KAIST
2023F MAS212 Linear Algebra
Summary

Choi Pooreunhaneul / Choi Pooha
University: KAIST
Department: Mathematical Science
Github: Choi Pooha

December 26, 2023

CONTENTS

Chapter 1

Linear Equations

1.1 Fields

Note:-

- **This is summary note for 2023 Fall, KAIST MAS212 - Linear Algebra course taught by Prof. Jinhyun Park.**
- **I assumed that you are familiar enough to MAS109 - Introduction to Linear Algebra course, or just some notation used in basic linear algebra like matrix and row operations.**
- **I also assumed that you are familiar enough to some mathematical logic symbols.**
- **We used Kenneth Hoffman / Ray Kunze - Linear Algebra 2nd ed.**

Definition 1.1.1: Field

Algebraic structure $\mathbb{F}$ satisfying given properties are called field:

1. Addition is commutative: $\forall\{x, y\} \subset \mathbb{F}\ (x + y = y + z)$
2. Addition is associative: $\forall\{x, y, z\} \subset \mathbb{F}\ (x + (y + z) = (x + y) + z)$
3. $\forall x \in \mathbb{F}\ \exists! 0 \in \mathbb{F}\ (x + 0 = x)$
4. $\forall x \in \mathbb{F}\ \exists!(-x) \in \mathbb{F}\ (x + (-x) = 0)$
5. Multiplication is commutative: $\forall\{x, y\} \subset \mathbb{F}\ (xy = yx)$
6. Multiplication is associative: $\forall\{x, y, z\} \subset \mathbb{F}\ (x(yz) = (xy)z)$
7. $\forall x \in \mathbb{F}\ \exists! 1 \in \mathbb{F}\ (x1 = x)$
8. $\forall x \in \mathbb{F}\ \exists! x^{-1} = 1/x \in \mathbb{F}\ (xx^{-1} = 1)$
9. $\forall\{x, y, z\} \in \mathbb{F}\ (x(y + z) = xy + xz)$

1.2 Systems of Linear Equations

This Chapter is Intentionally Skipped at Lectures.

1.3 Matrices and Elementary Row Operations

This Chapter is Intentionally Skipped at Lectures.

1.4 Row-Reduced Echelon Matrices

This Chapter is Intentionally Skipped at Lectures.

1.5 Matrix Multiplication

Definition 1.5.1: Matrix Multiplication

$C := [C_{ij}]$. $C_{ij} := \sum_{r=1}^{n} A_{ir} B_{rj}$.

Theorem 1.5.1

Matrix multiplication is associative, but not commutative.

1.6 Invertible Matrices

Definition 1.6.1: Invertible Matrices

P is invertible $\iff \exists! Q \ (PQ = QP = I)$.

Chapter 2

Vector Spaces

2.1 Vector Spaces

Definition 2.1.1: Vector Spaces

A vector space consists of the following:

1. field F of scalars
2. a set V of objects called vectors
3. $\forall \{\alpha, \beta, \gamma\} \subset V$, a rule called vector addition holds:
   - addition is commutative: $\alpha + \beta = \beta + \alpha$
   - addition is associative: $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$
   - $\exists ! 0 \in V \ (\alpha + 0 = \alpha)$
   - $\exists ! (-\alpha) \in V \ (\alpha + (-\alpha) = 0)$
4. $\forall \{\alpha, \beta\} \subset V \ \forall \{c_1, c_2\} \subset F$, a rule called scalar multiplication holds:
   - $1\alpha = \alpha$
   - $(c_1 c_2)\alpha = c_1(c_2\alpha)$
   - $c_1(\alpha + \beta) = c_1\alpha + c_1\beta$
   - $(c_1 + c_2)\alpha = c_1\alpha + c_2\alpha$

Definition 2.1.2: Linear Combinations

$\alpha \in V$ is said to be linear combination of the vectors $\alpha_1, \ldots, \alpha_n \in V$ if $\exists c_1, \ldots, c_n \in F$ s.t.

$$\alpha = c_1\alpha_1 + \cdots + c_n\alpha_n = \sum_{i=1}^{n} c_i\alpha_i$$

2.2 Subspaces

Definition 2.2.1: Subspaces

$W \subset V$ is called subspace if W satisfies vector space axioms.

Theorem 2.2.1

$((V\text{:f.d.v.s}/F) \wedge (\{0\} \subsetneq W \subset V)) \Rightarrow (W \text{ is subspace} \iff \forall\{\alpha, \beta\} \in V \;\; \forall c \in F \;\; (c\alpha + \beta \in V))$.

Proof. We have to check: $W \neq \emptyset \Rightarrow \exists w \in W \Rightarrow 0 \in W$. □

Theorem 2.2.2

$\{W_i\} :=$ collection of subspaces of F-v.s. V. Let $W := \cap W_i$. then W is also subspace.

Proof. All W_i has 0, thus $0 \in \cap W_i$, which implies $W \neq \emptyset$. Let $v_1, v_2 \in W$, $c \in F$. Then $\forall v_1, v_2 \; (v_1, v_2 \in W_i)$. Since W_i is subspace, $cv_1 + v_2 \in W_i$ for all i, thus also in W. □

Definition 2.2.2: Span

$V : F$-v.s. $S \subset V :=$ any nonempty subset. The span(S) is the intersetion of all subspaces of V that contains S.

Theorem 2.2.3

span(S) is set of All linear combination of S/F.

Proof. $W :=$ span(S) and let L be set of all lin. comb. of S/F. Then obviously, $L \subset W$ because $S \subset WW$ and W is subspace.

Conversely, note that $S \subset L$. If we prove L is subspace, then since $S \subset L$, $W =$ span$(S) \subset L$. Then L is apparently a subsapce. Thus $W = L$. □

2.3 Bases and Dimensions

Definition 2.3.1: Linearly Independent

$V : F$-v.s., and take S as subset of V. We say S is linearly independent if $\exists \alpha_1, \ldots, \alpha_n \in S$ and $c_1, \ldots, c_n \in F$, not all zero, s.t. $c_1\alpha_1 + \cdots + c_n\alpha_n = 0$ has nontrivial solution. If S is not linaerly dependent, we say it is linearly independent.

Theorem 2.3.1

$V : F$-v.s. $\alpha_1, \ldots, \alpha_n$ are linearly independent $\iff \forall i \in [n] \, \forall c_i \in F \;\; ((c_1\alpha_1 + \cdots + c_n\alpha_n = 0) \Rightarrow c_1 = c_2 = \cdots = c_n = 0$.

Proof. Exercise! □

Definition 2.3.2: Basis

V : F-v.s. A basis of V is a subset $S \subset V$ s.t. S is lin. indep. and $\text{span}(S) = V$.

Definition 2.3.3: Finite Dimensional

If basis S has property $|S| < \infty$, we say V is finite dimensional vector space.

Theorem 2.3.2

V : F-v.s. that is spanned by $\{\beta_1, \dots, \beta_n\} \subset V$. Then any lin. indep. set of vec. in V is finite and card. is no bigger than n.

Proof. E.T.S. that every subset S with more than n vec. are lin. dep. Suppose $S = \{\alpha_1, \dots, \alpha_m\}$, for distinct vec. with $m \geq n$. Since $\{\beta_1, \dots, \beta_n\}$ spans V, for each $1 \leq j \leq m$, $\alpha_j = \sum_{i=1}^{n} A_{ij}\beta_i$. Let $x_1, \dots, x_m \in F$ be arbitrary chosen. Then $x_1\alpha_1 + \cdots + x_m\alpha_m = \sum_{j=1}^{m} x_j\alpha_j = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} A_{ij}x_j\right)\beta_i$. Consider the system $[A_{ij}][\mathbf{x}^T] = 0$. This has at least 1 free variable, which leads system has nontrivial solution. □

Corollary 2.3.1

V: F-v.s. that has finite spanning set. Then any two basis of V have same card.

. Apply Theorem 2.3.2 to both side of two different basis. □

Lemma 2.3.1

$W \subsetneq V$ be finite dim. v.s. Then $\dim(W) < \dim(V)$.

Proof. Let S_0 be a basis of W. S_0 is lin. indep., so can enlarged it to get a basis of V. Since W is propersubset of V, $\exists v \in V \backslash W$. Take $S_1 = S_0 \cup \{v\}$, and repeat this. finite dimensional condition of V implies this algorithm terminates in finite times, and thus we can conclude $\dim(W) < \dim(V)$. □

Theorem 2.3.3

$W_1, W_2 \subset V$: finite v.s. Then $W_1 + W_2$ is a finitie dim. v.s. and $\dim(W_1) + \dim(W_2) = \dim(W_1 + W_2) - \dim(W_1 \cap W_2)$.

Proof. Choose $\{\alpha_1, \dots, \alpha_d\}$ a basis for $W_1 \cap W_2$. We can extend this into W_1 and W_2's basis. Take $\{\alpha_1, \dots, \alpha_d, \beta_{d+1}, \dots, \beta_a\}$ be basis for W_1 and $\{\alpha_1, \dots, \alpha_d, \gamma_{d+1}, \dots, \gamma_b\}$ be basis for W_2.

Claim 2.3.1

$\alpha_1, \dots, \alpha_d, \beta_{d+1}, \dots, \beta_a, \gamma_{d+1}, \dots, \gamma_b$ is a basis for $W_1 + W_2$.

Proof. Suppose for arbitrary lin. indep. set B, $\text{span}(B) = W_1 + W_2$. Let $x \in W_1 + W_2$. Then $x = w_1 + w_2$ where $w_1 \in \text{span}\{\alpha, \beta\}$ and $w_2 \in \text{span}\{\alpha, \gamma\}$, thus $x \in \text{span}\{\alpha, \beta, \gamma\}$. On the other hand, each vec. in B is already in $W_1 + W_2$. Thus $\text{span}(B) = W_1 + W_2$. □

Claim 2.3.2

This B is lin. indep.

Proof. Suppose we have $\sum a_i\alpha_i + \sum b_j\beta_j + \sum c_k\gamma_k = 0$ for alal scalars are 0. Then $\sum a_i\alpha_i + \sum b_j\beta_j = -\sum c_k\gamma_k$. Thus $\sum c_k\gamma_k \in W_1 \cap W_2$ where $\{\alpha_1, \ldots, \alpha_d\}$ is basis for $W_1 \cap W_2$ and γ are indep. with α. Thus $\forall k \in \mathbb{N}$ $(c_k = 0)$. Simarly, we can see that all scalars are 0. Thus B is indep. □

This two claim leads $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$. □

2.4 Coordinates

Definition 2.4.1: Coordinates (Ordered Basis)

An ordered basis for F-v.s. V is a sequence of vec. that forms a basis.

Lemma 2.4.1

V : f.d.v.s./F. Suppose $B = \{v_1, \ldots, v_n\}$ is an ordered basis of V. Then for each $x \in V$, $\exists!$ expression of the form $x = x_1v_1 + \cdots + x_nv_n$ for some $x_i \in F$.

Proof. Existence of expression of form is trivial since B is basis of V.

For uniqueness, suppose we have two expression. Then indepence condition of each v_i leads these expression have exactly same coefficients. □

Definition 2.4.2: Coordinate Matrix

V : f.d.v.s./F, B be ordered basis. We define $[x]_B = [x_1\, x_2 \cdots x_n]^T$ the coordinate matrix of x w.r.t. the basis B.

Theorem 2.4.1

V : f.d.v.s./F, B and B' be two different ordered basis of V. Then $\exists!$ invertible mat. P s.t. $\forall x \in B$, $[x]_B = P[x]_{B'}$, also $[x]_{B'} = P^{-1}[x]_B$.

Proof. Let $B := \{\alpha_1, \ldots, \alpha_n\}$ and $B' := \{\beta_1, \ldots, \beta_n\}$. For $\beta_j \in B'$, since B is a basis, $\beta_j = \sum_{i=1}^n P_{ij}\alpha_i$ and this P_{ij} are uniquely decided. Let $P := [P_{ij}]$. Let $x \in V$. Write $[x]_B = [x_1 \ldots x_n]^T$, $[x]_{B'} = [x'_1 \ldots x'_n]^T$. Then $x = \sum_i \left(\sum_j x'_j P_{ij}\right)\alpha$. By uniqueness, we can derive $[x]_B = P[x']_B$. Since B and B' are lin. indep., x=0 implies $[x]_B = [x]_{B'} = 0$. Thus P is invertible. □

2.5 Summary of Row-Equivalence

This Chapter is Intentionally Skipped at Lectures

2.6 Computations Concerning Subspace

This Chapter is Intentionally Skipped at Lectures

Chapter 3

Linear Transformations

3.1 Linear Transformations

Definition 3.1.1: Linear Transformation

$T : V_1 \to V_2$ for v.s. V_1, V_2 is funciton called linear transformation if this function satisfies $T(cx_1 + x_2) = cT(x_1) + T(x_2)$ for $x_i \in V_i$, $c \in F$.

Exercise 3.1.1

If T is a linear trans., then $T(0) = 0$.

Proof. $T(0) + T(0+0) = 2T(0)$. □

Exercise 3.1.2 If T is a linear trans., then $T(-x) = -T(x)$.

Theorem 3.1.1

V, W : f.d.v.s./F, $\{\alpha_1, \ldots, \alpha_n\}$ be basis of V and $\{\beta_1, \ldots, \beta_m\}$ be any given subset of W. Then $\exists! T : V \to W$ s,t, $T(\alpha_i) = \beta_i$.

Proof. Define $T_0(x_1\alpha_1 + \cdots + x_n\alpha_n) := \sum_{i=1}^{n} x_i\beta_i$. This is lin. trans. Thus existence is proven.

For uniqueness, if there if another U s.t. $U(\alpha_i) = \beta_i$, then $U(\sum x_i\alpha_i) = \sum x_i U(\alpha_i) = \sum x_i\beta_i = T_0(\sum x_i\alpha_i)$. Thus $U = T_0$. □

Definition 3.1.2: Null Space and Range

$T : V \to W$: lin. trans. of v.s./F. $N(T) \subset V$, $R(T) \subset W$ where $N(T) := \{v \in V \mid Tv = \}$ and $R(T) := \{w \in W \mid \exists v \in V \ (w = T(v))\}$.

Definition 3.1.3

$\text{nullity}(T) := \dim_F(N(T))$, $\text{rank}(T) := \dim_F(R(T))$.

Theorem 3.1.2

V : f.d.v.s./F, $T : V \to W$: lin. trans. Then $\text{rank}(T) + \text{nullity}(T) = \dim(V)$.

Proof. Begin with $N(T)$. Choose basis $\{v_1, \dots, v_k\}$ of $N(T)$ and choose $v_{k+1}, \dots, v_n \in V$ s.t. $\{v_1, \dots, v_n\}$ is a basis of V.

Claim 3.1.1

$T(v_{k+1}), \dots, T(v_n)$ is a basis of $R(T)$.

Proof. For linear independence, suppose $c_{k+1}T(v_{k+1}) + \cdots + c_n T(v_n) = 0$. Then $T(c_{k+1}v_{k+1} + \cdots + c_n v_n) = 0$, so $c_{k+1}v_{k+1} + \cdots + c_n v_n \in N(T)$. Since $\{v_1, \dots, v_k\}$ is a basis of $N(T)$, $c_{k+1}v_{k+1} + \cdots + c_n v_n = a_1 v_1 + \cdots + a_k v_k$. Since $\{v_1, \dots, v_n\}$ is basis, those are lin. indep. Thus all coefficients are 0, thus $T(v_{k+1}), \dots, T(v_n)$ are indep. □

Claim 3.1.2

$\operatorname{span}\{T(v_{k+1}), \dots, T(v_n)\} = R(T)$

Proof. Exercise! □

Thus $\dim(R(T)) = n - k$. □

Theorem 3.1.3

For $m \times n$ mat. A, row rank is equal to column rank.

Proof. $V := F^n$ and $W := F^m$. $T : V \to W$ is lin. trans. Then col. rank = dim. of spans of col. = $\dim(R(T))$ = $\operatorname{rank}(T)$. Also, $\operatorname{nullity}(T) = \dim(N(T)) = n - \operatorname{rank}(T)$ = number of rows with leading 1's in RREF = number of cols. with leading 1's in RREF = dim. of col. space of A. Thus row rank is equal to col. rank. □

3.2 The Algebra of Linear Transformations

Definition 3.2.1: $L(V, W)$

$L(V, W)$ is set of all lin. trans. from V to W.

Theorem 3.2.1

$V, W : F$-v.s. Then $L(V, W)$ is itself vec. space over F.

Proof. Let $T, U \in L(V, W)$. Define $T + U : V \to W$ by $(T + U)(v) = T(v) + U(v)$.

Claim 3.2.1

$cT + U \in L(V, W)$

Proof. $(cT + U)(av_1 + v_2) = cT(av_1 + v_2) + U(av_1 + v_2)$ where both T and U is lin. trans. Thus trivially it is lin. trans. □

□

Theorem 3.2.2

$V : n$-dim. v.s./F, $W : m$-dim. v.s./F. Then $\dim_F(L(V, W)) = nm$.

Proof. Suppose $B = \{\alpha_1, \ldots, \alpha_n\}$ is basis of V, $B' = \{\beta_1, \ldots, \beta_m\}$ is basis of W. For each (p,q) where $1 \leq p \leq m$ and $1 \leq q \leq r$, define $E^{p,q}(\alpha_i) = 0$ if $i \neq q$ and β_p if $i = q$. Then these are lin. indep. trans. $V \to W$ and they span $L(V,W)$. □

Lemma 3.2.1

$U \circ T$ is a lin. trans. in $L(V,Z)$ where $U : V \to W$ and $T : W \to Z$.

Proof. Exercise! □

Definition 3.2.2: Endomorphism (Linear Operator)

For the case $T : V \to V$, we say T is an endomorphism or linear operator.

Definition 3.2.3

$T : V \to W$ be lin. trans. Then

- one-to-one or injective if $T(v) = 0 \Rightarrow v = 0$. (nonsingular)
- onto or surjective if $T(V) = W$
- T is invertible if $\exists U : W \to V$ s.t. $U \circ T = T \circ U = Id$

Exercise 3.2.1

T is injective and surjective $\iff$ T is invertible.

Exercise 3.2.2

$T : V \to W$ is a nonsingular lin. trans. Then any lin. indep. subset S of V is sent to lin. indep. set $T(S)$.

Exercise 3.2.3

Suppose $T : V \to W$ is invertible. Then $\dim(V) = \dim(W)$ for f.d.v.s. V and W.

Theorem 3.2.3

Suppose V, W as f.d.v.s./F and $\dim(V) = \dim(W)$. Let $T : V \to W$ be a lin. trans. TFAE:

i) T is invertible

ii) T is nonsingular, i.e., T is injective

iii) T is onto, i.e., T is surjective

Proof. $\mathrm{rank}(T) + \mathrm{nullity}(T) = n$. T is nonsingular $\iff$ $\mathrm{nullity}(T) = 0$ $\iff$ $\mathrm{rank}(T) = n$ $\iff$ $R(T) = W$ $\iff$ T is onto. □

Definition 3.2.4: General linear Group

$G =$ invertible endo. on V. with inverse $\circ$. Then $G = GL(V)$ is the general linear group of V.

Definition 3.2.5: Group

If some algebraic structure is associative with identity, we say this algebraic structure is group.

3.3 Isomorphism

Definition 3.3.1: Isomorphism

$V, W : F$—v.s. We say a lin. trans. $T : V \to W$ is an isomorphism if T is an invertible lin. trans.

Theorem 3.3.1

V : n-d.v.s./F. Then V is isomorphic to F^n ($V \simeq F^n$).

Proof. $B := \{\alpha_1, \ldots, \alpha_n\}$ is basis of V. Define $T : V \to F^n$, i.e., $v \mapsto [v]_B$.

Claim 3.3.1

This is isomorphism $\iff$ T is injective.

Proof. Suppose $T(v) = 0$. Then $v = 0$. □

□

3.4 Representation of Transformation by Matrices

Theorem 3.4.1

$V, W : F$-v.s. and B, B' be basis, where $T : V \to W$ be lin. trans. Then $\exists ! m \times n$ mat. A. s.t. $[Tv]_{B'} = A[v]_B$.

Theorem 3.4.2

V, W, Z : f.d.v.s./F, B, B', B'' be basis. Let $U \circ T : V \to Z$ be lin. trans. If $A_1 = [T]_{B,B'}$ and $A_2 = [T]_{B',B''}$, then $[U \circ T]_{B,B''} = A_2 \circ A_1$.

Theorem 3.4.3

T : endo. on f.d.v.s.V/F, where B_1, B_2 be two different basis of V. Let P be mat. s.t. $[v]_{B_1} = P[v]_{B_2}$. Then $[T]_{B_2} = P^{-1}[T]_{B_1} P$.

Definition 3.4.1: Similar

We say M and N are similar if $\exists$ invertible P s.t. $N = P^{-1} M P$.

3.5 Linear Functionals

Definition 3.5.1: Linear Functional

$V : F$-v.s. A lin. trans. $T : V \to F$ is called a linear functional.

Example 3.5.1

Definite integral and functions, especially constant function are linear functional.

Definition 3.5.2: Dual Vector Space

$V : F$-v.s. We normally write $V^* = L(V, F)$ the dual vector space of V.

Note:-

For finite dimensional V, $\dim(V^*) = \dim(V)$. But if V is infinite dimensional, $\dim(V^*)$ can be extremely large.

Lemma 3.5.1

$V : n$-d.v.s./F. Let $\{\alpha_1, \ldots, \alpha_n\}$ be a basis of V. Define $f \in V^*$ by declaring $f_i(\alpha_j) = \delta_{ij}$. Then $\{f_1, \ldots, f_n\}$ is basis of V^*.

Proof. Because $\dim(V^*) = \dim(V) = n$, E.T.S. that $f_1, \ldots, f_n$ are lin. indep. Suppose $\exists c_1 f_1 + \cdots + c_n f_n = 0$ for some $c_i \in F$ in V^*. Since $f_i(\alpha_j) = \delta_{ij}$, we can derive $c_j f_j(\alpha_j) = 0$. Thus $c_1 = \ldots = c_n = 0$, which implies $\{f_1, \ldots, f_n\}$ is basis. $\square$

Definition 3.5.3: The Dual Basis

$\{f_1, \ldots, f_n\} \subset V^*$ is called the dual basis of the basis $\{\alpha_1, \ldots, \alpha_n\}$ of V.

Lemma 3.5.2

$V : n$-d.v.s./F. $\{\alpha_1, \ldots, \alpha_n\}$ is basis of V. Let $\{f_1, \ldots, f_n\}$ is the dual basis. Then

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

ii) For each $v \in V$ $v = \sum_{i=1}^n f_i(v) \alpha_i$

Proof. i): Since $f \in \operatorname{span}\{f_1, \ldots, f_n\}$, $\exists$ expression $f = \sum_{i=1}^n x_i f_i$ for some $x_i \in F$. Evaluate at $\alpha_j : f(\alpha_j) = x_j$.

ii): Since $v \in \operatorname{span}\{\alpha_1, \ldots, \alpha_n\}$, $\exists$ expression $v = \sum_{i=1}^n y_i \alpha_i$. Apply the dual basis. $\square$

Note:-

$V : n$-d.v.s./F. Let $f \in V^*$. Suppose $f \neq 0$ and $f : V \to F$ be surjective. $N_f := N(f)$. We know $\dim(N(f)) + \dim(R(f)) = \dim(V)$. Since $\dim(R(f)) = 1$, $\dim(N(f)) = n - 1$.

Definition 3.5.4: Hyperspace

V : f.d.v.s./F. subspace W which has property $\dim(W) = \dim(V) - 1$ is called hyperspace.

Definition 3.5.5: Annihilator

V : F-v.s. S be a nonempty subspace. The annihilator of S, $S^\circ = Ann(S)$ is defined to be $S^\circ := \{f \in V^* \mid \forall \alpha \in S \ (f(\alpha) = 0)\}$.

Exercise 3.5.1

$Ann(S)$ is subspace of V^*.

Example 3.5.2

If $S = \{0\}$, then $Ann(S) = V^*$.

Example 3.5.3

If $S = V$, then $Ann(S) = \{0\}$.

Theorem 3.5.1

V : n-d.v.s./F, and W be subspace. Then $\dim(W) + \dim(W^\circ) = \dim(V) = n$.

Proof. $k := \dim(W)$ with $\{\alpha_1, \dots, \alpha_n\} \subset W$. Choose $\alpha_{k+1}, \dots, \alpha_n \in V$ s.t. $\{\alpha_1, \dots, \alpha_n\}$ is basis of V. Let $\{f_1, \dots, f_k, f_{k+1}, \dots, f_n\}$ be the dual basis.

Claim 3.5.1

$\{f_{k+1}, \dots, f_n\}$ is a basis of W°

Proof. Let's see if $f_{k+1}, \dots, f_n \in W^\circ$. Indeed, by the constructure of the dual basis, all f_i for $i \geq k+1$ vanishes on α_i for $1 \leq i \leq k$. Thus $f_{k+1}, \dots, f_n \in W^\circ$.

Lin. indep. is obvious since this is part of basis of V^*. □

Claim 3.5.2

$\text{span}\{f_{k+1}, \dots, f_n\} = W^\circ$

Proof. $f \in W^\circ \subset V$. So $f = \sum_{i=1}^n f(\alpha_i) f_i$. Since $f \in W^\circ$, $f(\alpha_i) = 0$ for all $\alpha_i \in W$, $1 \leq i \leq k$. Thus $f = \sum_{i=1}^n f(\alpha_i) f_i$. □

□

Corollary 3.5.1

V : n-d.v.s./F. W be k-dim. subspace. Then W is intersection of $n-k$ hyperspaces in V of the form N_f for some $0 \neq f_i \in V^*$.

Proof. Basis of W can be extended to basis of V. Take $\{f_1, \dots, f_n\} \subset V^*$ be the dual basis of $\{\alpha_1, \dots, \alpha_n\}$. Then $W = \cap_{i=k+1}^n N_{f_i}$. □

Corollary 3.5.2

V : n-d.v.s./F. W be hyperspace. Then $W = N_f$ for some $0 \neq f \in V^*$.

Exercise 3.5.2

W_1, W_2 be subspaces. $V : n$-d.v.s./F. Then $W_1 = W_2 \iff W_1^\circ = W_2^\circ$.

3.6 The Double Dual

Definition 3.6.1: Double Dual

$V : F$-v.s. $V^{**} = L(V^*.F) = L(L(V,F),F)$.

Note:-

Dual is not natural in general, but double dual is natural. Define $L_\alpha \in V^{**}$ as: $L_\alpha : V^* \to F : f \mapsto f(\alpha)$.

Note:-

Define $\mathfrak{L} : V \to V^{**} : \alpha \mapsto L_\alpha$.

Claim 3.6.1

$\mathfrak{L}$ is a lin. trans.

Proof. Suppose $\alpha_1, \alpha_2 \in V, c \in F$. $\mathfrak{L}(c\alpha_1+\alpha_2) = L_{c\alpha_1+\alpha_2}(f) = f(c\alpha_1+\alpha_2) = cf(\alpha_1)+f(\alpha_2)$. □

Claim 3.6.2

$\mathfrak{L}$ is injective.

Proof. Suppose for some $\alpha \in V$, we have $\mathfrak{L}(\alpha)L_\alpha \in V^*$ is $0 \iff \forall f \in V^*\ (L_\alpha(f) = 0) \iff \forall f \in V^*\ (f(\alpha) = 0) \iff \alpha = 0$. Thus $\mathfrak{L}$ is injective. □

Note:-

Thus $\mathfrak{L}$ is not surjective in general for infinite dimensional V.

Theorem 3.6.1

V : f.d.v.s./F. Then $\mathfrak{L}$ is an iso. of vec. spaces.

Proof. $V : n$-d.v.s./F. Then $\dim(V^*) = \dim(V^**) = n$. Thus $\mathfrak{L}$ is injective. lin. trans. from n-dim. to n-dim. is automatically surjective. □

Definition 3.6.2: Proper Subspace

V : v.s./F. Then $W \subset V$ is proper if it is not equal to V.

Definition 3.6.3: Maximal

V : v.s./F. A proper subspace $W \subsetneq V$ is said to be maximal if there is no intermediate subspace between W and V, i.e., if there is subspace $W \subset Z \subset V$, then either $W = Z$ or $V = Z$.

Note:-

If $\dim(V) = n$, then proper maximal subspace has dim. $n-1$.

Definition 3.6.4: Generalization of Hyperspace

V : v.s./F. A hyperspace of V is a proper maximal subspace of V.

Theorem 3.6.2

V : F-v.s. Suppose $f \in V^*\backslash\{0\}$. Then, $N_f = \{x \in v \mid f(x) = 0\}$ is hyperspace in V.

Proof. N.T.S. N_f is proper maximal subspace of V. It is proper since $N_f = V$ implies $f \equiv 0$, which is contradiction. E.T.S. that $\forall \alpha \in V\backslash N_f$, $\text{span}\{N_f, \alpha\} = V$. For this, E.T.S. that $\forall \beta \in V$ $(\beta \in \text{span}\{N_f, \alpha\})$. Let $c := \frac{f(\beta)}{f(\alpha)}$. Note that $\alpha \notin N_f$ is $f(\alpha) \neq 0$. Let $\gamma := \beta - c\alpha$. Then $f(\gamma) = f(\beta) - cf(\alpha) = 0$. Thus $\gamma \in N_f$. Then $\beta = \gamma + c\alpha \in \text{span}\{N_f, \alpha\}$ since $\gamma \in N_f$. Thus N_f is hyperspace. □

Theorem 3.6.3

V : F-v.s. Let W be hyperspace. Then $\exists f \in V^*\backslash\{0\}$ $(W = N_f)$.

Proof. Since it's proper, $\exists \alpha \in V\backslash W$. $W \subsetneq \text{span}\{W, \alpha\} \subset V$. Since W is maximal, $\text{span}\{W, \alpha\} = V$. Then $\forall \beta \in V$ can be written as $\beta = \gamma + c\alpha$ for some $\gamma \in W$, $c \in F$.

Claim 3.6.3

This γ and c are uniquely decided by β.

Proof. Suppose $\beta = \gamma + c\alpha = \gamma' + c\alpha'$. Then $\gamma - \gamma' = (c' - c)\alpha$. If $c' - c \neq 0$, then $\alpha \in W$, which is contradiction. Thus this expression is unique. □

Claim 3.6.4

$c := g(\beta) \in F$. Then $g : V \to F : \beta \mapsto g(\beta)$ is linear.

Proof. N.T.S. $g(d\beta_1 + \beta_2) = dg(\beta_1) + g(\beta_2)$. Let $\beta_1 = \gamma_1 + c_1\alpha$, $\beta_2 = \gamma_2 + c_2\alpha$ where $c_i = g(\beta_i)$. Then $\beta_1 + \beta_2 = \gamma_1 + \gamma_2 + (c_1 + c_2)\alpha$. By the uniqueness of the expression, $g(\beta_1 + \beta_2) = c_1 + c_2 = g(\beta_1) + g(\beta_2)$. Also, $g(d\beta_1) = dc_1 = dg(\beta_1)$. Thus g is linear. □

Since $g \in V^*$, $N_g = W$, thus our statement holds. □

3.7 The Transpose of a Linear Transformation

Definition 3.7.1: Transpose

$T : V \to W$: lin. trans. of F-v.s. We define the transpose $T^t : W^* \to V^*$. Then $V \to W \to F$ is defined as $g \circ T \in V^*$. So $T^t(g) = g \circ T$.

Lemma 3.7.1

T^t is lin. trans.

Proof. $T^t(cg_1 + g_2) = (cg_1 + g_2) \circ T = cg_1 \circ T + g_2 \circ T = cT^t(g_1) + T^t(g_2)$. □

Theorem 3.7.1

$T : V \to W$: lin. trans. Then

i) $N(T^t) = Ann(R(T))$

ii) If V, W is f.d.v.s./F, rank$(T^t) =$ rank(T) and $R(T^t) = Ann(N(T))$

$$
\begin{tikzcd}
T: & V \arrow[r] \arrow[d, leftrightarrow] & W \arrow[d, leftrightarrow] & T^*: & W^* \arrow[r] \arrow[d, leftrightarrow] & V^* \arrow[d, leftrightarrow] & Ann(R(T)) = N(T^t) \\
& N(T) & R(T) & & R(T^t) & N(T^t) & Ann(N(T)) = R(T^t)
\end{tikzcd}
$$

Proof. i): $N(T^t)$=$Ann(R(T))$, $g \in N(T^t) \iff T^t(g) = 0 \iff g \circ T = 0 \iff g \in Ann(R(T))$.

ii): Let $\dim(V) = n$, $\dim(W) = m$. Let $r :=$ rank(T). Then $\dim(Ann(R(T))) = m - r$. By i), $Ann(R(T)) = N(T^t) \Rightarrow \dim N(T^t) = m - r \Rightarrow \dim(R(T^t)) = r$ by rank-nullity. Next, N.T.S. $R(T^t) = Ann(N(T))$. Let $f \in R(T^t)$. Then $f = T^t(g)$ for some $g \in W^*$. Then $f = g \circ T$. Now if $\alpha \in N(T)$, $f(\alpha) = g \circ T(\alpha) = 0$, thus $f \in Ann(N(T))$. So $R(T^t) \subset Ann(N(T))$. But since both have same dim., $R(T^t) = Ann(N(T))$. $\square$

Chapter 4

Polynomials

4.1 Algebras

Definition 4.1.1: Algebra

F-algebra A or linear algebra A/F is an F-v.s. with a product structvue $A \times A \to A$ which has ass., dis., comm. where multiplication is not necesserily comm. If A has an element $1_A \in A$ s.t. $\forall \alpha \in A$ $(1_A \cdot \alpha = \alpha \cdot 1_A = \alpha)$ then we say A is an F-algebra with 1.

Example 4.1.1

(i) $F[x]$: finite polynomial with coeff. in F is F-algebra with unity 1.
(ii) $F[[x]]$: formal power series in x with coeff. in F : $\sum_{i=1}^{\infty} a_i x^i$ form is F-algebra with unity 1.
(iii) Suppose $n \geq 1$ with field F. $M_{n\times n}(F)$: F-algebra with unity $1_A = I_n$
(iv) V : F-v.s. $A = L(V,V)$ is F-algebra with unity $1_A = Id_V$ with $+$ and $\circ$.

4.2 The Algebra of Polynomials

Note:-

$f, g \in F[x]$. $f := \sum a_i x_i$, $g := \sum b_j x_j$ We say $f = g \iff \forall i = j$ $(a_i = b_j)$. But this is not equiv. to say that $\forall \alpha \in F$ $(f(\alpha) = g(\alpha))$.

Example 4.2.1

$F = \mathbb{Z}/p$. Then Fermat's Little Theorem says $\forall \alpha \in F$ $(\alpha^p \equiv \alpha)$. Consider $f = 1 + x^p$ and $g = 1 + x$. Then $f \neq g$ but $f(\alpha) = g(\alpha)$.

Definition 4.2.1: Degree of Polynomials

Suppose $f \in F[x] \backslash \{0\}$. Degree of f is defined to be n if $f = a_0 + \cdots + a_n x^n$ with $a_n \in F \backslash \{0\}$. Note that we don't define degree of 0.

Definition 4.2.2: Monic

$f \in F[x] \backslash \{0\}$ is monic if the coeff. of highest deg. is 1.

Exercise 4.2.1

$f, g \in F[x] \backslash \{0\}$. Then $fg \in F[x] \backslash \{0\}$ where $\deg(fg) = \deg(f) + \deg(g)$ and if f, g is monic, fg either.

Definition 4.2.3: Evaluation

A is an F-algebra and $f(x) \in F[x]$ where $f = \sum_{i=0}^{n} a_i x^i$. Let $\alpha \in A$ be a fixed element. Define $f(\alpha) = \sum_{i=0}^{n} a_i \alpha^i$ and we call it the evaluation of α in $f(x)$.
$ev_\alpha : F[x] \to A : f(x) \mapsto f(\alpha)$. $f_1 + f_2$, $f_1 f_2$, cf_1 are all respected.

Definition 4.2.4: Homomorphism

Let A_1 and A_2 be both F-algebras. A function $\varphi : A_1 \to A_2$ is called a homomorphism of F-algebra if:

1. It is an F-lin. trans.
2. $\varphi(\alpha_1 \alpha_2) = \varphi(\alpha_1)\varphi(\alpha_2)$

Theorem 4.2.1 Euclidean Algorithm on $F[x]$

$f, g \in F[x]$ for nonzero g with property $\deg(f) \geq \deg(g)$. $\exists q \in F[x]$ $(r = f - qg)$. we have either $r = 0$ or $r \neq 0$ for $\deg(r) < \deg(g)$.

Note:-
In modern algebra, a ring with this property is called an Euclidean domain.

Definition 4.2.5: Divisibility

If $r = 0$, $f = qg$. Then we denote this situation as $g \mid f$.

Lemma 4.2.1

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

Proof. $f = qg + r = q(x - c) + r$. Then $f(c) = r$, so $(x - c) | f \iff r = 0$. These are called a zero, solution, or root of f. □

Exercise 4.2.2

$f(x) \in F[x]$, $\deg(f) = n \geq 1$. Then f has at most n roots.

4.3 Lagrange Interpolation

This Chapter is Intentionally Skipped at Lectures

4.4 Polynomial Ideals

Definition 4.4.1: Ideals

F : field. $F[x]$: polynomial ring over F. An ideal $M \subset F[x]$ is an F-subspace s.t. if $f \in F[x]$ and $g \in M$, then $fg \in M$.

Example 4.4.1

$M = (x)$: poly. divisible by x.

Definition 4.4.2: Principal Ideal

An ideal of the form $M = (g_0)$: poly. divisible by g_0 is called a principal ideal.

Theorem 4.4.1

F : field. $M \subset F[x]$: a nonzero ideal. Then M is a principal ideal given by a monic.

Proof. Since $M \neq 0$, M does contain nonzero poly. So, the set of deg. of nonzero poly. in $\mathbb{N}_0$ is nonempty. Let $g_0 \in M$ hs the minimal possible deg. If $g_0 = a_d x^d + \cdots a_1 x + a_0$, then $\frac{1}{a_d} g_0 = x^d + \cdots$ with the same deg. So using this instead, call it g_0, the g_0 is monic.

Claim 4.4.1

$M = (g_0)$.

Proof. $g_0 \subset M$ is obvious.

$(M \subset g_0)$: N.T.S. $\forall f \in M$ $(f = qg_0)$. By the Euclidean algorithm, $\exists q, r \in F[x]$ $(f = g_0 q + r)$. Suppose $r \neq 0$. Then $f = qg_0 + r$ with $\deg(r) < \deg(g_0)$. But $r = f - qg_0$ where $f, g_0 \in M$, $r \in M$. This is contradiction to minimality of g. Thus $r = 0$, which means f is multiple of g_0. $\square$

$\square$

Note:-

By putting g_0 monic, g_0 is also unique.

Corollary 4.4.1

$p_1, p_2, \cdots, p_n \in F[x]$ not all zero. Then $\exists!$ monic $g_0 \in F[x]$ s.t.

i) $p_1 F[x] + \cdots + p_n F[x] = (g_0)$

ii) $\forall i$ $(g_0 \mid p_i)$

iii) if $f \mid p_i$ for all i, then $f \mid g_0$. Such g_0 is called G.C.D. of p_i.

Proof. Check $p_1 F[x] + \cdots + p_n F[x]$ is an ideal. By this, $M \neq 0 \Rightarrow \exists! g_0$ $((g_0) = M)$. Also, $(p_i) \subset M = (g_0) \Rightarrow p_{\in}(g_0) \Rightarrow g_0 \mid p_i$. Also, $f \mid p_i \Rightarrow p_i = f h_i$ thus $g_0 = f h_1 F[x] + \cdots + f h_n F[x] \Rightarrow f | g_0$. $\square$

Definition 4.4.3: Coprime (Relatively Prime)

p_i are coprime of relatively prime if $\gcd(p_1,\dots,p_n) = (1)$.

4.5 The Prime Factorization of a Polynomial

Definition 4.5.1: Reducible

F : field. $f \in F[x]\backslash\{0\}$. We say f is reducible if $f = gh$ for some $g, h \in F[x]$ where $\deg(g), \deg(h) \geq 1$. If we can't, we say it is irreducible.

Definition 4.5.2: Prime Element

We say f is a prime element if it has property that whenever $f \mid gh$, either $f \mid g$ or $f \mid h$.

Example 4.5.1

F : field. f : poly. of deg. 1 in $F[x]$ is irreducible.

Example 4.5.2

$F : \mathbb{R}$. $f(x) = x^2 + ax + b$. f is irreducible $\iff$ f has a root in $\mathbb{R}$ $\iff$ $D \geq 0$.

Example 4.5.3

$F : \mathbb{F}_p = \mathbb{Z}/p$. Then there are many irreducible poly. of deg. d.

Theorem 4.5.1

Let $p(x) \in F[x]\backslash\{0\}$. Then it is irreducible $\iff$ it is prime.

Proof. $(\Leftarrow)$: Suppose it is reducible. $p = gh$ for some $g, h \in F[x]$ with $\deg. \geq 1$. Since p is prime, $p \mid g$ or $p \mid h$. But then, $\deg(p) \leq \deg(g)$ or $\deg(p) \leq \deg(h)$. But this is impossible since $\deg(g), \deg(h) < \deg(p)$.

$(\Rightarrow)$: $\gcd(p, g) = (d) \Rightarrow d \mid p \Rightarrow p$ is irreducible, so $d = 1$ or $d = p$. If $d = p$, $d \mid g$ leads $p \mid g$. If $d = 1$, $\exists p_0, g_0$ $(pp_0 + gg_0 = 1)$. Thus $php_0 + ghg_0 = h$ leads $p \mid h$. □

Theorem 4.5.2

F : field. Every non-constant poly. $f(x) \in F[x]$ factors into a product of irreducible poly. $f = p_1 p_2 \cdots p_r$, and this is unique up to relabeling.

Sketch. For convenience, assume f is monic. If $\deg(f) = 1$, $f(x) = x - a$ for $a \in F$. Since $(x - a)$ irreducible, it just holds.

Suppose $\deg(f) > 1$. We use induction. Suppose theorem holds $\forall g$ $(\deg(g) < \deg(f))$. If f itself is irreducible, $f = f$. If f is reeducible, $f = gh$ for some non-constant $g, h \in F[x]$ of $\deg(g), \deg(h) < \deg(f)$. By induction hypothesis, $g = p_1 \cdots p_r$, $h = q_1 \cdots q_s$. By putting together, $f = p_1 \cdots p_r q_1 \cdots q_s$. So existence is proven.

For uniqueness, suppose $f = p_1 \cdots p_r = q_1 \cdots q_s$. Then $p_1 \mid q_1 \cdots q_s$. Being prime, $p_1 \mid q_j$ for some j. Since q_j is irreducible, $cq_j = p_1$. By cancelling, repeating, and relabeling, we can deduce factorization is unique. □

Definition 4.5.3: Formal Derivative

$f(x) \in F[x] = a_0 + a_1x + \cdots + a_nx^n$. Define $f' = a_1 + 2a_2x + \cdots + na_nx^{n-1}$ as formal derivative.

Lemma 4.5.1

$(f+g)' = f' + g'$ and $(fg)' = f'g + fg'$.

Theorem 4.5.3

$f \in F[x]$. Then, f is a product of distinct irreducible poly. $\iff$ f and f' are relatively Prime.

Sketch. ($\Leftarrow$): Suppose f and f' are relatively prime but $f = p^2h$ for irreducible p. Then $f' = 2pp'h + p^2h'$, which is contradiction.

($\Rightarrow$): Exercise! $\square$

Definition 4.5.4: Algebraically Closed

F is algebraically closed if every irreducible poly. in $F[x]$ is of deg. 1.
$\iff$ Every $f(x) \in F[x]$ of deg. $n \geq 1$ has precisely n roots with multiplicity.
$\iff$ Every non-constant $f \in F[x]$ factors into linear poly.

Example 4.5.4

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

Chapter 5

Determinants

5.1 Commutative Rings

Definition 5.1.1: Ring

R : a ring with two operation $+, \cdot$ s.t. $< R, + >$ form abelian group and $\cdot$ satisfies $a \cdot (b + c)$ and $(b + c) \cdot a$. A ring with unity is a ring with $1 \in R$ s.t. $\forall a \ (1 \cdot a = a \cdot 1 = a \in R)$.

5.2 Determinant Functions

Definition 5.2.1: n-Linear and Alternating

K : a ring. A function $D : K^{n \times n} \to K$. This is considered as a function on n rows and n columns.
i) We say D is n-linear if D is a linear function on the i-th row while fixing others. $D(ca_1 + a_1', a_2, \dots, a_n) = cD(a_1', a_2, \dots, a_n) + D(a_1, a_2, \dots, a_n)$.
ii) An n-linear function $D : K^{n \times n} \to K$ is called alternating if $D(A) = 0$ when $\forall i \neq j \ (a_i = a_j)$.

Exercise 5.2.1

$D : K^{n \times n} \to K$: alternating n-linear function. $A \in K^{n \times n}$. $A' :=$ matrix obtained by interchanging i, j-th rows and fix others. Then $D(A') = -D(A)$.

Proof. Using given property. Exercise! □

Definition 5.2.2: Determinant Function

K : commu. ring with 1. $D : K^{n \times n} \to K$ be a function. We say D determinant function if D is n-linear, alternating, and $D(I_n) = 1$.

Theorem 5.2.1
$\exists!$ such D that we call the determinant function.

Theorem 5.2.2
Concrete description of D in terms of permutation.

Definition 5.2.3: Minor

K : commu. ring with 1, $n > 1$. Let $A \in K^{n\times n}$ and (i,j) for $1 \le i,j \le n$. $A(i|j)4$ is $(n-1)\times(n-1)$ mat. with i-th row and j-th col. removed. We call this (i,j)-minor.

Definition 5.2.4

$D(A(i|j)) = D_{ij}(A)$.

Theorem 5.2.3

$n > 1$, $D : K^{(n-1)\times(n-1)} \to K$, alternating $(n-1)$-linear function. Let $1 \le j \le n$. $A \in K^{n\times n}$. Define $E_j(A) := \sum_{i=1}^{n}(-1)^{i+j}A_{ij}D_{ij}(A)$. Then E_j is an alternating n-linear function on $K^{n\times n}$. Also, if $D : K^{(n-1)\times(n-1)} \to K$ is a determinant function, so is E_j.

Proof. $A : n\times n$ mat. Note that $D_{ij}(A)$ is indep. of the entries of i-th row and j-th col. D is $(n-1)$-linear on $K^{(n-1)\times(n-1)}$, so $D_{ij}(A)$ is linear, further more $A_{ij}D_{ij}(A)$ is n-linear. Thus E_j is n-linear being a lin. comb. of n-linear functions. To prove alternating, suppose A has two equal rows at α_k, α_{k+1}. Take $i \neq k, k+1$. Then $D_{ij}(A) = 0$ because $A(i|j)$ has two identical rows and D is alternating. Then $E_j(A) = (-1)^{k+j}D_{kj}(A) + (-1)^{k+1+j}D_{k+1j}(A)$. Here, $A_{kj} = A_{k+1j}$, $D_{k+1j} = D_{kj}$, thus 0. This shows E_j is alternating n-linear. Also, since $I_n(i|j) = I_{n-1}$, we can see trivailly $E_j(I_n) = 1$. □

Corollary 5.2.1

For all $n \in \mathbb{N}$, $\exists$ det, function.

Proof. If $n = 1$, $D_1 = Id_k$ is a det. function. Suppose $n > 1$ and cor. holds for $1 \le i < n$. Then D_{n-1} is a det. function, thus we can take $D_n = E_j$ written in terms of D_{n-1}. □

5.3 Permutations and the Uniqueness of Determinants

Definition 5.3.1: Permutation

A permutation σ of S is a bijective function $\sigma : S \to S$. We have $|S|!$ permutations.

Definition 5.3.2: Transposition

$\tau \in S_n$ is called transposition if it interchange just the values of 2 members.

Note:-

Every peprmutation can be written as a product of disjoint cycles. Also, every cycle is a product of non-disjoint transpositions.

Theorem 5.3.1

S_n be the permutations on n letters. $\sigma \in S_n$. For any permutation, the number of transpositions needed to express $\sigma \mod 2$ is an invariant of σ. Also, we define $\text{sgn}(\sigma)$ as 1 if mod is even, -1 if odd.

Corollary 5.3.1
$\sigma_1, \sigma_2 \in S_n$. Then $\text{sgn}(\sigma_1\sigma_2) = \text{sgn}(\sigma_1)\text{sgn}(\sigma_2)$.

Theorem 5.3.2 The Uniqueness of Determinant
Let $D : K^{n\times n} \to K$ be a function that is alternating n-linear with $D(I_n) = 1$. Then D is unique with $D = \sum_{\sigma\in S_n} \text{sgn}(\sigma)A(1,\sigma_1)\cdots A(n,\sigma_n)$.

Proof. Suppose $e_1, \ldots, e_n$ as rows of I_n and α_i as rows of A. Then $\alpha_i = \sum_{j=1}^n A_{ij}e_j$, so $D(A) = D(\alpha_1, \ldots, \alpha_n) = D(\sum_{j=1}^n A_{1j}e_j, \ldots, \alpha_n) = \sum_{j=1}^n A_{ij}D(e_j, \ldots, \alpha_n)$, thus
$D(A) = \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n A_{1j_1} \cdots A_{nj_n} D(e_{j_1}, \ldots, e_{j_n})$. If any $j_p = j_q$, then $D(e_1, \ldots, e_n) = 0$. Thus all entries in this are different. So j_i are permutation of $\{1, \ldots, n\}$. If σ is the permutation, $D(e_{j_1}, \ldots, e_{j_n}) = \text{sgn}(\sigma)D(I_n)$. Therefore det. function is unique. □

Theorem 5.3.3
$\det(AB) = \det(A)\det(B)$.

Hint. B is fixed. Define $D(A) := \det(AB)$ as n-linear algternating. Then $D(A) = \det(A)D(I_n)$. □

5.4 Additional Properties of Determinants

Corollary 5.4.1
$\det(A^t) = \det(A)$

Proof. $\det(A^t) = \sum_\sigma \text{sgn}(\sigma)A(\sigma_1, 1)\cdots A(\sigma_n, n)$. Take $i = \sigma^{-1}j$. $A(\sigma i, j) = A(j, \sigma^{-1}j)$. Thus $\det(A^t) = \sum_{\sigma^{-1}} \text{sgn}(\sigma^{-1})A(1, \sigma^{-1}1)\cdots A(n, \sigma^{-1}n) = \det(A)$. □

Corollary 5.4.2
$A : n \times n$ mat. and B : i-th row $\leftarrow r_i + cr_j$. Then $\det(A) = \det(B)$.

Proof. $\det(r_1 + cr_2, r_2) = \det(r_1, r_2) + 2\det(r_2, r_2) = \det(r_1, r_2) = \det(A)$. □

Theorem 5.4.1
If
$$M = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$
for block mat. A, B, C, then $\det(M) = \det(A)\det(B)$.

. Fix A, B, then $D(A, B, C)$ is a function of C. Then D is alternating, linear function. □

Note:-
Now, we can use cofactor expansion to derive determinant.

Definition 5.4.1: Adj

$\mathrm{adj}(A) := C^t$ where each entries of C are cofactor expansion of A, i.e., $C = [C_{ij}]$.

Corollary 5.4.3

$A \cdot \mathrm{adj}(A) = \mathrm{adj}(A) \cdot A = \det(A) I_n$.

Corollary 5.4.4

$A^{-1} = \mathrm{adj}(A)/\det(A)$.

5.5 Modules

This Chapter is Intentionally Skipped at Lectures

5.6 Multilinear Functions

This Chapter is Intentionally Skipped at Lectures

5.7 The Grassman Ring

This Chapter is Intentionally Skipped at Lectures

Chapter 6

Elementary Canonical Forms

6.1 Introduction

6.2 Characteristic Values

Definition 6.2.1: Characteristic Value and Vectors, Spaces

T : endo. on f.d.v.s V/F. A characteristic value of T is $c \in F$ s.t. $\exists \alpha \in V \backslash \{0\}$ s.t. $T\alpha = c\alpha$. This α is also called a characteristic vector of T associated to c. Also, $E_c := \{\alpha \in V \mid T\alpha = c\alpha\}$ is called the characteristic space of T associated to c.

Theorem 6.2.1

T : endo. on f.d.v.s. V/F. TFAE:

i) c is a characteristic value of T

ii) Operator $T - cI$ is singular (not invertible)

iii) $\det(T - cI) = 0$

Proof. ii) $\iff$ iii) is trivial. If i) holds, $\exists v \in V \backslash \{0\}$ $(Tv = cv) \Rightarrow (T - cI)v = 0$. Thus this is not injective, so singular. Thus i) $\iff$ ii). □

Definition 6.2.2: Characteristic Polynomials

$f(x) := \det(xI - A) \in F[x]$ is called characteristic polynomial of T. Then f is monic with $\deg(f) = n$ for $n \times n$ mat. A and $\forall c$ which is characteristic values, $f(c) = 0$.

Exercise 6.2.1

Check the choice of basis doesn't affect the char. poly. of T.

Proof. $B := P^{-1}AP$. $\det(xI - B) = \det(xI - P^{-1}AP) = \det(P^{-1}(xI - A)P) = \det(xI - A)$. □

Definition 6.2.3: Diagonalizable

T : endo. on f.d.v.s. V/F. If $\exists \mathfrak{B} = \{v_1, v_2, \ldots, v_n\}$ s.t. each v_i are char. vec. of T, we say T is diagonalizable.

Note:-

$[T]_{\mathfrak{B}} = \begin{bmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{bmatrix}$ with (may be) repititions. Then $[T]_{\mathfrak{B}}$ is diagonal mat. Furthermore, we can see $f(x) = \det(xI - [T]_{\mathfrak{B}})$ is decomposed complety into a product of linear factors.

Example 6.2.1

$A : n \times n$ mat. on f.d.v.s. $V/\mathbb{R}$. If char. poly. has no real sol., then it is not diagonalizable.

Lemma 6.2.1

T : endo. on f.d.v.s. V/F. Suppose $c_1, c_2, \ldots, c_k$ are all possible distinct char. values of T and $W_i := \text{Null}(T - c_i I)$. Then $W := W_1 + \cdots + W_k \Rightarrow \dim(W) = \dim(W_1) + \cdots + \dim(W_k)$.

Proof. Trivially $\dim(W) \leq \dim(W_1) + \cdots + \dim(W_k)$. Thus we have to check $\geq$ part. Suppose $\forall \beta_i \in W_i$ $(\beta_1 + \cdots + \beta_k = 0)$. We will show $\forall \beta_i = 0$. Suppose $\beta_1 + \beta_2 = 0$. Then $T\beta_1 + T\beta_2 = c_1\beta_1 + c_2\beta_2 = 0$. We can derive $(c_1 - c_2)\beta_2 = 0$. Since $c_1 \neq c_2$, $\beta_2 = 0$ thus $\beta_1 = 0$. Inductively, we can derive $\forall \beta_i = 0$. Thus $\dim(W) = \dim(W_1) + \cdots + \dim(W_k)$. □

Theorem 6.2.2

T : endo. on n-d.v.s. V/F. $c_1, c_2, \ldots, c_k$ are all possible distinct char. values of T and $W_i := \text{Null}(T - c_i I)$. TFAE:

i) T is diagonalizable

ii) Char. poly. $p(x) = \prod_{i=1}^{k}(x - c_i)^{d_i}$ where $d_i = \dim(W_i)$

iii) $d_1 + d_2 + \cdots + d_k = n = \dim(V)$

Proof. i)$\Rightarrow$ ii): $\exists \bigcup_{i=1}^{k} \mathfrak{B}_i$, basis of V where each $\mathfrak{B}_i$ are the part belonging to c_i. Then, $\text{span}(\mathfrak{B}_i) = W_i$, $\dim(W_i) = d_i \Rightarrow p(x) = \prod_{i=1}^{k}(x - c_i)^{d_i}$ where $d_i = \dim(W_i)$.

ii)$\Rightarrow$ iii): Trivial.

iii)$\Rightarrow$ i): $W_1 + \cdots + W_k = W \Rightarrow d_1 + \cdots + d_k = n$. Thus $W = V$. Thus V has a basis consisting of char. vec., so diagonalizable. □

6.3 Annihilating Polynomials

Theorem 6.3.1

T : endo. on n-d.v.s. V/F. $p(x)$ as char. poly. of T, and $m(x)$ as min. poly. of T. Ignoring multiplicities, $p(x)$ and $m(x)$ has same sol. in F.

Proof. $m(c) = 0 \Rightarrow m(x) = (x - c)q(x)$. m is minimal implies $q(T) \neq 0$. Thus $\exists \beta \in V$ s.t. $q(T)\beta \neq 0$. This leads $(T - cI)q(T)\beta = 0$ since $(T - cI)q(T)\beta = m(T)\beta = 0\beta$. Thus $q(T)\beta$ is char. vec., which leads c as a char. value of T, so $p(c) = 0$.
Now if $p(c) = 0$, $\exists \alpha \in V \backslash \{0\}$ s.t. $T\alpha = c\alpha$. Thus $T^n\alpha = c^n\alpha$. So for any poly. $f(x) \in F[x]$, $f(T)\alpha = f(c)\alpha$. In particular, $m(T)\alpha = m(c)\alpha \Rightarrow m(c)\alpha = 0\alpha \Rightarrow m(c) = 0$. □

Corollary 6.3.1

$p(x) = \prod_{i=1}^{k}(x-c_i)^{d_i} \Rightarrow m(x) = \prod_{i=1}^{k}(x-c_i)^{r_i}$ where $1 \le r_i \le d_i$.

Theorem 6.3.2 Cayley-Hamilton

T : endo. on n-d.v.s. V/F. $p(x)$ as char. poly. of T. Then $p(T) = 0$. In particular, $m(x)\,|\,p(x)$.

Proof. $K := \{h(T) \mid h(x) \in F[x]\}$ be image of $ev_T : F[x] \to L(v, v)$. Let $\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\}$ be a basis of V. $A := [T]_{\mathfrak{B}}$ so that $T\alpha_i = \sum_{j=1}^{n} A_{ji}\alpha_j \ (i \in [n]) \Rightarrow \sum_{j=1}^{n}(\delta_{ij}T - A_{ji}I)\alpha_j = 0$. Then $B := [B_{ij}]$ where $B_{ij} := (\delta_{ij}T - A_{ji}I)$. We know $\mathrm{adj}(B)\cdot B = B\cdot\mathrm{adj}(B) = \det(B)I$. By construction, $\sum_{j=1}^{n} B_{ij}\alpha_j = 0 \Rightarrow \sum_{j=1}^{n}\mathrm{adj}(B)_{ki}B_{ij}\alpha_j = 0$. Taking sums over i leads $0 = \sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{adj}(B)_{ki}B_{ij}\alpha_j = \sum_{j=1}^{n}\left(\sum_{i=1}^{n}\mathrm{adj}(B)_{ki}B_{ij}\right)\alpha_j = \sum_{j=1}^{n}\delta_{kj}\det(B)\alpha_j = \det(B)\alpha_k$. Since $\{\alpha_1, \ldots, \alpha_n\}$ is basis, $\det(B) = 0$, which is char. poly. of T. □

6.4 Invariant Subspaces

Theorem 6.4.1

T : endo. on f.d.v.s. V/F. $c_1, c_2, \ldots, c_k$ are all possible distinct char. values of T. Then T is diagonalizable $\iff m(x) = (x-c_1)(x-c_2)\cdots(x-c_k)$.

Proof. Only for ($\Rightarrow$) here: Let $f(x)$ be a char. poly. of T. Then $m(x)\,|\,f(x)$. Thus $m(x) = (x-c_1)^{e_1}(x-c_2)^{e_2}\cdots(x-c_k)^{e_k}$.

Claim 6.4.1

$(T-c_1I)(T-c_2I)\cdots(T-c_kI) = 0$

Proof. Since T is diagonalizable, it has a basis $\{\alpha_1, \ldots, \alpha_n\}$ consisting of char. vec. Thus $T\alpha_j = c_{i(j)}\alpha_j$ where $c_{i(j)} \in \{c_1, \ldots, c_k\}$. This leads $(T - c_{i(j)}I)\alpha_j = 0$. Take $S := (T-c_1I)\cdots(T-c_kI)$. Then for each $j \in [n]$, $S(\alpha_j) = 0$. since each α_i form basis, $\forall v \in V \ (S(v) = 0)$. Thus Claim 6.4.1 holds. □

Oppisite of this proof is at Theorem 6.4.3. □

Corollary 6.4.1

T : endo. on n-d.v.s. V/F. Suppose T has n distinct char. values. If $f(x) = \prod_{i=1}^{n}(x-c_i)$ where distinct c_i, then $m(x) = f(x)$ thus it is diagonalizable.

Definition 6.4.1: T-Invariant Subspaces

T : endo. on n-d.v.s. V/F. Take subspace W. We say W is T-invariant or invariant under T if $T(W) \subset W$. If W is T-invariant, then T induces a endo. on W, denoted as $T|_W$.

$$\begin{array}{rccc} T: & V & \longrightarrow & V \\ & \updownarrow & & \updownarrow \\ T|_W: & W & \longrightarrow & W \end{array}$$

Example 6.4.1

$W = 0$ is trivailly T-invariant. Also, char. space E_c is T-invariant.

Lemma 6.4.1

Suppose W is T-invariant. $m(x)$ as min. poly. and $f(x)$ as char. poly. of T. Then $m_W(x) \mid m(x)$ and $f_W(x) \mid f(x)$ for each restriction to W.

Proof. Choose a basis $\mathfrak{B}' = \{\alpha_1, \ldots, \alpha_k\}$ of W and extend it to $\mathfrak{B} = \{\alpha_1, \ldots, \alpha_k, \alpha_{k+1} \ldots, \alpha_n\}$ which is a basis of V. Since W is T-inv., $T\alpha_i \in \text{span}\{\mathfrak{B}'\}$. So $A = [T]_{\mathfrak{B}} = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$ where $B = [T|_W]_{\mathfrak{B}'}$. Furthermore, $f(x) = \det(xI - A) = \det(xI - B) \cdot \det(xI - D)$. clearly, $f_W(x) \mid f(x)$.

Note that $A^r = \begin{bmatrix} B^r & C_r \\ 0 & D^r \end{bmatrix}$. Therefore, $\forall p(x) \in F[x]$ $(p(T) = 0)$, we can see $p_W(x) \mid p(x)$. Especially, $m_W(x) \mid m(x)$. $\square$

Definition 6.4.2: T-Conductors

T : endo. on f.d.v.s. V/F. W be T-inv. subspaces. Suppose $\alpha \in V$. We define T-conductor as $S_T(\alpha; W) := \{g(x) \in F[x] \mid g(T)\alpha \in W\}$.

Lemma 6.4.2

$S_T(\alpha; W)$ is a nonzero ideal.

Proof. char. poly. $f(x)$ satisfies $f(T) = 0 \in W \Rightarrow f(x) \in S_T(\alpha; W)$. Trivially it is closed. Also, since polynomials are commutative and W is T-inv., it satisfies properties of ideals. $\square$

Definition 6.4.3: T-Conductor as Generator

The unique monic poly. generator of $S_T(\alpha; W)$ is also often called the T-conductor of α to W.

Corollary 6.4.2

Min. poly. and char. poly. is in $S_T(\alpha; W)$, thus generator of that conductor divides both.

Definition 6.4.4: Triangulable

T : endo. on f.d.v.s. V/F. We say T is triangulable if V has a basis $\mathfrak{B}$ s.t. $[T]_{\mathfrak{B}}$ is an upper triangular mat.

Corollary 6.4.3

T is diagonalizable $\Rightarrow$ T is triangulable.

Lemma 6.4.3

T : endo. on f.d.v.s. V/F. Suppose $m(x) = \prod_{i=1}^{k} (x - c_i)^{r_i}$ where c_i are all distinct and $r_i \geq 1$. If W is T-inv. subspace, then $\exists \alpha \in V \backslash W$ $((T - cI)\alpha \in W)$ for some char. value $c = c_i$.

Proof. Let $\beta \in V \backslash W$ and let $g(x)$ be the min. T-conducting poly. taking β to W. Then $g(x) \mid m(x)$. Since $\beta \notin W$, $\deg(g(x)) \geq 1$. Then $g(x) = \prod_{i=1}^{k}(x-c_i)^{e_i}$ for $e_i \leq r_i$. since $\deg(g) \geq 1$, $\exists j\ (e_j \geq 1)$, so $(x-c_j) \mid g(x) \Rightarrow g(x) = (x-c_j)h(x)$. $\alpha := h(T)\beta$. This cannot be in W since $g(x)$ is the min. deg. fellow in $S_T(\beta\,;W)$. But $(T-c_jI)\alpha = g(T)\beta \in W$. Thus $(x-c_j) = S_T(\alpha\,;W)$. □

Theorem 6.4.2

T : endo. on n-d.v.s. V/F. T is triangulable $\iff$ $m(x) = \prod_{i=1}^{k}(x-c_i)^{r_i}$ for $r_i \geq 1$.

Proof. ($\Rightarrow$): Since T is triangulable, $\exists \mathfrak{B}$ s.t. $[T]_{\mathfrak{B}}$ is triangular. Thus char. poly. $f(x) = \prod_{i=1}^{k}(x-c_i)^{e_i}$ for $\sum e_i = n$, $e_i \geq 1$ and distinct c_i. Since $m(x) \mid f(x)$ our statement holds.

($\Leftarrow$): Suppose $m(x) = \prod_{i=1}^{k}(x-c_i)^{r_i}$. We use the Lemma 6.4.3 repeatedly over different choices of W. Take $W = 0$ then $\exists \alpha_1 \in V \backslash W\ ((T-d_1)\alpha_1 = 0)$ for some d_1. Take $W_1 = \mathrm{span}\{\alpha_1\}$. Then $\exists \alpha_2 \in V \backslash W_1\ ((T-d_2)\alpha_2 = 0)$. Repeating this, we can derive $T\alpha_1 = d_1\alpha_1$, $T\alpha_2 = *\alpha_1 + d_2\alpha_2$, and so on, thus $[T]_{\{\alpha_1,\ldots,\alpha_n\}} = \begin{bmatrix} d_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & d_k \end{bmatrix}$, which is upper triangular mat. □

Theorem 6.4.3

T : endo. on n-d.v.s. V/F. T is diagonalizable $\iff$ $m(x) = (x-c_1)(x-c_2)\cdots(x-c_k)$.

Proof. Forward is at Theorem 6.4.1. ($\Leftarrow$): Let $W \subset V$ be subspace spanned by all char. vec. Suppose $W \subsetneq V$ toward contradiction. Since $T\alpha = c\alpha$ for each char. vec. α of T, W is T-inv. So by Lemma 6.4.3, $\exists \alpha \in V \backslash W\ ((T-c_jI)\alpha =: \beta \in W)$. Note that $\beta \in W \backslash \{0\}$. So we can write $\beta = \beta_1 + \cdots + \beta_k$ where $\beta_i \in E_{c_i}$. Here, $T\beta_i = c_i\beta_i$, and $T^k\beta_i = c_i^k\beta_i$. Thus $f(T)\beta = f(T)\beta_1 + \cdots + f(T)\beta_k$. $m(x) := (x-c_j)h(x)$ where $h(x) = \prod_{i\neq j}(x-c_i)$. Clearly $h(c_j) \neq 0$. Consider $h(x) - h(c_j) = (x-c_j)q(x) \Rightarrow h(T)\alpha - h(c_j)\alpha = q(T)(T-c_jI)\alpha = q(T)\beta \in W$. Also, $m(T)\alpha = (T-c_jI)h(T)\alpha = 0 \Rightarrow h(T)\alpha \in E_{c_j} \subset W$. Thus $(h(T)\alpha \in W) \wedge (q(T)\beta \in W)$ implies $h(c_j)\alpha \in W$, so $h(c_j) = 0$. This is contradiction to the fact that min. poly. has distinct roots, so $W = V$, which means V has basis consisting of char. vec., and T is diagonalizable. □

Corollary 6.4.4

If F is algebraically closed, then T is always triangulable.

6.5 Simultaneous Triangulation; Simultaneous Diagonalization

Definition 6.5.1: Commuting Family

T_i : endo. on n-d.v.s. V/F. We say $\mathscr{F}$ is a commuting family of endo. if $\forall T_i, T_j \in \mathscr{F}\ (T_iT_j = T_jT_i)$.

Definition 6.5.2: $\mathscr{F}$-Invariant

If $\forall T_i \in \mathscr{F}$ ($W$is T_i-invariant), then we say W is $\mathscr{F}$-inv.

Lemma 6.5.1

Suppose $\mathscr{F}$ is a commuting family of triangulable endo. Suppose $W \subsetneq V$, which is $\mathscr{F}$-inv. Then $\exists \alpha \in V \backslash W$ $(\forall T_i \in \mathscr{F}$ $((T_i - cI)\alpha \in \operatorname{span}\{W, \alpha\}))$.

Proof. We may assume $\{T_1, \ldots, T_r\}$, a maximal lin. indep. subset of $\mathscr{F}$. Applying Lemma 6.4.3 to T_1, $\exists \beta_1 \in V \backslash W$ $\exists c_1 \in F$ $((T_1 - c_1 I)\beta_1 \in W)$. Let $V_1 = \{\beta \in V \mid (T_1 - c_1 I)\beta \in W\}$. $\beta_1 \in V_1$, so it is nonempty and $W \subsetneq V_1 \subset V$. Here, by construction, V_1 is $\mathscr{F}$-inv. since $\forall T_i \in \mathscr{F}$ $((T_1 - c_1 I)T\beta = T(T_1 - c_1 I)\beta \in W)$.

Now, take $V_1 \subset V$ and let $U_2 := T_2|_{V_1}$. Applying Lemma 6.4.3 to $V_1 \backslash W$ and U_2, $\exists \beta_2 \in V_1 \backslash W$ $\exists c_2 \in F$ $((T_2 - c_2 I)\beta_2 \in W)$. So, $\beta_2 \notin W$, $(T_1 - c_1 I)\beta_2 \in W$, $(T_2 - c_2 I)\beta_2 \in W$. Take $V_2 = \{\beta \in V_1 \mid (T_2 - c_2 I)\beta \in W\}$. Then $(\beta_2 \notin W) \wedge (\beta_2 \in V_2)$. By repeating, we can get $W \subsetneq \cdots \subset V_1 \subset V$. Thus terminates in finite steps since $\dim(V) < \infty$. □

Corollary 6.5.1

V: f.d.v.s./F and $\mathscr{F}$ as comuuting family of triangulable endo. Then $\exists \mathfrak{B}$ s.t. $[T_i]_{\mathfrak{B}}$ are all upper triangular mat.

Proof. Exercise. Use our argument for a single operator and use Lemma 6.4.3 for commuting families. □

Corollary 6.5.2

V: f.d.v.s./F and $\mathscr{F}$ as comuuting family of diagonalizable endo. Then $\exists \mathfrak{B}$ s.t. $[T_i]_{\mathfrak{B}}$ are all diagonal mat.

Corollary 6.5.3

Suppoer F is algebraically closed and $\mathscr{F}$ as commuting family of endo. Then $\exists$ simultaneously triangulating basis.

6.6 Direct-Sum Decompositions

Definition 6.6.1: Independent

V: v.s./F. We say subspaces, just say W_i, are indep. if there common elements are just 0.

Definition 6.6.2: Internal Direct Sum

If $W = \sum_{i=1}^{k} W_i$ and each W_i are indep., then we say the sum is direct and we write it as $W = \bigoplus_{i=1}^{k} W_i$.

Exercise 6.6.1

If $W = \bigoplus_{i=1}^{k}$, then $\exists!$ expression of $w \in W$ w.r.t. each $w_i \in W_i$.

Definition 6.6.3: Projection

V: f.d.v.s./F. Supopose we have endo. $E : V \to V$ s.t. $E^2 = E$. Then we say E is a projection.

Example 6.6.1

$V := V_1 \oplus V_2$. $P_1 : V \mapsto V_1$ and $P_2 : V \mapsto V_2$. Then those classical 'projection' is actually a projection we defined above.

Lemma 6.6.1

Let E be a projection. Then for $V := V_1 \oplus V_2$ and $P_1 : V \mapsto V_1$, E really is a classical 'projection', i.e., $E = P_1 : V \mapsto V_1$.

Proof. $V_1 := R(E)$, $V_2 := N(E)$.

Claim 6.6.1

$V = V_1 \oplus V_2$

Proof. Let $v \in V$. Then $v = E(v)+v-E(v)$. $E(v) \in R(E)$. Also, $E(v-E(v)) = E(v)-E^2(v) = 0$, so $(v-E(v)) \in N(E)$. Thus $V = R(E)+N(E)$. To show this is direct, suppose we have $v_1+v_2 = 0$ for $(v_1 \in R(E)) \wedge (v_2 \in N(E))$. Then $v_1 = -v_2 \in R(E)$ and $\exists \alpha \in V\ (v_1 = R(\alpha))$. $E(v_1) = -E(v_2) = 0$ and $E(v_1) = E^2(\alpha) = E(\alpha) = v_1$. Since $E(v_1) = 0$, $v_1 = 0$. Thus $v_2 = 0$, which leads sum is direct. □

Now if $v \in V_1 \oplus V_2$, write $v = v_1 + v_2$, then $E(v) = E(v_1) = v_1$. So $E = P_1$. □

Theorem 6.6.1

V: f.d.v.s./F and $V = \bigoplus_{i=1}^{k} W_i$. Then $\exists E_i : V \mapsto W_i$ s.t.

i) Each E_i are projection

ii) $\forall i \neq j\ (E_i E_j = 0)$

iii) $I = \sum E_i$

iv) The range of E_i is W_i

Converse also holds. Furthermore, only i), ii), and iii) leads our theorem.

Proof. i), ii), and iv) are trivial by definition. For iii), take $\alpha \in V$. $\alpha = \sum E_i \alpha \Rightarrow I = \sum E_i$. Conversely, suppose we have E_i $i \in [k]$ s.t. they satisfy those first three conditions. We can take W_i as $R(E_i)$. Then, $V = W_1 + \cdots + W_k$. We have to show this is direct. By iii), we have $\alpha = \sum E_i \alpha$. This expression is unique since if $\alpha = \alpha_1 + \cdots + \alpha_k$ for $\alpha_i \in W_i$, then using i) and ii), we can derive $E_j \alpha = \sum_{i=1}^{k} E_j \alpha_i = E_j^2 \beta_j = E_j \beta_j = \alpha_j$ if we take $\alpha_i = E_i \beta_i$. □

6.7 Invariant Direct Sum

Theorem 6.7.1

T : endo. on n-d.v.s. V/F. $V = \bigoplus_{i=1}^{k} W_i$. Let $E_i : V \mapsto V$ be projection to W_i. Then W_i are T-inv. $\iff$ T commutes with E_i.

Proof. ($\Leftarrow$): Suppose T commutes with all E_i. Let $\alpha_i \in W_i = R(E_i)$. N.T.S. $T\alpha_i \in W_i$. We can write $\alpha_i = E_i\beta$. So $T\alpha_i = TE_i\beta = E_iT\beta$, which leads $T\alpha_i \in R(E_i) = W_i$. Since α_i was arbitrary element is W_i, W_i is T-inv.

($\Rightarrow$): Let $\alpha \in V$. We can say $\alpha = v_1 + \cdots + v_k$ for each $v_i \in W_i$ uniquely. $W_i := R(E_i)$, so each $v_i = E_i(\alpha)$. So $\alpha = E_1(\alpha) \cdots + E_k(\alpha) \Rightarrow T\alpha = TE_1(\alpha) + \cdots TE_k(\alpha)$. Since $E_i(\alpha) \in W_i$ is T-inv., $T(E_i\alpha) = E_i(\beta_i) \in W_i \Rightarrow T\alpha = E_1(\beta_1) + \cdots + E_k(\beta_k)$. For $i \neq j$, $E_jTE_i\alpha = E_jE_i\beta_i = 0$. For $i = j$, $E_jTE_j\alpha = E_j\beta_j$. Thus $E_jT\alpha = E_jTE_1\alpha + \cdots + E_jTE_k\alpha = E_j\beta_j = TE_j\alpha$. Thus $E_jT = TE_j$ since α is arbitrary. $\square$

Theorem 6.7.2

T : endo. on n-d.v.s. V/F. If T is diagonalizable and if $c_1, \ldots, c_k$ are the distinct char. values of T, then $\exists E_i$ on V s.t.

i) $T = c_1E_1 + \cdots + c_kE_k$

ii) $I = \sum E_i$

iii) $\forall i \neq j \ (E_iE_j = 0)$

iv) $E_i^2 = E_i$

v) The range of E_i is the char. space for T associated with c_i

Converse also holds. Furthermore, only i), ii), and iii) leads our theorem.

Proof. ($\Rightarrow$): Suppose diagonalizable with char. values c_i. $W_i := E_{c_i} = N(T - c_iI)$. Since T is diagonalizable, $V = \bigoplus_{i=1}^{k} W_i$. Thus ii)~v) are trivial. Now, $\alpha = \sum E_i\alpha \Rightarrow T\alpha = \sum TE_i\alpha = \sum T\alpha_i = \sum c_i\alpha_i = \sum c_iE_i\alpha$. Since α is arbitrary, $T = \sum c_iE_i$.

($\Leftarrow$): Using ii) and iii) to obtain iv). using i) and iv) to obtain $R(E_i) \subset N(T - c_iI)$. Since we assumed $E_i \neq 0$, c_i is char. value of T. Take $i) - c \times ii)$. Then $(T - cI) = (c_1 - c)E_1 + \cdots + (c_k - c)E_k$. so if $(T - cI)\alpha = 0$, we must have $(c_i - c)E_i\alpha = 0$. If $\alpha \neq 0$, then $E_i\alpha \neq 0$ for some i, so in this case, $c_i = c$. Certainly T is diagonalizable, since every nonzero vector in $R(E_i)$ is a char. vec. of T, and $I = \sum E_i$ shows these char. vec. span V. Now we have to show $N(T - c_iI) = R(E_i)$. This is clear since if $T\alpha = c_i\alpha$, then $\sum_{j=1}^{k}(c_j - c_i)E_j\alpha = 0$ hence $(c_j - c_i)E_j\alpha = 0$ for each j, and then $E_j\alpha = 0$ for $j \neq i$. Since $\alpha = \sum E_i\alpha$ and $E_j\alpha = 0$ for $j \neq i$, $\alpha = E_i\alpha$, which shows $\alpha \in R(E_i)$. $\square$

6.8 The Primary Decomposition Theorem

Theorem 6.8.1 Primary Decomposition Theorem

T : endo. on f.d.v.s. V/F. $\exists$ a decomposition of V into $V = \bigoplus_{i=1}^{k} W_i$ s.t. $W_i = N(p_i(T)^{r_i})$ where $m(x) = \prod_{i=1}^{k} p_i(x)^{r_i}$ for $r_i \geq 1$ and irreducible, distinct p_i. Also, each W_i are

T-inv., and $T_i := T|_{W_i}$ has min. poly. $p_i(T)^{r_i}$.

Proof. When $k = 1$, it is trivial. Suppose $k > 1$. Define $f_i(x) := \frac{m(x)}{p_i(x)^{r_i}} = \prod_{j \neq i} p_j(x)^{r_j}$. Then $\gcd(f, p_i^{r_i}) = 1$. Since each f_i are also relatively prime, $\exists g_1, \ldots, g_k$ $(f_1 g_1 + \cdots + f_k g_k = 1)$. Define $h_i(x) := f_i(x) g_i(x)$. For $i \neq j$, $m \mid f_i f_j$ thus $f_i(T) f_j(T) = 0$. Note that $\sum h_i(T) = I$. Define $E_i := h_i(T)$. Then $\sum E_i = I$ and $\forall i \neq j$ $(E_i E_j = 0)$ since $E_i E_j = f_i(T) g_i(T) f_j(T) g_j(T) = 0$. Thus we can see E_i are projection. Thus $V = \bigoplus_{i=1}^{k} R(E_i)$ and each are T-inv.

Claim 6.8.1

$R(E_i) = W_i = N(p_i(T)^{r_i})$

Proof. Let $\alpha \in R(E_i)$. Then $\alpha = E_i \alpha \Rightarrow p_i(T)^{r_i} \alpha = p_i(T)^{r_i} f_i(T) g_i(T) \alpha = 0$ since $p_i(T)^{r_i} f_i(T) g_i(T) = m(T) g_i(T) = 0$. Thus $R(E_i) \subset N(p_i(T)^{r_i})$. Conversely, let $\alpha \in N(p_i(T)^{r_i})$. Note that if $i \neq j$, $p_i^{r_i} \mid f_j$ thus $p_i^{r_i} \mid f_j g_j = h_j$, thus $f_j(T) g_j(T) \alpha = h_j(T) \alpha = 0$. In other words, $\forall i \neq j$, α is in V whose projection about E_j is 0. Thus α has only $R(E_i)$ component. Thus $N(p_i(T)^{r_i}) \subset R(E_i)$, consequently $R(E_i) = N(p_i(T)^{r_i})$. □

Now we have to show T_i has min. poly. as $p_i(x)^{r_i}$. Note that $W_i = N(p_i(T)^{r_i})$ implies $p_i(T)^{r_i}|_{W_i} = 0$. Thus $m_i(x) \mid p_i(x)^{r_i}$. So $m_i(x) = p_i^{s_i}$ for $1 \leq s_i \leq r_i$. E.T.S. $s_i = r_i$. Let $g(x)$ be poly. s.t. $g(T_i) = 0$.

Claim 6.8.2

$p_i(x)^{r_i} \mid g(x)$

Proof. $g(T_i) = 0 \iff g(T) f_i(T) = 0$. So min. poly. of T divides $g(x) f_i(x)$. Since $\gcd(p_i^{r_i}, f_i) = 1$, $m(x) \mid g(x) f_i(x)$ leads $p_i^{r_i} \mid g(x)$. In particular, $m_i(x)$ is divisible by $p_i^{r_i}$, thus $r_i = s_i$. □

□

Corollary 6.8.1

$E_1, \ldots, E_k$ be projection associated to primary decomposition of V w.r.t. T. Then each E_i is a poly. in T. In particular, if $U : V \mapsto V$ is another endo. commuting with T, then, U commutes with each E_i so W_i are U-inv.

Theorem 6.8.2

T : endo. on f.d.v.s. V/F. If T is triangulable, $\exists$ diagonalizable D and nilpotent N s.t. $T = D + N$ and $DN = ND$. Such D and N are uniquely determined by T.

Proof. $m(x) = \prod (x - c_i)^{r_i}$ for distinct c_i. Take $R(E_i) = W_i := N((T - c_i I)^{r_i})$ as like Theorem 6.8.1. Take $D := \sum c_i E_i$ and $N = T - D$.

Claim 6.8.3

N is nilpotent

Proof. $I = \sum E_i \Rightarrow T = \sum T E_i \Rightarrow N = T - D = \sum (T - c_i I) E_i$. Since each E_i are poly. in T and $E_i E_j = 0$, $N^r = \sum (T - c_i I)^r E_i$. By choosing $r = \max(r_1, \ldots, r_k)$, $N^r = 0$. □

D and N are commute since they are poly. in T. Thus existence is proven.

For uniqueness, suppose we have $T = D' + N' = D + N$. Then $D - D' = N' - N$. We know $D - D'$ is diagonalizable. Now suppose $N^r = N'^{r'} = 0$. Then $(N' - N)^A = \sum_{i=0}^{A} \binom{A}{i} N'^i N^{A-i}$. Taking $A > r + r'$ leads $(N' - N)^A = 0$. Take $\alpha := N' - N = D' - D$. Then α is diagonalizable and nilpotent, which leads $\alpha = 0$. Thus $D = D'$ and $N = N'$. $\square$

Chapter 7

The Rational and Jordan Forms

7.1 Cyclic Subspaces and Annihilaters

Definition 7.1.1: T-Cyclic Subspaces

T : endo. on f.d.v.s. V/F. Take $\alpha \in V$. Then the T-cyclic subspace generated by α is denoted as $Z(\alpha; T) := \{g(T)\alpha \in V \mid g(x) \in F[x]\}$. Just in case $Z(\alpha; T) = V$, we say V is cyclically generated by α and T, and α is a cyclic vector for T.

Note:-

$Z(\alpha; T)$ is always T-invariant. Also, $Z(\alpha; T)$ is very sensitive to choice of α. If $\alpha = 0$, nothing no show. If α is a char. vec., then $T\alpha = c\alpha$, so $Z(\alpha; T) = \text{span}\{\alpha\}$, which implies 1-dimensional. Also note that converse holds.

Definition 7.1.2: T-Annihilaters

The T-annihilater, denoted as $M(\alpha; T) := \{g(x) \in F[x] \mid g(T)\alpha = 0\}$.

Note:-

Note that annihilator is just a special case of conductor, which takes $W = 0$. We can also see that monic generator of annihilater divides minimal poly.

Theorem 7.1.1

T : endo. on f.d.v.s. V/F. p_α : T-annihilator of α. Then

i) $\deg(p_\alpha) = \dim(Z(\alpha; T))$

ii) If $\deg(p_\alpha) = k$, then $\{\alpha, T\alpha, \ldots, T^{k-1}\alpha\}$ forms a basis of $Z(\alpha; T)$

iii) Let $U := T|_{Z(\alpha;T)} : Z \mapsto Z$. Then min. poly. of U is $p_\alpha(x)$.

Proof. Take $g(x) = p_\alpha q(x) + r(x)$ by Euclidean algorithm for $\deg(p_\alpha) = k$. Note that $(p_\alpha) = M(\alpha; T)$. Thus $p_\alpha q \in M(\alpha; T) \Rightarrow g(T)\alpha = p_\alpha(T)q(T)\alpha + r(T)\alpha = r(T)\alpha \Rightarrow Z(\alpha; T) = \text{span}\{\alpha, T\alpha, \ldots, T^{k-1}\alpha\}$. Thus $\dim(Z(\alpha; T)) \leq k$.

Claim 7.1.1

$\{\alpha, T\alpha, \ldots, T^{k-1}\alpha\}$ is a linearly independent set.

Proof. Suppose not. Then there is nonzero coefficients satisfying $\sum c_i T^i \alpha = 0$. Clearly $g(x) = \sum c_i x^i$ has deg $< k$. But p_α is the nonzero poly. of min. deg. in $M(\alpha; T)$, while $g(x) \in M(\alpha; T)$ with degree less than $p_\alpha(x)$. This is contradiction, so this set is linearly independent. □

Thus by Claim 7.1.1, $Z(\alpha; T)$ is k-dimensional with $\deg(p_\alpha) = k$. i) and ii) done.
For iii), need to check $p_\alpha(U) = 0$ and it really is poly. with min. deg.
An arbitrary element of $Z(\alpha; T)$ is of the form $g(T)\alpha$ for some $g(x) \in F[x]$. Thus $p_\alpha(U)g(T)\alpha = p_\alpha(T)g(T)\alpha = g(T)p_\alpha(T)\alpha = 0$. Our first condition holds. Second condition is immediate from the minimality of the degree of p_α in $M(\alpha; T)$. □

Lemma 7.1.1 Companion Matrices

T : endo. on f.d.v.s. V/F. $W = Z(\alpha; U) \subset V$ where $U := T|_{Z(\alpha; T)}$. Then w.r.t. the basis

$$\{\alpha, T\alpha, \dots, T^{k-1}\alpha\} = \mathfrak{B} \text{ of } Z,\ [U]_{\mathfrak{B}} = \begin{bmatrix} 0 & & & -c_0 \\ 1 & \ddots & & -c_1 \\ & \ddots & \ddots & \vdots \\ & & 1 & -c_{k-1} \end{bmatrix} \text{ where } p_\alpha = x^k + \sum_{i=0}^{k-1} c_i x^i.$$

This matrix is called companion matrix.

Proof. $\mathfrak{B} := \{\alpha, T\alpha, \dots, T^{k-1}\alpha\} = \{\alpha_1, \dots, \alpha_k\}$. Then $U\alpha_1 = T\alpha = \alpha_2$, $U\alpha_2 = T^2\alpha = \alpha_3$, and so on, $U\alpha_{k-1} = T^{k-1}\alpha = \alpha_k$. By our supposition of p_α, $p_\alpha(U)\alpha = U^k\alpha + \sum_{i=0}^{k-1} c_i U^i \alpha$. Thus we can derive companion matrix of above form. □

Theorem 7.1.2

U has a cyclic vec. $\iff$ there is some ordered basis s.t. U is represented by the companion mat. of the min. poly. for U.

Corollary 7.1.1

If A is the companion mat. of a monic poly. p, then p is both min. and char. poly. of A.

7.2 Cyclic Decompositions and the Rational Form

Definition 7.2.1: Complementary Subspaces

T : endo. on f.d.v.s. V/F. $W \subset V$ as T-inv. subspaces. If $\exists$ T-inv. subspace $W' \subset V$ s.t. $V = W \oplus W'$, then we say W' is a complementary T-inv. subspaces of W.

Definition 7.2.2: T-Admissible

T : endo. on f.d.v.s. V/F. A subspace is T-admissible if W is T-inv. and $\exists f(x) \in F[x]\ \exists \beta \in V\ \exists \gamma \in W\ (f(T)\beta \in W \Rightarrow f(T)\beta = f(T)\gamma)$.

Lemma 7.2.1

T : endo. on f.d.v.s. V/F. Suppose W is T-inv. If its complementary T-inv. subspace exists, then W is T-admissible.

Proof. W is trivially T-inv. Suppose $f(T)\beta \in W$ for $(f(x) \in F[x]) \wedge (\beta \in V)$. Since $V = W \oplus W'$, $\beta = \gamma + \gamma'$ for unique $(\gamma \in W) \wedge (\gamma' \in W')$. Then $f(T)\beta = f(T)\gamma + f(T)\gamma'$. Since $f(T)\beta$ and $f(T)\gamma$ are T-inv. and in W, $f(T)\gamma'$ should be in W. Independence of W and W' implies thus $f(T)\gamma' = 0$. Thus $f(T)\beta = f(T)\gamma$, so W is T-admissible. □

Theorem 7.2.1 Cyclic Decomposition Theorem

T : endo. on f.d.v.s. V/F. Let $W_0 \subset V$ be any proper T-admissible subspace. $\exists \alpha_1, \ldots, \alpha_r \in V\backslash\{0\}$ with respective T-annihilators $p_1, \ldots, p_r$ s.t.

i) $V = W_0 \oplus (\bigoplus_{i=1}^{r} Z(\alpha_i\,;\,T))$

ii) $p_k \mid p_{k-1}$

Furthermore, the integer r and p_i are uniquely determined by i), ii), and the fact that no α_k is 0.

Proof. We will divide our proof to 4 steps. During our proof, we intentionally denote $f(T)\beta$ as $f\beta$.

Before: Let $\beta \in V\backslash W$. Consider $S(\beta\,;\,W) := \{g(x) \in F[x] \mid g(T)\beta \in W\}$. Then $\exists$ monic poly. generator f s.t. $f(T)\beta \in W$. By T-admissibility, $\exists \gamma \in W$ s.t. $f(T)\beta = f(T)\gamma$. Let $\alpha := \beta - \gamma$, then $f(T)\alpha = 0$. Since $\gamma \in W$, we can see that $S(\alpha\,;\,W) = S(\beta\,;\,W)$ and f is also the T-conductor of α to W. Since $f(T)\alpha = 0$, $f \in M(\alpha\,;\,T)$. Thus $(f) = S(\alpha\,;\,W) \subset M(\alpha\,;\,T)$. Conversely, if $g \in M(\alpha\,;\,T)$, $g(T)\alpha = 0 \in W$ so $M(\alpha\,;\,T) \subset S(\alpha\,;\,T)$. Thus $S(\alpha\,;\,W) = M(\alpha\,;\,T)$ and f is also a T-annihilater.

Claim 7.2.1

$W \cap Z(\alpha\,;\,T) = 0$.

Proof. Suppose $g(T)\alpha \in W \cap Z(\alpha\,;\,T)$. Then $g \in S(\alpha\,;\,W) = M(\alpha\,;\,T)$ implies $g(T)\alpha = 0$. Thus $W \cap Z(\alpha\,;\,T) = 0$, so $W + Z(\alpha\,;\,T) \Rightarrow W \oplus Z(\alpha\,;\,T)$. □

Step 1: Let's make following observation: Let $W \subset V$ be a proper T-inv. subspace. Then $\max_{\alpha \in V} S(\alpha\,;\,W)$ is obtained by some $\beta \in V$, so that $\deg(S(\beta\,;\,W))$ is maximized.

For the above β, $W + Z(\beta\,;\,T)$ is T-inv. and strictly larger than W. Applying this observation to the given $W_0 \subset V$: T-inv. proper subspaces. Then we obtain $\beta_1 \in V$ s.t. $\deg(S(\beta_1\,;\,T))$ is maximized among $\deg(S(\beta\,;\,W))$. Again, take $W_2 = W_1 + Z(\beta_2\,;\,T)$, which leads $W_0 \subsetneq W_1 \subsetneq \cdots \subsetneq W_r = V$.

From this, we can derive at least $V = W_0 + \sum_{i=1}^{r} Z(\beta_i\,;\,T)$. Know Let's say $(p_k) := S(\beta_k\,;\,W_{k-1})$ has the maximum deg. among the conductors.

Step 2: Take W_i, β_i, p_i $i \in [r]$ as above. Fix $1 \le k \le r$ and let $\beta \in V$. Suppose $(f) = S(\beta\,;\,W_{k-1})$. Write $f\beta = \beta_0 + \sum_{i=1}^{k-1} g_i\beta_i$ for some $g_i \in F[x]$, $\beta_i \in W_i$.

Claim 7.2.2

$\beta_0 = f\gamma_0$ for some $\gamma_0 \in W_0$ and $f \mid g_i$.

Proof. If $k = 1$, it means W_0 is T-admissible, so nothing to proof. Thus suppose $k > 1$. By the Euclidean algorithm, $g_i = fh_i + r_i$. We want to prove all $r_i = 0$. Let $\gamma := \beta - \sum_{i=1}^{k-1} h_i\beta_i$. Then $\beta - \gamma = \sum_{i=1}^{k-1} h_i\beta_i \subset W_{k-1}$. This leads $S(\gamma\,;\,W_{k-1}) = S(\beta\,;\,W_{k-1})$. Also, $f\gamma = f\beta - \sum_{i=1}^{k-1} fh_i\beta_i = f\beta - \sum_{i=1}^{k-1}(g_i - r_i)\beta_i = \beta_0 + \sum_{i=1}^{k-1} g_i\beta_i - \sum_{i=1}^{k-1} g_i\beta_i + \sum_{i=1}^{k-1} r_i\beta_i$. Thus $f\gamma = \beta_0 + \sum_{i=1}^{k-1} r_i\beta_i \;\cdots\; (1)$. Toward contradiction, some $r_j \neq 0$ and say that j is the largest between such numbers.

$f\gamma = \beta_0 + \sum_{i=1}^{k-1} r_i\beta_i$ for nonzero r_i. Clearly $\dim(r_i) < \dim(f)$ $\cdots$ (2). Consider conductor $(p) := S(\gamma\,;W_{j-1})$. With $W_{j-1} \subset W_{k-1}$, $S(\gamma\,;W_{j-1}) \subset S(\gamma\,;W_{k-1}) = (f)$. Thus $f\,|\,p$, i.e., $p = fq$ for some $q \in F[x]$. Applying g to (2) leads $p(\gamma) = g\beta_0 + \sum_{i=1}^{j-1} gr_i\beta_i + gr_j\beta_j$ where $p(\gamma) \in W_{j-1}$, $g\beta_0 \in W_0 \subset W_{j-1}$, $gr_i\beta_i \in W_i \subset W_{j-1}$. This eq. leads $gr_j\beta_j \in W_{j-1}$, and thus $\deg(gr_j) \geq \deg(S(\beta_j\,;W_{j-1})) = \deg(p_j)$ by definition, and $\deg(p_j) \geq \deg(S(\gamma\,;W_{j-1}))$ by mazimality condition of β_j, where $\deg(S(\gamma\,;W_{j-1})) = \deg(p) = \deg(fg)$. Consequently, $\deg(r_i) \geq \deg(f)$, which is contradiction. Thus all $r_i = 0$, and all $f\,|\,g_i$, and (1) says $f\gamma = \beta_0 \in W_0$. Since W_0 is T-admissible, $\exists\gamma_0 \in W_0$ s.t. $f\gamma = \beta_0 = f\gamma_0$. □

Step 3: Now we will find $\{\alpha_1,\ldots,\alpha_r\}$ in V which satisfies i) and ii).

Take $\{\beta_1,\ldots,\beta_r\}$ as ***Step 1***. Fix $1 \leq k \leq r$. Apply ***Step 2*** to the vec. $\beta = \beta_k$ and the T-conductor $f = p_k$. We obtain $p_k\beta_k = p_k\gamma_0 + \sum_{i=1}^{k-1} p_k h_i\beta_i$ for $\gamma_0 \in W_0$. Let $\alpha_k := \beta_k - \gamma_0 - \sum_{i=1}^{k-1} h_i\beta_i$. Since $\beta_k - \alpha_k \in W_{k-1}$, $S(\alpha_k\,;W_{k-1}) = S(\beta_k\,;W_{k-1}) = (p_k)$, and since $p_k\alpha_k = 0$, we have $W_{k-1} \cap Z(\alpha_k\,;T) = \{0\}$. Because each α_k satisfies this condition, $W_k = W_0 \oplus (\bigoplus_{i=1}^k Z(\alpha_i\,;T))$ and that p_k is the T-annihilater of α_k.

Since $p_i\alpha_i = 0$ for each i, we have the trivial relation $p_k\alpha_k = 0 + p_1\alpha_1 + \cdots + p_{k-1}\alpha_{k-1}$. Apply ***Step 2*** with β_i replaced by α_i and with $\beta = \alpha_k$, we can conclude p_k divides each p_i with $i < k$.

Step 4: We will show r and each poly. p_r are uniquely determined by the conditions.

Take γ_i, g_i $i \in [s]$ that satisfies conditions either. We will show $r = s$ and $p_i = g_i$.

The poly. g_1 is determined as the T-conductor of V into W_0. Let $S(V\,;W_0)$ be the collection of poly. f s.t. $\forall\beta \in V$ $(f\beta \in W_0)$, i.e., poly. f s.t. $R(f(T)) \subset W_0$. Then $S(V\,;W_0)$ is nonzero ideal. g_1 is the monic generator of this. Each $\beta \in V$ has the form $\beta = \beta_0 + f_1\gamma_1 + \cdots + f_s\gamma_s$ and so $g_1\beta = g_1\beta_0 + \sum_{i=1}^{s} g_1 f_i\gamma_i$. Since each g_i divides g_1, we have $g_1\gamma_i = 0$ for all i and $g_1\beta = g_1\beta_0 \in W_0$. Thus $g_1 \in S(V\,;W_0)$. Since g_1 is the monic poly. of least deg. which sends γ_1 into W_0, we see that g_1 is the monic poly. of least deg. in the ideal $S(V\,;W_0)$. By the same argu., p_1 also, so $p_1 = g_1$. Now note three facts:

1. $fZ(\alpha\,;T) = Z(f\alpha\,;T)$
2. If $V = \bigoplus_{i=1}^k V_i$, where each V_i is T-inv., $fV = fV_1 \oplus \cdots \oplus fV_k$.
3. If α and γ have the same T-annihilator, then $f\alpha$ and $f\gamma$ have the same T-annihilater and thus $\dim(Z(f\alpha\,;T)) = \dim(Z(f\gamma\,;T))$.

Now, proceed induction to show that $r = s$ and $p_i = g_i$. Suppose $r \geq 2$. Then $\dim(W_0) + \dim(Z(\alpha_1\,;T)) < \dim(V)$ Since $p_1 = g_1$, we know $\dim(Z(\alpha_1\,;T)) = \dim(Z(\gamma_1\,;T))$. Thus $\dim(W_0) + \dim(Z(\gamma_1\,;T)) < \dim(V)$. Then

$$p_2V = p_2W_0 \oplus Z(p_2\alpha_1\,;T)$$
$$p_2V = p_2W_0 \oplus Z(p_2\gamma_1\,;T) \oplus \cdots \oplus Z(p_2\gamma_s\,;T)$$

satisfies our desire. Furthermore, we conclude that $p_2\gamma_2 = 0$ and g_2 divides p_2. The argument can be reversed to show that p_2 divides g_2. Thus $g_2 = p_2$. □

Corollary 7.2.1

If, W is T-admissible, it has complementary T-inv. subspace. So with Lemma 7.2.1, if and only if condition holds.

Theorem 7.2.2

T : endo. There is $\alpha \in V$ s.t. T-annihilator of α is equal to min. poly.

Proof. With $W_0 = 0$, apply cyclic decomposition. Take $\alpha = \alpha_1$. T-conductor fo α_1 to W_0 is T-annihilater of α_1, which is the min. poly. $\square$

Theorem 7.2.3

If T has cyclic vec., then char. poly. of T is equal to min. poly. of T.

Theorem 7.2.4 Generalized Cayley-Hamilton Theorem

T : endo. on f.d.v.s. V/F. m be min. poly. and p be char. poly. Then

i) $p \mid f$

ii) p and f have the same prime factors except for multiplicities

iii) If $p = f_1^{r_1} \cdots f_k^{r_k}$, then $f = f_1^{d_1} \cdots f_k^{d_k}$ where d_i is the nullity of $f_i(T)^{r_i}$ divided by the deg. of f_i.

Proof. $i)$: trivial from Cayley-Hamilton Theorem.

$ii)$: Cyclic decompose with W_0 says $\exists \alpha_1 \sim \alpha_r$ s.t. $V = \bigoplus_{i=1}^r Z(\alpha_i\,;\,T)$ with $m(x) = p_1(x)$ which is T-annihilater of α_1. $p_i \mid p_{i-1}$. Take $T_i := T|_{Z(\alpha_i\,;\,T)}$. Since $Z(\alpha_i\,;\,T)$ is a cyclic vec. space with cyclic vec. α_i, p_i is min. poly. for T_i is also char. poly. of T_i. Thus char. poly. $f(x) = \prod_{i=1}^r p_i$ and any prime factor of $m(x)$ divides $f(x)$ by i) while if a prime factor divides f, it divides one of p_i. Thus $p_i \mid p_{i-1} \mid \cdots \mid p_1 = m(x)$. Thus each prime factor of f also divides $m(x)$.

$iii)$: Apply primary decomposition: $W_i = N(f_i(T)^{r_i})$. Take $T_i := T|_{W_i}$. Then $f_i(x)^{r_i}$ is the min. poly. of T_i. Applying ii) to T_i its min. poly. Thus char. poly. of T_i is $f_i^{d_i}$ with $d_i \geq r_i$. Here, $\dim(W_i)$ is $d_i \cdot \deg(f_i)$. So $d_i = \frac{\dim(W_i)}{\deg(f_i)} = N(f_i(T)^{r_i}) / \deg(f_i)$. $\square$

Corollary 7.2.2

T : nilpo. endo. on n-d.v.s. V/F. Then char. poly. of T is x^n.

Proof. T is nilpo. $\Rightarrow \exists N$ s.t. $T^N = 0 \Rightarrow$ min. poly. $m(x) \mid x^N \Rightarrow m(x) = x^r$. Thus $f(x) = x^n$. $\square$

7.3 The Jordan Form

Note:-

How to find Jordan form?

Solution. ***Step 1:*** char. poly. $f(x) = \prod_{i=1}^k (x - c_i)^{d_i}$ for distinct c_i and $m(x) = \prod_{i=1}^k (x - c_i)^{r_i}$ for $1 \leq r_i \leq d_i$. Take $W_i = N((T - c_i I)^{r_i})$ as primary decomposition theorem. Then $V = \bigoplus_{i=1}^k W_i$. $T_i := T|_{W_i}$ where $m_i(x)$ of T_i is $(x - c_i)^{r_i}$.

Step 2: For each W_i, let $N_i := (T_i - c_i I) : W_i \to W_i$. Then N_i is nilpotent operator on W_i. Note that $T_i = N_i + c_i I$. Consider each W_i the cyclic decomposition of W_i w.r.t. N_i. So,

$W_i = \bigoplus_{k=1}^{s_i} Z(\alpha_k\,;N_i)$. Take $\beta_j = \{\alpha_j, N_i\alpha_j, \ldots, N_i^{k_j-1}\alpha_j\}$. Then

$$[N_i|_{Z(\alpha_j\,;N_i)}]_{\beta_j} = \begin{bmatrix} 0 & & & 0 \\ 1 & \ddots & & \vdots \\ & \ddots & \ddots & \vdots \\ & & 1 & 0 \end{bmatrix} \qquad \Rightarrow \qquad [T_i|_{Z(\alpha_j\,;N_i)}]_{\beta_j} = \begin{bmatrix} c_u & & & 0 \\ 1 & \ddots & & \vdots \\ & \ddots & \ddots & \vdots \\ & & 1 & c_i \end{bmatrix}. \tag{7.1}$$

Take $\mathfrak{B}^i = \cup\beta_j$. Then

$$[T_i|_{W_i}]_{\mathfrak{B}^i} = \begin{bmatrix} \square & & \\ & \ddots & \\ & & \square \end{bmatrix}$$

where each box is of the form at 7.1 R.H.S. Then finally take $B = \cup\mathfrak{B}^i$. This leads what we call Jordan form, where each small blocks are elementary Jordan blocks. $\square$

7.4 Computation of Invariant Factors

This Chapter is Intentionally Skipped at Lectures.

7.5 Summary; Semi-Simple Operators

This Chapter is Intentionally Skipped at Lectures.

Chapter 8

Inner Product Spaces

8.1 Inner Products

Definition 8.1.1: Inner Product

An inner product $(-,-)$ on V is a function $(-,-): V \times V \mapsto F$ satisfying:

1. $(-,\ )$ is linear functionoal with $(c\alpha+\beta,\gamma) = c(\alpha,\gamma)+(\beta,\gamma)$
2. $(\beta,\alpha) = \overline{(\alpha,\beta)}$
3. $\forall \alpha \in \mathbb{F}\backslash\{0\}$ $((\alpha,\alpha) > 0)$

Note:-

If $F = \mathbb{R}$, $(\beta,\alpha) = (\alpha,\beta)$. Thus 1. and 2. leads $(\ ,-)$ is also linear. Thus $(\ ,\ )$ is symmetric bilinear form.

But If $F = \mathbb{C}$, then $(\alpha, c\gamma) = \overline{c(\gamma,\alpha)}$. In this case, we call bbC is sesqui-linear. Also, $(\alpha,\alpha) = \overline{(\alpha,\alpha)}$, thus $(\alpha,\alpha) \in \mathbb{R}$.

Example 8.1.1 (Standard Inner Product)

$V := \mathbb{C}^n$, $[x_i],[y_i] \in \mathbb{C}^n$. Then $([x_i],[y_i]) = \sum_{i=1}^n x_i\bar{y}_i$ is called the standard inner product.

Example 8.1.2 (Positive Definite)

$F = \mathbb{R}^n$. A : $n\times n$ real mat. s.t. $\forall x \in \mathbb{R}^n$, $x^TAx > 0$. Then A is called positive definite. When A is symmetric pos. def., then $(x,y)_A := x^TAy$.

Exercise 8.1.1

Prove that $(x,y)_A$ is an inner product on $\mathbb{R}^n$.

Theorem 8.1.1

$F = \mathbb{R}$, $V = \mathbb{R}^n$. Let $(\ ,\ ): V \times V \mapsto F$ be an arbitrary inn. prod. on V. Then $\exists$ a sym. pos. def. mat. A s.t. $(\ ,\ ) = (\ ,\ )_A$.

Proof. Choose a basis, the standard basis for convenient. $(e_i,e_j) =: g_{ij}$. Define $A := [g_{ij}]$. Let $x,y \in \mathbb{R}^n$. Then $x = \sum_{i=1}^n x_ie_i$ and $y = \sum_{j=1}^n y_je_j$. $(x,y) = \sum_i\sum_j x_iy_je_ie_j = \sum_i\sum_j x_ig_{ij}y_j =$

$\sum_i x_i \sum_j g_{ij} y_j = [x^T]_{\mathfrak{B}} A [y]_{\mathfrak{B}}$. □

Definition 8.1.2: Hermitian Matrix

$n \times n$ mat. A is called Hermitian if $A^* = A = [a_{ij}]$ where $[A^*]_{ij} = [\overline{a_{ji}}] = \overline{A^T}$.

Theorem 8.1.2

$V = \mathbb{C}^n$. Let $(\ ,\ ) : V \times V \mapsto F$ be an inn. prod. on V. Then $(x, y) = x^* A y$ for some Hermitian pos. def. mat. A and vice versa.

Example 8.1.3

$V = C([a, b] \mapsto \mathbb{C})$: $\mathbb{C}$-v.s. of continuous functions on $[a, b]$. Define $f, g \in V$ as $(f, g := \int_a^b f(t)\overline{g(t)}dt)$. Then it is an inn. prod. on V of ∞-dim.

Definition 8.1.3: Quadratic Form

V : v.s. with inn. prod. Define the quadratic form $\alpha \mapsto ||\alpha||^2 = (\alpha, \alpha) \geq 0$ and $\alpha = 0 \iff ||\alpha||^2 = 0$.

Exercise 8.1.2 Polarization Identity

$||\alpha + \beta||^2 = ||\alpha||^2 + 2\Re(\alpha, \beta) + ||\beta||^2$. $F = \mathbb{R}$, $(\alpha, \beta) = \frac{1}{4}(||\alpha + \beta||^2 - ||\alpha - \beta||^2)$. $F = \mathbb{C}$, $(\alpha, \beta) = \frac{1}{4}(||\alpha + \beta||^2 - ||\alpha - \beta||^2 + i||\alpha + i\beta||^2 - i||\alpha - i\beta||^2)$.

8.2 Inner Product Spaces

Definition 8.2.1: Inner Product Space

$F = \mathbb{R}$ or $F = \mathbb{C}$. A vector space V/F with a specified inn. prod. called inner product space.

Note:-

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

Theorem 8.2.1

V : inn. prod. space. $\forall \alpha, \beta \in V\ \forall c \in F$, we have properties:

1. $||c\alpha|| = |c|||\alpha||$
2. $||\alpha|| > 0$ for $\alpha \neq 0$
3. $|(\alpha, \beta)| \leq ||\alpha||||\beta||$
4. $||\alpha + \beta|| \leq ||\alpha|| + ||\beta||$

Proof. 1 ane 2 are obvious from definition of inn. prod. For 3, just take $\alpha \neq 0$. Let $\beta^{\parallel} := \frac{(\beta,\alpha)}{||\alpha||^2}\alpha$, $\beta^{\perp} := \beta - \beta^{\parallel}$.

This is because: $(\beta^\perp, \alpha) = 0 \iff (\beta, \alpha) = c(\alpha, \alpha)$, thus $c = \frac{(\beta,\alpha)}{(\alpha,\alpha)}$.

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

For 4, $\|\alpha+\beta\|^2 = (\alpha+\beta, \alpha+\beta) = (\alpha, \alpha) + (\alpha, \beta) + (\beta, \alpha) + (\beta, \beta) \le (\|\alpha\| + \|\beta\|)^2$. Since both side are positive, we can just take off square. □

Note:-

The "angle" is defined as inner product. Using 3, we can derive $-1 \le \frac{(\alpha,\beta)}{\|\alpha\|\|\beta\|} \le 1$. Then for nonzero α, β, define angle θ as:

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

Definition 8.2.2: Orthogoanl and Orthogonal, Orthonormal Set

V : inn. prod. space. We say $\alpha, \beta \in V$ are orthogonal or perpendicular if their inn. prod. is 0. If $S \subset V$, S is orthogonal set if $\forall \alpha \ne \beta \in S$, $(\alpha, \beta) = 0$. If all element of S satisfies $\|\alpha\| = 1$, we say S is orthonormal.

Theorem 8.2.2

Suppose $S \subset V$ be orthogonal set. Then S is lin. indep.

Proof. Take $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ as distinct vectors in S. Then $\beta = c_1\alpha_1 + \cdots + c_m\alpha_m$. So $(\beta, \alpha_k) = (\sum_j c_j\alpha_j, \alpha_k) = \sum_j c_j(\alpha_j, \alpha_k) = c_k(\alpha_k, \alpha_k)$. Since $(\alpha_k, \alpha_k) \ne 0$, $c_k = \frac{(\beta,\alpha_k)}{\|\alpha_k\|^2}$ for $1 \le k \le m$. When $\beta = 0$, this leads $\forall c_i = 0$, so S is lin. indep. set. □

Theorem 8.2.3 Gram-Schmidt

Let V be an inn. prod. space and let $\beta_1, \ldots, \beta_n$ be any indep. vec. in V. Then we can construct orthogonal vectors $\alpha_1, \ldots, \alpha_n$ in V s.t. for each $k \in [n]$ the set $\{\alpha_1, \ldots, \alpha_k\}$ is a basis for the subspace spanned by $\beta_1, \ldots, \beta_k$.

Proof. We can apply Gram-Schmidt orthogonalization process. $\alpha_1 := \beta_1$. $\beta_2 = \beta_2^{\perp\{\alpha_1\}} + \beta_2^{\|\{\alpha_1\}}$. $\alpha_2 = \beta_2^{\perp\{\alpha_1\}} = \beta_2 - \beta_2^{\|\{\alpha_1\}} = \beta_2 - \frac{(\beta_2,\alpha_1)}{\|\alpha_1\|^2}\alpha_1$. $\beta_3 = \beta_3^{\perp\{\alpha_1,\alpha_2\}} + \beta_3^{\|\{\alpha_1,\alpha_2\}}$. $\alpha_3 = \beta_3^{\perp\{\alpha_1,\alpha_2\}} = \beta_3 - \frac{(\beta_3,\alpha_1)}{\|\alpha_1\|^2}\alpha_1 - \frac{(\beta_3,\alpha_2)}{\|\alpha_2\|^2}\alpha_2$, and so on. We can take orthogonal basis with this process. □

Corollary 8.2.1

All f.d. inn. prod. space has orthogonal basis.

Definition 8.2.3: Best Approximation

V : inn. prod. space. $W \subset V$, $\beta \in V \backslash W$. A best approximation of β to W is $\alpha \in W$ s.t. $\forall \gamma \in W$ $(\|\beta - \alpha\| \le \|\beta - \gamma\|)$.

Theorem 8.2.4

Let W be a subspaces of an inn. prod. space V and let β be a vec. in V. Then

1. $\alpha \in W$ is a best approx. to β by vec. in $W \iff \beta - \alpha$ is orthogonal to every vec. in W

2. If a best approx. to β by vec. in W exists, it is unique
3. If W is f.d. and $\{\alpha_1, \ldots, \alpha_n\}$ is any ortho. basis for W, then the vec. $\alpha = \sum_k \frac{(\beta,\alpha_k)}{\|\alpha_k\|^2}\alpha_k$ is the unique best approx. to β by vec. in W

Proof. Note that $\forall\gamma \in W$, $\beta-\gamma = (\beta-\alpha)+(\alpha-\gamma)$, and $\|\beta-\gamma\|^2 = \|\beta-\alpha\|^2 + 2\Re(\beta-\alpha, \alpha-\gamma) + \|\alpha-\gamma\|^2$. Now suppose $\beta-\alpha$ is ortho. to every vec. in W. Then since $(\alpha-\gamma) \in W$, we can see $\|\beta-\gamma\|^2 = \|\beta-\alpha\|^2 + \|\alpha-\gamma\|^2 \geq \|\beta-\alpha\|^2$.

Conversely, suppose $\forall\gamma \in W$ $(\|\beta-\gamma\| \geq \|\beta-\alpha\|)$. Then from above we can find that $\forall\gamma \in W$ $(2\Re(\beta-\alpha, \alpha-\gamma)+\|\alpha-\gamma\|^2 \geq 0)$. Since every vec. in W may be expressed in the form $\alpha-\gamma$ with $\gamma \in W$, we see that $2\Re(\beta-\alpha, \tau)+\|\tau\|^2 \geq 0$. We may take $\tau = -\frac{(\beta-\alpha,\alpha-\gamma)}{\|\alpha-\gamma\|^2}(\alpha-\gamma)$. Then the equality reduces to the statement $-\frac{|(\beta-\alpha,\alpha-\gamma)|^2}{\|\alpha-\gamma\|^2} \geq 0$, which holds iff $(\beta-\alpha, \alpha-\gamma) = 0$. This completes the proof of 1. and ortho. condition is evidently satisfied by at most one vec. in W, thus proves 2.

Now suppose W is f.d. and let $\{\alpha_1, \ldots, \alpha_n\}$ be ortho. basis for W. We know $\beta-\alpha$ is ortho. to each elements of basis, i.e., to every vec. in W, so α it is best approx. to β, which leads $\|\beta-\gamma\| \geq \|\beta-\alpha\|$. Therefore $\alpha \in W$ and it is best approx. to β. $\square$

Definition 8.2.4: Orthogonal Complement

$W^\perp := \{\beta \in V \mid \alpha \perp \beta \forall\alpha \in W\}$.

Exercise 8.2.1

V : f.d. inn. prod. space. Then $V = W \oplus W^\perp$

Proof. $\beta \in V$. Then $E\beta$ is best approx. lies in W. It is easy to see that this is proj. Also, since $\alpha - E\alpha$ and $\beta - E\beta$ are each ortho. to W, $c(\alpha-E\alpha)+(\beta-E\beta) = (c\alpha+\beta)-(cE\alpha+E\beta) \in W^\perp$. Thus E is linear transformation by uniqueness of ortho. proj.

Note that $(\beta \in W^\perp) \iff (E\beta = 0)$. The eq. $\beta = E\beta + (\beta - E\beta)$ shows $V = W + W^\perp$. Also, $W \cap W^\perp = \{0\}$, so $V = W \oplus W^\perp$. $\square$

8.3 Linear Functionals and Adjoints

Theorem 8.3.1

V : f.v.s./$\mathbb{R}$ or $\mathbb{C}$, V^* : dual vec. space. Let $f \in V^*$. Then $\exists!\beta \in V$ $(f(-) = (-, \beta))$.

Proof. Choose ortho basis $\{\alpha_1, \ldots, \alpha_n\}$ of V. For uniqueness, try $\beta = \sum_{i=1}^n c_i\alpha_i$. We can see $f(\alpha_j) = (\alpha_j, \sum_{i=1}^n c_i\alpha_i) = \sum_{i=1}^n \overline{c_i}(\alpha_j, \alpha_i) = \overline{c_j}(\alpha_j, \alpha_j)$. If such β exists, then it must be $\beta = \sum_{i=1}^n \frac{\overline{f(\alpha_i)}}{\|\alpha_i\|^2}\alpha_i$.

So take this as β. Now let's prove $f(-) = (-, \beta)$. We can see $(\alpha_j, \beta) = \sum_{i=1}^n \frac{f(\alpha_i)}{\|\alpha_i\|^2}(\alpha_i, \alpha_j) = \frac{f(\alpha_j)}{\|\alpha_j\|^2}(\alpha_j, \alpha_j) = f(\alpha_j)$. Thus such inn. prod. which corresponds to linear functional exists and unique. $\square$

Note:-

Usually V and V^* are not naturally related. But if V has inn. prod., then we can have an isomorphism.

Theorem 8.3.2

T : endo. on f.d.v.s.$V/\mathbb{R}$ or $\mathbb{C}$. Then $\exists! T : V \to T$ s.t. $(T\alpha, \beta) = (\alpha, T^*\beta)$ where T^* is a unique linear operator. If $F = \mathbb{R}$, T^* is transpose and if $F = \mathbb{C}$, T^* is conjugate transpose.

Proof. Fix $\beta \in V \Rightarrow (-, \beta) \in V^*$. Let's modify it a bit to get what we want. Theorem 8.3.1 says that $\exists! \beta' \in V$ $((T(-), \beta) = (-, \beta'))$. Define $T^* : V \to V : \beta \mapsto \beta'$. This mapping is well-defined. Also, easy to show that T^* is linear, and since for any $\beta \in V$, $T^*\beta$ is uniquely determined, thus uniqueness holds. □

Theorem 8.3.3

T : endo. on f.d.v.s.$V/\mathbb{F}$ or $\mathbb{C}$, $\mathfrak{B} = \{\alpha_1, \dots, \alpha_n\}$ be an orthonormal basis. Let $A := [T]_{\mathfrak{B}} = [A]_{ij}$. Then $A_{ij} = (T\alpha_j, \alpha_i)$.

Proof. $\alpha \in V$. $\alpha = \sum_{i=1}^n (\alpha, \alpha_i)\alpha_i$. A is defined by A_{ij} s.t. $T(\alpha_j) = \sum_{i=1}^n A_{ij}\alpha_i$. Since $T\alpha_j = \sum_{i=1}^n (T\alpha_j, \alpha_i)\alpha_i$, $A_{ij} = (T\alpha_j, \alpha_i)$. □

Corollary 8.3.1

T : endo. on f.d.v.s.$V/\mathbb{F}$ or $\mathbb{C}$, $\mathfrak{B} = \{\alpha_1, \dots, \alpha_n\}$ be an orthonormal basis. Then $[T^*]_{\mathfrak{B}} = ([T]_{\mathfrak{B}})^*$ where L.H.S. is adjoint s.t. $(T\alpha, \beta) = (\alpha, T^*\beta)$ and R.H.S. is conjugate transpose.

Proof. $A := [T]_{\mathfrak{B}} = [A_{ij}]$, $B := [T^*]_{\mathfrak{B}} = [B_{ij}]$. Then $A_{ij} = (T\alpha_j, \alpha_i)$ and $B_{ij} = (T^*\alpha_j, \alpha_i)$. Then $\overline{B_{ij}} = (\alpha_i, T^*\alpha_j) \Rightarrow \overline{B_{ji}} = (\alpha_j, T^*\alpha_i) = A_{ij}$. □

Exercise 8.3.1

$$(T_1 + T_2)^* = T_1^* + T_2^*,\ (cT)^* = \bar{c}T^*,\ (T_1T_2)^* = T_2^*T_1^*.$$

Definition 8.3.1: Hermitian

T : endo. on f.d.v.s.$V/\mathbb{R}$ or $\mathbb{C}$. We say T is Hermitian or self-adjoint if $T = T^*$.

8.4 Unitary Operators

Definition 8.4.1: Preserve

$T : V \to W$ on inn. prod. space V and W. Then we say T preserves the inn. prod. if $\forall \alpha, \beta \in V$ $(T\alpha, T\beta) = (\alpha, \beta)$. We say this is isometry.

Definition 8.4.2: Isomorphism of Inner Product Spaces

An isomorphism of inn. prod. space is a linear transf. s.t. it is an isomorphism of vec. spaces and preserves the inn. prod.

Theorem 8.4.1

$T : V \to W$ with same dim f.d. inn. prod. spaces. TFAE:

i) T preserves inn. prod.

ii) T is an isomorphism of inn. prod. spaces

iii) For arbitrary orthonormal basis $\mathfrak{B}$ of V, $T\mathfrak{B}$ is an orthonormal basis for W

iv) For some orthonormal basis mfB of V, $T\mathfrak{B}$ is an orthonormal basis for W

Proof. i) $\Rightarrow$ ii): Suppose $\exists \alpha \in N(T)$. Then $(T\alpha, T\alpha) = \|T\alpha\|^2 = \|\alpha\|^2 = 0$. Thus $\alpha = 0$. Since $\dim(V) = \dim(W)$, T is one-to-one, Thus T is an isomorphism.

ii) $\Rightarrow$ iii): Let $\mathfrak{B}$ an arbitrary orthonormal basis $\{\alpha_1, \dots, \alpha_n\}$. Then $(\alpha_i, \alpha_j) = \delta_{ij}$. Since T preserves, $(T\alpha_i, T\alpha_j) = \delta_{ij}$. Isomorphic condition of T implies then $T\mathfrak{B}$ is basis for W while $\{T\alpha_1, \dots, T\alpha_n\}$ is an orthonormal set.

iii) $\Rightarrow$ iv): Trivial.

iv) $\Rightarrow$ i): Let $\mathfrak{B}$ an orthonormal basis of V s.t. $T\mathfrak{B}$ is also an orthonormal basis.

Claim 8.4.1

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

Proof. $\alpha := \sum x_i \alpha_i$, $\beta := \sum y_i \alpha_i$. Then $T\alpha = \sum x_i T\alpha_i$ and $T\beta = \sum y_i T\alpha_i$. We can see $(T\alpha, T\beta) = (\sum x_i T\alpha_i, y_j T\alpha_j) = \sum_j \sum_i x_i \overline{y_j} (\alpha_i, \alpha_j)$

while $(\alpha, \beta) = (\sum x_i \alpha_i, \sum y_j \alpha_j) = \sum_j \sum_i x_i \overline{y_j} (\alpha_i, \alpha_j)$, and both are δ_{ij}. □

□

Theorem 8.4.2

$T : V \to W$ on inn. prod. space with preserving $\iff$ $\|T\alpha\| = \|\alpha\|$.

Proof. ($\Rightarrow$): Trivial since $\|T\alpha\|^2 = (T\alpha, T\alpha) = (\alpha, \alpha) = \|\alpha\|^2$.

($\Leftarrow$): By using polarization identity, we can easily derive this direction. □

Definition 8.4.3: Unitary Operator

T is unitary operator if it is an isomorphism on inn. prod. space.

Theorem 8.4.3

$U : V \to V$ on inn. prod. space. Then U is unitary $\iff$ U^* exists and $UU^* = U^*U = I$.

Proof. ($\Rightarrow$): If U is unitary, then, isomorphism, so $\exists U^{-1} : V \to V$ and $(U\alpha, \beta) = (U\alpha, I\beta) = (U\alpha, UU^{-1}\beta) = (\alpha, U^{-1}\beta)$. Thus $U^{-1} = U^*$.

($\Leftarrow$): Suppose $\exists U^* : V \to V$ s.t. $UU^* = U^*U = I$. Then U is invertible where $U^* = U^{-1}$. Then $(U\alpha, U\beta) = (\alpha, U^*U\beta) = (\alpha, \beta)$. □

Definition 8.4.4: Unitary

$A : n \times n$ mat. on $\mathbb{R}$ or $\mathbb{C}$. We say A is unitary if $AA^* = A^*A = I$.

Theorem 8.4.4

$U : V \to V$ on inn. prod. space. Then U is unitary $\iff$ $[U]_{\mathfrak{B}}$ for orthonormal basis $\mathfrak{B}$ is a unitary mat.

Proof. $[U]_{\mathfrak{B}}$ is unitary $\iff$ U is unitary. Then iff condition follows from Theorem 8.4.3. □

Corollary 8.4.1

If U_1 and U_2 are unitary, then U_1U_2 also. Furthermore, U_1^{-1} is also unitary.

Definition 8.4.5: Unitary Group - Optional

For f.d.inn. prod. space, let $U(V)$ be a collection of all unitary op. on V. This is a group, i.e., closed under mat. multiplication.

Note:-

OPTIONAL.

When $V=\mathbb{C}^n$, $U(\mathbb{C}^n)=U(n)$: the n-th unitary group.

$V=\mathbb{R}^n, A: n\times n$ mat. on $\mathbb{R}$ s.t. $AA^t=A^tA=I$. Then $O(n)$ is the real orthogonal group.

$V=\mathbb{C}^n, A: n\times n$ mat. on $\mathbb{C}$ s.t. $AA^t=A^tA=I$. Then $O(n,\mathbb{C})$ is the complex orthogonal group.

$SU(n)=\{A\in U(n)\,|\,\det(A)=1\}$ is special unitary group.

$SO(n)=\{A\in O(n)\,|\,\det(A)=1\}$ is special orthogonal group. For example, $SO(2)$ is rotation and $SO(3)$, with $SO(3)\rtimes\mathbb{R}^3$ is rigid motion.

8.5 Normal Operators

Definition 8.5.1: Normal

T : endo on f.d.inn. prod. space. V/F. We say T is normal if $TT^*=T^*T$.

Note:-

Q. When do we have an orthonormal basis $\mathfrak{B}$ on V s.t. vec. in $\mathfrak{B}$ are also char. vec. of T?

Theorem 8.5.1

T : endo on f.d.inn. prod. space. V/F. Suppose T is normal. For char. vec. α of T, $c\in F$ is char. value $\iff$ $\bar{c}$ is char. value for T^* with char. vec. α.

Proof.

Claim 8.5.1

If U is normal, then $||Uv||=||U^*v||$.

Proof. $||Uv||^2=(Uv,Uv)=(v,U^*Uv)=(v,UU^*v)=(U^*v,U^*v)=||U^*v||^2$. □

$\forall c\in F$, $U:=T-cI$ is normal for normal T. Then $U^*=T^*-\bar{c}I$. $UU^*=U^*U$ is obvious. Thus $||(T-cI)\alpha||=||(T^*-\bar{c}I)\alpha||$ by Claim 8.5.1. Thus $(T-cI)\alpha=0 \iff (T^*-\bar{c}I)\alpha=0$. □

Theorem 8.5.2

T as Theorem 8.5.1 but not normal. Suppose $\exists$ orthonormal basis $\mathfrak{B}$ s.t. $[T]_{\mathfrak{B}}$ is upper triangular. Then T is normal $\iff$ $[T]_{\mathfrak{B}}$ is diagonal.

Proof. ($\Leftarrow$): Let $A:=[T]_{\mathfrak{B}}$. $A^*=[T^*]_{\mathfrak{B}}$. A is diagonal, so A^* also. Trivially $AA^*=A^*A$, thus $TT^*=T^*T$, i.e., T is normal.

($\Rightarrow$): Suppose T is normal. We are given that A is upper triangular. Let $\mathfrak{B} = \{\alpha_1, \dots, \alpha_n\}$. Then

$$A = [T]_{\mathfrak{B}} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ 0 & \ddots & \vdots \\ 0 & 0 & a_{nn} \end{bmatrix}$$

where T is normal, and α_1 is char. vec., where a_{11} are char. value w.r.t. α_1. By Theorem 8.5.1, $T^*\alpha_1 = \overline{a_{11}}\alpha_1$. On the other hand, since $[T^*]_{\mathfrak{B}} = A^*$, $T^*\alpha_1 = \overline{a_{11}}\alpha_1 + \overline{a_{12}}\alpha_2 + \cdots + \overline{a_{1n}}\alpha_n$. Thus

$$A = [T]_{\mathfrak{B}} = \begin{bmatrix} a_{11} & \cdots & 0 \\ 0 & \ddots & \vdots \\ 0 & 0 & a_{nn} \end{bmatrix}.$$

Applying this algorithm to each α_i leads A is diagonal. □

Lemma 8.5.1

T : endo on f.d.inn. prod. space. $V/\mathbb{R}$ or $\mathbb{C}$. Let $W \subset V$ be T-inv. subspace. Then $W^\perp$ is automatically T^*-inv.

Proof. Let $\beta \in W^\perp$. N.T.S. $T^*\beta \in W^\perp$, i.e., $\forall \alpha \in W$ $((\alpha, T^*\beta) = (T\alpha, \beta) = 0)$. Since W is T-inv., this clearly holds. □

Theorem 8.5.3

T : endo on f.d.inn. prod. space. $V/\mathbb{C}$. Then $\exists$ orthonormal basis $\mathfrak{B}$ for V s.t. $[T]_{\mathfrak{B}}$ is upper triangular mat.

Proof. We prove it by induction on $n = \dim(V)$. If $n = 1$, it is obvious. So suppose $n > 1$ and assume Theorem 8.5.3 holds for any inn. prod. space with dim $< n$. Since $F = \mathbb{C}$, applying Fundamental Theorem of Algebra to T^*, $\exists$ char. value $c \in \mathbb{C}$, and a char. vec. α s.t. $T^*\alpha = c\alpha$. By replacing α to $\frac{\alpha}{\|\alpha\|}$, α itself has length 1. Define $W = \operatorname{span}\{\alpha\}^\perp$. Since $\operatorname{span}\{\alpha\}$ is T^*-inv, which leads $W = \operatorname{span}\{\alpha\}^\perp$ is T-inv. by the Lemma 8.5.1. Then we can see

$$\begin{array}{rccccl} T: & V & \longrightarrow & V & & \dim(V) = n \\ & \updownarrow & & \updownarrow & & \\ T|_W: & W & \longrightarrow & W & & \dim(W) = n-1 \end{array}$$

By induction hypothesis, $\exists$ orthonormal basis $\mathfrak{B}' = \{\alpha_1, \dots, \alpha_{n-1}\}$ s.t. $[T|_W]_{\mathfrak{B}'}$ is upper triangular. Take $\alpha_n := \alpha$, and $\mathfrak{B} = \mathfrak{B}' \cup \{\alpha_n\}$. Then

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

Thus $[T]_{\mathfrak{B}}$ is upper triangular. □

Corollary 8.5.1

T : endo on f.d.inn. prod. space. $V/\mathbb{C}$ where T is normal. Then V has orthonormal basis consisting of char. vec. of T. In particular, T is diagonalizable.

Corollary 8.5.2

With Theorem 8.5.2 and Theorem 8.5.3, if $A \in M_{n\times n}(\mathbb{C})$, $\exists$ unitary mat. $P \in U(n)$ s.t.

$P^{-1}AP$ is upper triangular. In case $AA^* = A^*A$, $P^{-1}AP$ is diagonal, i.e., A is normal implies A is unitary diagonalizable.

Example 8.5.1

T : endo on f.d.inn. prod. space. V/F. If T is hermitian, i.e., self-adjoint, then T is normal. Also, if T is unitary operator, it is normal.