, there are total  $p(n)$  nilpotent  $n \times n$  linear operators on  $V$  that are not similar to each other where  $p(n)$  denotes the number of partitions of  $n$ .

### Lemma 7.3.1

Let  $V$  be an  $n$ -dimensional vector space over  $F$  and let  $N \in L(V)$  be nilpotent. Let  $V = \oplus_{i=1}^r Z(\alpha_i; T)$  be the cyclic decomposition obtained by applying Theorem 7.2.1 to  $W_0 = \{0\}$ . Let  $k_i \triangleq \deg p_i$  where  $p_i$  is the  $T$ -annihilator of  $\alpha_i$ . Then,  $\mathcal{B} = \{N^{k_1-1}\alpha_1, \dots, N^{k_r-1}\alpha_r\}$  is a basis for  $\ker N$ .

**Proof.** The fact that  $\mathcal{B} \subseteq \ker N$  is evident as  $p_i = x^{k_i}$ . Also,  $\mathcal{B}$  is linearly independent as each vector is in independent subspaces  $(Z(\alpha_i; N))$ .

Take any  $\alpha \in \ker N$ . Then,  $\alpha = \sum_{i=1}^r f_i \alpha_i$  for some  $f_i \in F[x]$ . WLOG,  $\deg f_i < \deg p_i = k_i$ . Then,  $N\alpha = \sum_{i=1}^r Nf_i(N)\alpha_i = 0$ , and thus  $Nf_i(N)\alpha_i = 0$  for each  $i \in [r]$ . Therefore,  $p_i \mid x f_i$ . Moreover, as  $p_i = x^{k_i}$ , and as  $\deg f_i < k_i$ , we have  $f_i = c_i x^{k_i-1}$  for some  $c_i \in F$ . Then, we have  $\alpha = \sum_{i=1}^r c_i N^{k_i-1} \alpha_i$ ;  $\ker N \subseteq \text{span } \mathcal{B}$ .  $\square$

### Definition 7.3.1: Jordan Form

Let  $V$  be an  $n$ -dimensional vector space over  $F$  and let  $T \in L(V)$  be triangulable. Let  $f(x) = \prod_{i=1}^k (x - c_i)^{d_i}$  and  $m(x) = \prod_{i=1}^k (x - c_i)^{r_i}$  be the characteristic polynomial and the minimal polynomial of  $T$ , respectively.

#### Note:-

Assume  $T$  is triangulable.

**Step 1** Let  $f(x) = \prod_{i=1}^k (x - c_i)^{d_i}$  be the characteristic polynomial. Let  $m(x) = \prod_{i=1}^k (x - c_i)^{r_i}$  be the minimal polynomial. ( $1 \leq r_i \leq d_i$ ) Take  $W_i = \ker(T - c_i I)^{r_i}$ . By Theorem 6.7.1, we have  $V = \oplus_{i=1}^k W_i$ . Let  $T_i = T|_{W_i}$ . Each  $W_i$  is  $T$ -invariant and  $T_i$ 's minimal polynomial is  $(x - c_i)^{r_i}$ .

**Step 2** For each  $W_i$ , let  $N_i = T_i - c_i I \in L(W_i)$  so  $N_i$  is nilpotent. So,  $T_i = N_i + c_i I$ . We consider, for each  $W_i$ , the cyclic decomposition of  $W_i$  with respect to  $N_i$ .

For each  $W$ , we have  $W = Z(\alpha_1; N) \oplus \dots \oplus Z(\alpha_s; N)$ . Take  $\mathcal{B}_j = \{\alpha_j, N\alpha_j, \dots, N^{k_j-1}\alpha_j\}$ .

So  $\left[N|_{Z(\alpha_j; N)}\right]_{\mathcal{B}_j}$  is  $\delta_{i-1,j}$ . So  $\left[T|_{Z(\alpha_j; N)}\right]_{\mathcal{B}_j}$  is  $\delta_{i-1,j} + c_i \delta_{ij}$ . (< fix this) (This is called a *elementary Jordan block*.)  $\mathcal{B}^i = \bigcup_j \mathcal{B}_j$

# Chapter 8

## Inner Product Spaces

### 8.1 Inner Products

#### Definition 8.1.1: Inner Product

Fix the field  $F$  to  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . An inner product  $(-, -)$  on  $V$  is a function  $(-, -): V \times V \rightarrow F$  satisfying

- (i)  $(-, -)$  is linear over  $F$ .
- (ii)  $(\beta, \alpha) = \overline{(\alpha, \beta)}$
- (iii) If  $\alpha \neq 0$ ,  $(\alpha, \alpha) > 0$ .

#### Note:-

- If  $F = \mathbb{R}$ , (i) and (ii) say that  $(-, -)$  is also linear over  $F$ . Thus, an inner product is symmetric and bilinear.
- If  $F = \mathbb{C}$ ,  $(\alpha, c\gamma) = \overline{c}(\alpha, \gamma)$ , i.e.,  $(-, -)$  is sesqui-linear.
- If  $F = \mathbb{C}$ ,  $(\alpha, \alpha) = \overline{(\alpha, \alpha)}$ , i.e.,  $(\alpha, \alpha) \in \mathbb{R}$ .

#### Example 8.1.1

- For  $[x_i], [y_i] \in \mathbb{C}^n$ , the inner product defined by  $([x_i], [y_i]) = \sum_{i=1}^n x_i \overline{y_i}$  is called the *standard inner product*.
- $F = \mathbb{R}$ , let  $A \in \mathbb{R}^{n \times n}$  such that  $x^T A x > 0$  for all  $x \in \mathbb{R}^n \setminus \{0\}$ . We say  $A$  is positive definite. The function  $(x, y) = x^T A y$  is an inner product.

#### Definition 8.1.2: Conjugate Transpose

For an  $n \times n$  matrix  $A$  over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ , the  $n \times n$  matrix  $A^*$  over  $F$  defined by

$$(A^*)_{ij} = \overline{(A)_{ji}}$$

for each  $i, j \in [n]$  is called the *conjugate transpose* of  $A$ .

#### Definition 8.1.3: Self-adjoint Matrix

A matrix  $A \in \mathbb{C}^{n \times n}$  is said to be *self-adjoint* if  $A^* = A$ .

### Theorem 8.1.1

Let  $V = \mathbb{C}^n$ . Let  $(-, -): V \times V \rightarrow \mathbb{C}$  be an inner product. Then, there exists a self-adjoint positive definite matrix  $A \in \mathbb{C}^{n \times n}$  such that  $(x, y) = x^* A y$ .

**Proof.** Choose a basis, e.g., the standard basis  $\{e_1, \dots, e_n\}$ . Let  $(A)_{ij} = (e_i, e_j)$ . Let  $x, y \in \mathbb{C}^n$  and write  $x = \sum_{i=1}^n x_i e_i$  and  $y = \sum_{j=1}^n y_j e_j$ . Then,

$$\begin{aligned} (x, y) &= \left( \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \overline{y_j} (A)_{ij} x_i = y^* A x. \end{aligned}$$

□

### Corollary 8.1.1

Let  $V = \mathbb{R}^n$ . Let  $(-, -): V \times V \rightarrow \mathbb{R}$  be an inner product. Then, there exists a symmetric positive definite matrix  $A \in \mathbb{R}^{n \times n}$  such that  $(x, y) = x^* A y$ .

### Example 8.1.2

Let  $V = \mathcal{C}([a, b], \mathbb{C})$ . Define, for  $f, g \in V$ ,  $(f, g) = \int_a^b f(t) \overline{g(t)} dt$ . That is an inner product on  $V$ .

## 8.2 Inner Product Space

### Definition 8.2.1: Inner Product Space

A vector space  $V$  over  $F = \mathbb{R}$  or  $F = \mathbb{C}$  equipped with a specified inner product is called an *inner product space*.

### Notation 8.1: Norm

We write

$$\|v\| = \sqrt{(v, v)}.$$

This is called a *norm* of  $v$ .

### Theorem 8.2.1

Let  $V$  be an inner product space.

- (i)  $\|c\alpha\| = |c| \cdot \|\alpha\|$  for all  $c \in F$  and  $\alpha \in V$ .
- (ii)  $\|\alpha\| > 0$  for all  $\alpha \in V \setminus \{0\}$ .
- (iii)  $|(\alpha, \beta)| \leq \|\alpha\| \cdot \|\beta\|$  for all  $\alpha, \beta \in V$ . (*Cauchy-Schwarz*)
- (iv)  $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$  for all  $\alpha, \beta \in V$ . (*Triangle Inequality*)

**Proof.**

(i) ✓

(ii) ✓

(iii) If  $\alpha = 0$ , we have nothing to prove, so suppose  $\alpha \neq 0$ . Let

$$\beta^\parallel \triangleq \frac{(\beta, \alpha)}{(\alpha, \alpha)} \alpha \quad \text{and} \quad \beta^\perp \triangleq \beta - \frac{(\beta, \alpha)}{(\alpha, \alpha)} \alpha.$$

Then,  $(\beta^\perp, \alpha) = 0$  and  $\beta^\parallel + \beta^\perp = \beta$ . Let  $c = (\beta, \alpha)/(\alpha, \alpha)$ . We have

$$0 \leq \|\beta^\perp\|^2 = (\beta - c\alpha, \beta^\perp) = (\beta, \beta) - |c|^2(\alpha, \alpha) = \|\beta\|^2 - \frac{|(\alpha, \beta)|^2}{\|\alpha\|^2}.$$

Rearranging the inequality gives the result.

$$(iv) \quad \|\alpha + \beta\|^2 = (\alpha, \alpha) + (\alpha, \beta) + (\beta, \alpha) + (\beta, \beta) \leq \|\alpha\|^2 + 2\|\alpha\| \cdot \|\beta\| + \|\beta\|^2 = (\|\alpha\| + \|\beta\|)^2 \quad \square$$

**Note:-**

For  $\alpha, \beta \in V \setminus \{0\}$ , we define the *angle* between  $\alpha$  and  $\beta$  be  $\theta \in \mathbb{R}$  such that

$$\cos \theta = \frac{(\alpha, \beta)}{\|\alpha\| \cdot \|\beta\|}.$$

**Lemma 8.2.1 Polarization Identities**

Let  $V$  be an inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ .

(i) If  $F = \mathbb{R}$ ,

$$\forall \alpha, \beta \in V, (\alpha, \beta) = \frac{1}{4}\|\alpha + \beta\|^2 - \frac{1}{4}\|\alpha - \beta\|^2.$$

(ii) If  $F = \mathbb{C}$ ,

$$\forall \alpha, \beta \in V, (\alpha, \beta) = \frac{1}{4} \sum_{k=1}^4 i^k \|\alpha + i^k \beta\|^2.$$

**Proof.** We have the fact that, for every  $\alpha, \beta \in V$ ,

$$\|\alpha + \beta\|^2 = \|\alpha\|^2 + 2\Re(\alpha, \beta) + \|\beta\|^2.$$

Also, we have

$$\Im(\alpha, \beta) = \Re[-i(\alpha, \beta)] = \Re(\alpha, i\beta),$$

and thus  $(\alpha, \beta) = \Re(\alpha, \beta) + i\Re(\alpha, i\beta)$ . They can be shown from these equations.  $\square$

**Definition 8.2.2: Orthogonality**

Let  $V$  be an inner product space.

- For  $\alpha, \beta \in V$ , we say  $\alpha$  and  $\beta$  are *orthogonal* if  $(\alpha, \beta) = 0$ .
- For  $S \subseteq V$ , we say  $S$  is *orthogonal* if  $\alpha, \beta \in S$  and  $\alpha \neq \beta$ , then  $(\alpha, \beta) = 0$ .
- $\{\alpha_1, \dots, \alpha_n\} \subseteq V$  is said to be *orthonormal* if it is orthogonal and if  $\|\alpha_i\| = 1$ .

**Theorem 8.2.2**

Let  $V$  be an inner product space. Then, every orthogonal subset  $S$  of  $V \setminus \{0\}$  is linearly independent.

**Proof.** Take any distinct  $\alpha_1, \dots, \alpha_k \in S$ . Suppose  $\sum_{i=1}^k c_i \alpha_i = 0$  for some  $c_i$ . Then, for each  $k \in [n]$ ,

$$0 = \left( \sum_{i=1}^k c_i \alpha_i, \alpha_k \right) = \sum_{i=1}^k c_i (\alpha_i, \alpha_k) = c_k (\alpha_k, \alpha_k).$$

Since  $(\alpha_k, \alpha_k) > 0$ , it must be  $c_k = 0$ .  $\square$

**Note:-**

If  $S = \{\alpha_1, \dots, \alpha_m\} \subseteq V$  is orthogonal, we may explicitly write every  $\beta \in \text{span } S$  in the form of

$$\beta = c_1 \alpha_1 + c_2 \alpha_2 + \dots + c_m \alpha_m$$

by setting  $c_i = \frac{(\beta, \alpha_i)}{\|\alpha_i\|^2}$ .

**Theorem 8.2.3 Gram–Schmidt**

Let  $V$  be an inner product space. Suppose  $\{\beta_1, \dots, \beta_n\}$  is a linearly independent subset of  $V$ . Then, there exists an orthogonal set  $\{\alpha_1, \dots, \alpha_n\}$  of vectors such that, for each  $k \in [n]$ ,  $\{\alpha_1, \dots, \alpha_k\}$  is a basis of  $\text{span}\{\beta_1, \dots, \beta_k\}$ .

**Proof.** Take  $\alpha_1 = \beta_1$ . Then, for each  $k \in \{2, \dots, n\}$ , set

$$\alpha_k \triangleq \beta_k - \sum_{i=1}^{k-1} \frac{(\beta_k, \alpha_i)}{\|\alpha_i\|^2} \alpha_i.$$

Then,  $\{\alpha_1, \dots, \alpha_n\}$  satisfies the condition. □

**Corollary 8.2.1**

Every finite dimensional inner product space has an orthonormal basis.

**Proof.** Normalize vectors from Theorem 8.2.3. □

**Definition 8.2.3: Orthogonal Complement**

Let  $V$  be an inner product space. Let  $W$  be a subset of  $V$  and  $\beta \in V$ . The subspace of  $V$

$$W^\perp = \{\beta \in V \mid \forall \alpha \in W, (\alpha, \beta) = 0\}$$

is called the *orthogonal complement* of  $W$ .

**Definition 8.2.4: Best Approximation**

Let  $V$  be an inner product space. Let  $W$  be a subspace of  $V$  and  $\beta \in V$ . A *best approximation* of  $\beta$  to  $W$  is a vector  $\alpha \in W$  such that

$$\forall \gamma \in W, \|\beta - \alpha\| \leq \|\beta - \gamma\|.$$

**Theorem 8.2.4**

Let  $V$  be an inner product space. Let  $W$  be a subspace of  $V$  and  $\beta \in V$ .

- (i)  $\alpha \in W$  is a best approximation of  $\beta$  to  $W$  if and only if  $\beta - \alpha \in W^\perp$ .
- (ii) If a best approximation of  $\beta$  to  $W$  exists, then it is unique.
- (iii) If  $W$  is finite-dimensional and  $\{\alpha_1, \dots, \alpha_n\}$  is an orthonormal basis for  $W$ , then  $\alpha = \sum_{k=1}^n (\beta, \alpha_k) \alpha_k$  is the best approximation of  $\beta$  to  $W$ .

**Proof.** For any  $\gamma \in V$ , we have  $\beta - \gamma = (\beta - \alpha) + (\alpha - \gamma)$  and thus

$$\|\beta - \gamma\|^2 = \|\beta - \alpha\|^2 + 2\Re(\beta - \alpha, \alpha - \gamma) + \|\alpha - \gamma\|^2.$$

(i) ( $\Leftarrow$ ) Suppose  $\beta - \alpha \in W^\perp$ . Then, for any  $\gamma \in W$ , as  $\alpha - \gamma \in W$ ,

$$\|\beta - \gamma\|^2 = \|\beta - \alpha\|^2 + \|\alpha - \gamma\|^2 \geq \|\beta - \alpha\|^2.$$

( $\Rightarrow$ ) Suppose  $\alpha$  is a best approximation of  $\beta$  to  $W$ . For any  $\gamma \in W$ , we have

$$2\Re(\beta - \alpha, \alpha - \gamma) + \|\alpha - \gamma\|^2 \geq 0$$

by the inequality above and  $\|\beta - \gamma\|^2 \geq \|\beta - \alpha\|^2$ . In other words, for any  $\tau \in W$ ,

$$2\Re(\beta - \alpha, \tau) + \|\tau\|^2 \geq 0.$$

Putting

$$\tau = -\frac{(\beta - \alpha, \alpha - \gamma)}{\|\alpha - \gamma\|^2}(\alpha - \gamma),$$

we have

$$-2\frac{|(\beta - \alpha, \alpha - \gamma)|^2}{\|\alpha - \gamma\|^2} + \frac{|(\beta - \alpha, \alpha - \gamma)|^2}{\|\alpha - \gamma\|^2} \geq 0,$$

which holds if and only if  $(\beta - \alpha, \alpha - \gamma) = 0$ .

(ii) Directly follows from the inequality in the proof of (i)-( $\Leftarrow$ ).

(iii) A calculation shows that  $\beta - \alpha$  is orthogonal to every vector in  $W$ . Hence, by (i),  $\alpha$  is the best approximation of  $\beta$  to  $W$ .  $\square$

### Definition 8.2.5: Orthogonal Projection

Let  $V$  be an inner product space. Let  $W$  be a subspace of  $V$ .

- Let  $\beta \in V$  and  $\alpha \in W$ . If  $\beta - \alpha \in W^\perp$ , then  $\alpha$  is called the *orthogonal projection of  $\beta$  on  $W$* .
- If every vector in  $V$  has an orthogonal projection on  $W$ , the mapping that assigns to each vector in  $V$  its orthogonal projection on  $W$  is called the *orthogonal projection of  $V$  on  $W$* .

#### Note:-

If a subspace  $W$  of  $V$  is finite-dimensional, then there always exists the orthogonal projection of  $V$  on  $W$ .

### Corollary 8.2.2

Let  $V$  be an inner product space. Let  $W$  be a finite-dimensional subspace of  $V$ . Let  $E: V \rightarrow W$  be the orthogonal projection of  $V$  on  $W$ . Then,  $\beta \mapsto \beta - E\beta$  is the orthogonal projection of  $V$  on  $W^\perp$ .

**Proof.** Take any  $\beta \in V$ . Then,  $\beta - E\beta \in W^\perp$  by Theorem 8.2.4 (i). Also, for any  $\gamma \in W^\perp$ ,  $\beta - \gamma = E\beta + (\beta - E\beta - \gamma)$ , and thus

$$\|\beta - \gamma\|^2 = \|E\beta\|^2 + \|\beta - E\beta - \gamma\|^2 \geq \|\beta - (\beta - E\beta)\|^2$$

as  $E\beta \in W$  and  $\beta - E\beta - \gamma \in W^\perp$ . Hence,  $\beta - E\beta$  is the best approximation of  $\beta$  to  $W^\perp$ .  $\square$

### Theorem 8.2.5

Let  $V$  be an inner product space. Let  $W$  be a finite-dimensional subspace of  $V$ . Let  $E$  be the orthogonal projection of  $V$  on  $W$ . Then  $E$  is an idempotent linear transformation



of  $V$  to  $W$ ,  $\ker E = W^\perp$ , and  $V = W \oplus W^\perp$ .

**Proof.** By the definition, we have

$$\forall \beta \in W, E\beta = \beta,$$

which means  $E(E\beta) = E\beta$  for all  $\beta \in V$ . ✓

Take any  $\alpha, \beta \in V$ . Then,  $(c\alpha + \beta) - (cE\alpha + E\beta) = c(\alpha - E\alpha) + (\beta - E\beta) \in W^\perp$ . Also,  $cE\alpha + E\beta \in W$ . By (ii) of Theorem 8.2.4,  $E(c\alpha + \beta) = cE\alpha + E\beta$ . ✓

The equation  $\beta = E\beta + (\beta - E\beta)$  shows that  $V = W + W^\perp$ . For independence, for any  $\alpha \in W \cap W^\perp$ , it must be  $(\alpha, \alpha) = 0$ ; thus  $\alpha = 0$ . ✓ □

### Notation 8.2: Projection

Let  $V$  be an inner product space. Let  $W$  be a finite-dimensional subspace of  $V$ . Then, as  $V = W \oplus W^\perp$ , we have a unique expression  $v = w + w'$  where  $w \in W$  and  $w' \in W^\perp$ . Then, we write

$$w = \text{proj}_W v \quad \text{and} \quad w' = \text{proj}_{W^\perp} v.$$

#### Note:-

Let  $\{\alpha_1, \dots, \alpha_k\}$  and  $\{\alpha_{k+1}, \dots, \alpha_n\}$  be orthogonal bases for  $W$  and  $W^\perp$ , respectively. Then,

$$\text{proj}_W \beta \triangleq \sum_{i=1}^k \frac{(\beta, \alpha_i)}{\|\alpha_i\|^2} \alpha_i \quad \text{and} \quad \text{proj}_{W^\perp} \beta \triangleq \sum_{i=k+1}^n \frac{(\beta, \alpha_i)}{\|\alpha_i\|^2} \alpha_i.$$

## 8.3 Linear Functionals and Adjoint

### Theorem 8.3.1

Let  $V$  be an  $n$ -dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . Then,

$$\forall f \in V^*, \exists! \beta \in V, f(-) = (-, \beta).$$

In other words, the linear transform  $V \rightarrow V^*$  defined by  $\beta \mapsto (-, \beta)$  is an isomorphism.

**Proof.** Choose an orthogonal basis  $\{\alpha_1, \dots, \alpha_n\}$  of  $V$ . Let  $\beta = \sum_{i=1}^n c_i \alpha_i$ . Then,

$$f(\alpha_j) = (\alpha_j, \sum_{i=1}^n c_i \alpha_i) = \sum_{i=1}^n \overline{c_i} (\alpha_j, \alpha_i) = \overline{c_j} \|\alpha_j\|^2.$$

Hence, if such  $\beta$  exists, it must be

$$\beta = \sum_{i=1}^n \frac{\overline{f(\alpha_i)}}{\|\alpha_i\|^2} \alpha_i.$$

So, it is only left to check if this  $\beta$  actually works. It can be easily checked by reversing the computation above. □

#### Note:-

Theorem 8.3.1 essentially says that  $V$  and  $V^*$  are naturally related if  $V$  is equipped with an inner product.

### Theorem 8.3.2

Let  $V$  be an  $n$ -dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ .

$$\forall T \in L(V), \exists! T^* \in L(V), \forall \alpha, \beta \in V, (T\alpha, \beta) = (\alpha, T^*\beta).$$

**Proof.** Fix  $\beta \in V$ . For the linear functional  $\alpha \mapsto (T\alpha, \beta)$ , Theorem 8.3.1 guarantees the unique existence of  $\beta' \in V$  such that  $(T(-), \beta) = (-, \beta')$ . Hence, we may define  $T^* \in L(V)$  by  $\beta \mapsto \beta'$  so we have  $(T\alpha, \beta) = (\alpha, T^*\beta)$  for all  $\alpha \in V$ . (It is a linear operator.)  $\square$

#### Note:-

Such  $T^*$  is called a (*Hermitian*) *transpose*.

### Theorem 8.3.3

Let  $V$  be an  $n$ -dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . Let  $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$  be an orthonormal basis for  $V$ . Let  $T \in L(V)$  and  $A = [T]_{\mathcal{B}}$ . Then,  $(A)_{ij} = (T\alpha_j, \alpha_i)$  for each  $i, j \in [n]$ .

**Proof.** Take any  $\alpha \in V$  and write  $\alpha = \sum_{i=1}^n (\alpha, \alpha_i) \alpha_i$ . Hence,

$$T(\alpha_j) = \sum_{i=1}^n (T\alpha_j, \alpha_i) \alpha_i.$$

By definition, we have  $T(\alpha_j) = \sum_{i=1}^n A_{ij} \alpha_i$ . Therefore,  $A_{ij} = (T\alpha_j, \alpha_i)$ .  $\square$

### Corollary 8.3.1

Let  $V$  be an  $n$ -dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . Let  $T \in L(V)$ . Let  $\mathcal{B}$  be an orthonormal basis for  $V$ . Then,  $[T^*]_{\mathcal{B}} = ([T]_{\mathcal{B}})^*$ , i.e., the matrix representation of  $T^*$  equals the Hermitian transpose of the matrix representation of  $T$ .

**Proof.** Let  $A = [T]_{\mathcal{B}}$  and  $B = [T^*]_{\mathcal{B}}$ . By the Theorem 8.3.3, we have

$$(A)_{ij} = (T\alpha_j, \alpha_i) \quad \text{and} \quad (B)_{ij} = (T^*\alpha_j, \alpha_i).$$

Then,  $\overline{(B)_{ij}} = (\alpha_i, T^*\alpha_j)$ , and thus  $\overline{(B)_{ji}} = (\alpha_j, T^*\alpha_i)$ . However, we have  $(T\alpha_j, \alpha_i) = (\alpha_j, T^*\alpha_i)$  by the definition of  $T^*$ , and thus  $(A)_{ij} = \overline{(B)_{ji}}$ .  $\square$

#### Note:-

Corollary 8.3.1 essentially says that all transposes we discussed represents the same object when orthonormal basis is given.

### Definition 8.3.1: Hermitian Operator

Let  $V$  be an  $n$ -dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . Let  $T \in L(V)$ . We say  $T$  is *Hermitian* (or, *self-adjoint*) if  $T = T^*$ . If  $F = \mathbb{R}$ , we sometimes call it *symmetric*.

### Definition 8.3.2: Isometry

Let  $V$  and  $W$  be inner product spaces over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ .  $T \in L(V, W)$  is said to *preserve the inner products*, or is called an *isometry*, if

$$\forall \alpha, \beta \in V, (\alpha, \beta) = (T\alpha, T\beta).$$

### Definition 8.3.3: Isomorphism of Inner Product Spaces

Let  $V$  and  $W$  be inner product spaces over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ .  $T \in L(V, W)$  is called an *isomorphism of inner product spaces* if it is an isomorphism of vector spaces and it preserves the inner products.

#### Theorem 8.3.4

Let  $V$  and  $W$  be finite-dimensional inner product spaces over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . Let  $T \in L(V, W)$ . TFAE.

- (i)  $T$  preserves the inner products.
- (ii)  $T$  is an isomorphism of inner product spaces.
- (iii) For any arbitrary orthonormal basis  $\mathcal{B}$  of  $V$ ,  $T(\mathcal{B})$  is an orthonormal basis for  $W$ .
- (iv) For some orthonormal basis  $\mathcal{B}$  for  $V$ ,  $T(\mathcal{B})$  is an orthonormal basis for  $W$ .

**Proof.** ((i)  $\Rightarrow$  (ii)) We need to show that  $T$  is an isomorphism of vector spaces. Take any  $\alpha \in \ker T$ . Then,  $0 = (T\alpha, T\alpha) = (\alpha, \alpha) = \|\alpha\|^2$ ; hence  $\alpha = 0$ . Therefore,  $T$  is injective, and thus is an isomorphism.

((ii)  $\Rightarrow$  (iii)) Let  $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$  be an orthonormal basis for  $V$ . Hence,  $(T\alpha_i, T\alpha_j) = (\alpha_i, \alpha_j) = \delta_{ij}$ ; and thus  $T(\mathcal{B}) = \{T\alpha_1, \dots, T\alpha_n\}$  is an orthonormal basis for  $W$ .

((ii)  $\Rightarrow$  (i) and (iii)  $\Rightarrow$  (iv) are obvious.

((iv)  $\Rightarrow$  (i)) Let  $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$  be an orthonormal basis for  $V$  such that  $T(\mathcal{B})$  is also an orthonormal basis for  $W$ . Then,  $(T\alpha_i, T\alpha_j) = \delta_{ij} = (\alpha_i, \alpha_j)$ .  $\square$

#### Theorem 8.3.5

Let  $V$  and  $W$  be inner product spaces over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . For  $T \in L(V, W)$ ,

$$T \text{ preserves the inner products} \iff \forall \alpha \in V, \|T\alpha\| = \|\alpha\|.$$

**Proof.** ( $\Rightarrow$ )  $\|T\alpha\|^2 = (T\alpha, T\alpha) = (\alpha, \alpha) = \|\alpha\|^2$ .

( $\Leftarrow$ ) (Polarization identity)  $(\alpha, \beta)$  can be expressed in terms of  $\|\alpha \pm \beta\|$  and  $\|\alpha \pm i\beta\|$ . (Exercise)  $\square$

### Definition 8.3.4: Unitary Operator

Let  $V$  be an inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ .  $T \in L(V)$  is called a *unitary operator* if it is an isomorphism of inner product spaces.

#### Theorem 8.3.6

Let  $V$  be an inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . For  $U \in L(V)$ ,

$$U \text{ is unitary} \iff U^* \text{ exists and } UU^* = U^*U = \text{Id}.$$

**Proof.** ( $\Rightarrow$ ) For  $U$  being an isomorphism, there exists the inverse  $U^{-1} \in L(V)$  of  $U$ . Then, for all  $\alpha, \beta \in V$ ,  $(U\alpha, \beta) = (U\alpha, UU^{-1}\beta) = (\alpha, U^{-1}\beta)$ . Hence,  $U^{-1} = U^*$ .

( $\Leftarrow$ )  $U$  is invertible and  $U^{-1} = U^*$ . Then, for all  $\alpha, \beta \in V$ ,  $(U\alpha, U\beta) = (\alpha, U^*U\beta) = (\alpha, \beta)$ . Hence,  $U$  preserves the inner product. Moreover,  $U$  is an isomorphism of vector spaces as well. Hence,  $U$  is unitary.  $\square$

**Note:-**

For finite-dimensional inner product spaces, for any  $T: V \rightarrow V$ , we have  $T^*$ . If the dimension is infinite, the existence of  $T^*$  is not immediate.

**Definition 8.3.5: Unitary Matrix**

Let  $A$  be an  $n \times n$  matrix over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . We say  $A$  is a *unitary matrix* if

$$AA^* = A^*A = I_n.$$

**Theorem 8.3.7**

Let  $V$  be a finite-dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . For  $U \in L(V)$ ,

$U$  is unitary  $\iff \exists$  orthonormal basis  $\mathcal{B}$  for  $V$ ,  $[U]_{\mathcal{B}}$  is a unitary matrix.

**Theorem 8.3.8**

Let  $V$  be a finite-dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . Let  $U_1, U_2 \in L(V)$  be unitary operators. Then,  $U_1 \circ U_2$  is also unitary.

**Proof.** For all  $\alpha, \beta \in V$ ,  $(U_1 U_2 \alpha, U_1 U_2 \beta) = (U_2 \alpha, U_2 \beta) = (\alpha, \beta)$ . □

**Theorem 8.3.9**

Let  $V$  be a finite-dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . If  $U \in L(V)$  is unitary, then  $U^{-1}$  is unitary as well.

**Proof.**  $U^{-1} = U^*$ , and  $UU^* = U^*U = \text{Id}$ . □

**Definition 8.3.6**

Let  $V$  be an  $n$ -dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . Let

$$U(V) \triangleq \{ T \in L(V) \mid T \text{ is unitary} \}.$$

$$O(n) \triangleq \{ A \in F^{n \times n} \mid AA^t = A^t A = I \}.$$

- When  $F = \mathbb{C}$ , the group  $(U(V), \circ) \simeq (U(\mathbb{C}^n), \circ)$  is called the  *$n$ -th unitary group*.
- When  $F = \mathbb{R}$ , the group  $(O(n), \cdot)$  is called the *real orthogonal group*.
- When  $F = \mathbb{C}$ , the group  $(O(n, \mathbb{C}), \cdot)$  is called the *complex orthogonal group*.
- $SU(n) = \{ A \in U(n) \mid \det A = 1 \}$  is called the *special unitary group*.
- $SO(n) = \{ A \in O(n) \mid \det A = 1 \}$  is called the *special orthogonal group*.

## 8.4 Normal Operators

**Definition 8.4.1: Normal Operator**

Let  $V$  be a finite-dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ .  $T \in L(V)$  is *normal* if  $TT^* = T^*T$ .

**Note:-**

If we can find an orthonormal basis  $\mathcal{B}$  for  $V$  such that the vectors in  $\mathcal{B}$  are also characteristic vectors of  $T$ , then  $[T]_{\mathcal{B}}$  is diagonal, and thus  $TT^* = T^*T$ . In other words,  $T$  having an orthonormal basis consisting of characteristic vectors implies  $T$  is normal. In this section, we prove the converse.

**Theorem 8.4.1**

Let  $V$  be an  $n$ -dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . Let  $T \in L(V)$  be normal. Then, for a characteristic vector  $\alpha \in V$ ,  $c \in F$  is a eigenvalue of  $T$  associated to  $\alpha$  if and only if  $\bar{c}$  is an eigenvalue of  $T^*$  associated to  $\alpha$ .

**Proof.** For any  $\beta \in V$  and any normal  $U \in L(V)$ , we have

$$\|U\beta\|^2 = (U\beta, U\beta) = (\beta, U^*U\beta) = (\beta, UU^*\beta) = (U^*\beta, U^*\beta) = \|U^*\beta\|^2.$$

Also, for any  $c \in F$ ,  $T - cI$  is normal. Hence, if  $c \in F$  is an eigenvalue and  $\alpha \in V$  is an eigenvector associated to  $c$ , we have  $\|(T^* - \bar{c}I)\alpha\| = \|(T - cI)\alpha\| = 0$ , and thus  $\bar{c}$  is an eigenvalue and  $\alpha$  is an eigenvector associated to  $\bar{c}$ .  $\square$

**Theorem 8.4.2**

Let  $V$  be an  $n$ -dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . Let  $T \in L(V)$ . Suppose  $\mathcal{B}$  is an orthonormal basis for  $V$  such that  $[T]_{\mathcal{B}}$  is upper triangular. Then,

$$T \text{ is normal} \iff [T]_{\mathcal{B}} \text{ is diagonal.}$$

**Proof.** Let  $A \triangleq [T]_{\mathcal{B}}$ . Then  $A^* = [T^*]_{\mathcal{B}}$  by Corollary 8.3.1.

( $\Leftarrow$ ) We apparently have  $AA^* = A^*A$ , i.e.,  $TT^* = T^*T$ .  $T$  is normal.

( $\Rightarrow$ )  $A$  is upper triangular. Let  $a_{ij} = (A)_{ij}$  so  $a_{ij} = 0$  if  $i > j$ . Let  $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ . We immediately have that  $T\alpha_1 = a_{11}\alpha_1$ , and thus  $T^*\alpha_1 = \overline{a_{11}}\alpha_1$  by Theorem 8.4.1. On the other hand, we have  $T^*\alpha_1 = \sum_{j=1}^n \overline{a_{1j}}\alpha_j$ . Hence,  $a_{1j} = 0$  for  $1 < j \leq n$ .

Repeating the argument above, we get  $a_{ij} = 0$  if  $i < j$ . Hence,  $A$  is diagonal.  $\square$

**Lemma 8.4.1**

Let  $V$  be an  $n$ -dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ . Let  $T \in L(V)$ . Let  $W$  be a  $T$ -invariant subspace of  $V$ . Then,  $W^\perp$  is  $T^*$ -invariant.

**Proof.** Take any  $\beta \in W^\perp$ . For all  $\alpha \in W$ , we have  $(\alpha, T^*\beta) = (T\alpha, \beta) = 0$  as  $W$  is  $T$ -invariant. Hence,  $T^*\beta \in W^\perp$ .  $\square$

**Theorem 8.4.3**

Let  $V$  be an  $n$ -dimensional inner product space over  $\mathbb{C}$ . Let  $T \in L(V)$ . Then, there exists an orthonormal basis  $\mathcal{B}$  for  $V$  such that  $[T]_{\mathcal{B}}$  is upper triangular.

**Proof.** We prove by induction on  $\dim V$ . Suppose that  $n = \dim V > 1$  and that the theorem holds for any inner product space with dimension less than  $n$ .

Since  $\mathbb{C}$  is algebraically closed, there exists an eigenvalue  $c \in \mathbb{C}$  and an eigenvector  $\alpha \in V$  such that  $T^*\alpha = c\alpha$ . WLOG,  $\|\alpha\| = 1$ . Then, as  $\text{span}\{\alpha\}$  is  $T^*$ -invariant, we have  $W \triangleq (\text{span}\{\alpha\})^\perp$  is  $T$ -invariant by Lemma 8.4.1. By the induction hypothesis on  $W$  and

$T|_W$ , there exists an orthonormal basis  $\mathcal{B}' = \{\alpha_1, \dots, \alpha_{n-1}\}$  for  $W$  such that  $[T|_W]_{\mathcal{B}'}$  is upper triangular. Then,  $\mathcal{B} = \mathcal{B}' \cup \{\alpha\}$  is a basis for  $V$  and

$$[T]_{\mathcal{B}} = \begin{bmatrix} [T|_W]_{\mathcal{B}'} & * \\ 0 & * \end{bmatrix}$$

is upper triangular. □

#### Corollary 8.4.1

Let  $V$  be an  $n$ -dimensional inner product space over  $\mathbb{C}$ . Let  $T \in L(V)$  be normal. Then, there exists an orthonormal basis for  $V$  consisting of eigenvectors of  $T$ .

#### Corollary 8.4.2

Let  $A \in \mathbb{C}^{n \times n}$ .

- (i) There exists a unitary matrix  $P \in U(n)$  such that  $P^{-1}AP$  is upper triangular.
- (ii) If  $A$  is normal, then  $P^{-1}AP$  is diagonal. In other words,  $A$  is *unitarily diagonalizable*.

#### Example 8.4.1

Let  $V$  be an  $n$ -dimensional inner product space over  $\mathbb{C}$ . Let  $T \in L(V)$ .

- (i) If  $T$  is Hermitian, then  $T$  is normal.
- (ii) If  $T$  is unitary, then  $T$  is normal.

*End.*